QUELQUES CONTRIBUTIONS À LA RESTAURATION D'IMAGES ET DE SIGNAUX PAR DES MÉTHODES VARIATIONNELLES ET PAR SÉPARATION AVEUGLE DE SOURCE

Mohammed El Rhabi

Habilitation à Diriger les Recherches (HDR)

Spécialité mathématiques appliquées

Traitement du signal et des images

QUELQUES CONTRIBUTIONS À LA RESTAURATION D'IMAGES ET DE SIGNAUX PAR DES MÉTHODES VARIATIONNELLES ET PAR SÉPARATION AVEUGLE DE SOURCE

Mohammed EL RHABI

Année universitaire 2018/2019

Le 11/06/2019 devant le jury composé de :

M. Yannick Berthoumieu, professeur, IMB, Bordeaux INP	Rapporteur
M. Jérôme Mars, professeur, Gipsa Lab, Grenoble INP	Rapporteur
M. Simon Masnou, professeur, ICJ, Université Lyon 1	Rapporteur
M. Albert Cohen, professeur, LJLL, Sorbonne Université	Président
M. Laurent Dumas, professeur, LMV, Université de Versailles	Examinateur
M. Abdelilah Hakim, professeur, LAMAI, Université Cadi Ayyad	Examinateur
M. Djamal Merad, professeur, LIS, Aix-Marseille Université	Examinateur
M. Pascal Monasse, professeur, Imagine-LIGM, École des Ponts	Examinateur
M. Éric Moreau, professeur, LIS, Aix-Marseille Université	Tuteur de l'HDR

École nationale des ponts et chaussées
6 et 8 avenue Blaise-Pascal - F-77455 Marne-la-Vallée cedex 2
Tél. : + 33 (0)1 64 15 30 00 - Fax : + 33 (0)1 64 15 34 09 - www.enpc.fr

À la mémoire de mon père Abida,
Au regard de ma fille Camélia.

Remerciements

Je tiens tout d'abord à remercier Yannick Berthoumieu, Jérôme Mars et Simon Masnou d'avoir accepté de relire ce mémoire et d'en être rapporteurs. Je remercie également tous les autres membres du jury : Albert Cohen, Laurent Dumas, Abdelilah Hakim, Djamal Merad et Pascal Monasse d'avoir accepté d'assister à la présentation de ce travail.

Je voudrais remercier tout particulièrement Eric Moreau sans qui cette synthèse n'aurait pas été possible. Il a été un guide scientifique très important dans mes choix thématiques après la thèse . Il a aussi soigneusement relu ce mémoire ; j'ai pu alors bénéficier de ses nombreuses remarques (scientifiques et autres) et il a largement contribué à la clarté rédactionnelle du manuscrit et sans toi, ce mémoire n'aurait probablement jamais vu le jour !

Je voudrais également remercier Saïd Raghay, professeur à l'université Cadi Ayyad à Marrakech et directeur du laboratoire de mathématiques appliquées et d'informatique (LAMAI) pour sa confiance en m'accordant l'honneur d'être membre associé du laboratoire qu'il dirige et pour les nombreuses collaborations scientifiques (co-encadrement de stages, PFE, thèses, entre autres) sans oublier les événements scientifiques que l'on a organisés ensemble à Marrakech (école d'été, semaines d'initiation à la recherche et à l'innovation), ton dynamisme, tes idées et ton intelligence sont des moteurs pour qui collabore avec toi …

Abdelilah Hakim est sans doute le mathématicien avec lequel j'ai eu les collaborations les plus fructueuses, j'apprécie beaucoup notre complémentarité et sans doute j'ai encore beaucoup à apprendre à tes côtés. Merci pour tout !

Comment ne pas mentionner Tahar Boulmezaoud qui m'a initié à la recherche et m'a donné l'envie de poursuivre avec persévérance dans cette voie, des remerciements ne suffiraient pas …

Abdelilah, Eric et Tahar, vous êtes des exemples scientifiques que j'essaie de suivre … vais-je tenir le rythme ? Je l'espère sincèrement !

Alain Maruani, Abdelkader Slifi, vos relectures du premier jet et surtout vos encouragements m'ont été précieux … comment avoir des moments de découragement lorsque vous êtes là ?

Sandrine Guillerm, ma partenaire au département d'Ingénierie Mathématique et Informatique à l'École des Ponts, ce fut un plaisir de travailler avec toi, il est rare d'avoir la chance d'avoir une collaboratrice aussi professionnelle … Ton travail, ta rigueur et ta gentillesse m'ont énormément permis de progresser et j'ai beaucoup apprécié ton soutien constant lors de la rédaction de ce mémoire. Merci mon Amie !

Mes élèves, mes chers IMI, vous avez été une source d'inspiration, j'ai souvent été en admiration face à votre intelligence … donc merci à vous aussi.

Merci aussi au Cermics (laboratoire de mathématiques appliquées à l'École des Ponts), en particulier à Gabriel Stoltz, Tony Lelièvre et Eric Cancès pour vos encouragements.
L'équipe Imagine-LIGM, aka « mes copains des Ponts » : Renaud Marlet, Pascal Monasse, Guillaume Obozinski, j'ai adoré notre collaboration !

Je ne peux évidemment pas oublier mon épouse Françoise avec qui rien n'est impossible. Othmane, mon fils et Camélia, ma fille, vous grandissez … quand je pense que j'ai soutenu ma thèse deux mois après la naissance d'Othmane … A vous de prendre la relève.
J'aimerais aussi remercier ma grande famille, ma mère, mes frères, mes sœurs et leurs enfants. Youssef, c'est ton tour l'hdr ;-) … A la mémoire de mon père, je suis certain que là où tu es, tu dois sourire ….

Table des matières

1 Curriculum Vitæ

1.1 Parcours académique

1998-2002 : Thèse de doctorat à l'université Pierre & Marie Curie (mention *Très honorable* [1]) présentée et soutenue publiquement le 04 décembre 2002 au laboratoire Jacques-Louis LIONS.

Rapporteurs :

— Jean-Claude NÉDÉLEC (Professeur, CMAP, école polytechnique) ;
— Barbara WOHLMUTH (Professeur, université de Stuttgart, Allemagne).

Membres du jury :

— Christine BERNARDI (Directrice de Recherche CNRS, université Pierre & Marie Curie, LJLL), membre et directrice de la thèse ;
— Tahar Z. BOULMEZAOUD (maître de conférences, université de Versailles St Quentin en Yvelines, LMA), membre et co-directeur de la thèse ;
— Monique DAUGE (Directrice CNRS, université de Renne I), membre ;
— Georges DELAUNAY (Professeur, université de Reims Champagne-Ardenne, DéCOM-CReSTIC), membre ;
— Yvon MADAY (Professeur, université Pierre & Marie Curie, LJLL), membre et président du jury ;
— Jean-Claude NÉDÉLEC (Professeur, CMAP, École Polytechnique), membre et rapporteur ;
— Pierre-Arnaud RAVIART (Directeur de Recherche CNRS, université de Pierre & Marie Curie, LJLL), membre ;
— Barbara WOHLMUTH (Professeur, université de Stuttgart, Allemagne), membre et rapporteur.

1998 : DEA « EDP & Calcul Scientifique » au laboratoire d'analyse numérique & EDP, université d'Orsay Paris XI. Stage sous la direction de Jacques LAMINIE (Univ. d'Orsay, Paris XI) : « Sur le préconditionnement par approximations de l'inverse ».

1997 : Maîtrise Ingénieries mathématiques (M.I.M), Université de Reims Champagne-Ardenne. Mémoire de TER (Travail d'Étude et de Recherche) sous la direction de Michel BRUNEAU (Professeur), Laboratoire de Mathématiques à l'université de Reims : « Sur quelques problèmes de statistiques appliquées à la physique ».

1996 : Licence de mathématiques fondamentales, université de Reims Champagne-Ardenne.

1995 : DEUG A, Sciences et Structure de la matière (mathématiques).

1. L'UPMC recommande de ne plus attribuer la mention "très honorable avec les félicitations du jury", décision du Conseil Scientifique du 04/11/2001 et du 02/07/2001, et du Conseil d'Administration du 08/10/2001.

1.2 Parcours professionnel

Depuis 2016 : Maître de conférences, Ecole des Ponts ParisTech, directeur académique, département Ingénierie Mathématique et Informatique, École des Ponts ParisTech

2014-2017 : Responsable des partenariats avec le Maroc, École des Ponts ParisTech

2013-2016 : Responsable académique, département Ingénierie Mathématique et Informatique, École des Ponts ParisTech

Depuis 2010 : Chercheur associé, LAMAI, FSTG, université Cadi Ayyad (Marrakech, Maroc)

2010-2011 : Chargé de mission scientifique, comité du SIMI-1 mathématiques et interaction. Agence Nationale de la Recherche (ANR).

2008-2010 : Directeur scientifique, Realeyes3d SA.

2006-2008 : Chargé de recherche, Realeyes3d SA, profil : mathématiques appliquées au traitement d'images.

2005-2006 : Chargé de recherche (modélisation mathématique) au Centre d'Enseignement et de Recherche en Mathématiques, Informatique et Calcul Scientifique (CERMICS), École Nationale des Ponts et Chaussées en partenariat avec l'École des Mines de Paris.

déc. 2002-2005 : Contrat post-doctoral (statistiques et traitement du signal) au CReSTIC (Centre de Recherche en STIC) & Enseignant à l'université de Reims Champagne-Ardenne, niveau : DEUG, Master.

1999-2001 : Attaché Temporaire d'Enseignement et de Recherche (ATER) à l'université Paris 12, Créteil, niveau : Licence, Master.

1997-1999 : Enseignant vacataire à l'université Paris-Sud, faculté Jean Monnet, niveau Licence.

1.3 Activités de recherche

Thèses co-encadrées et soutenues :

2013-2016 : Abdelghani Ghazdali, Laboratoire de Mathématiques Appliquées et Informatique, Faculté des Sciences et Techniques (FST), université Cadi Ayyad (Marrakech, Maroc, LAMAI-FST-UCA). Soutenue publiquement le 12 décembre 2015. Sujet « Séparation aveugle de sources indépendantes / dépendantes par minimisation de critères entre copules ». Professeur assistant titulaire (eq. français : maître de conférences) à l'ENSA (eq. français : INSA) de Khouribga (Maroc) depuis septembre 2016. Directeurs de thèse : Abdelilah Hakim, professeur à l'université Cadi Ayyad (Marrakech, Maroc), Amor Kéziou, maître de conférences, université de Reims Champagne-Ardenne, Hassan Fenniri, maître de conférences, université de Reims Champagne-Ardenne et Mohammed El Rhabi. Publications communes : [7],

2013-2016 : Amine Laghrib, LAMAI-FST-UCA. Soutenue publiquement le 19 décembre 2015. Sujet « La super résolution d'une séquence d'images par modèles variationnels - Recalage et restauration ». Professeur assistant titulaire (eq. français : maître de conférences) à l'université Sultan Moulay Slimane

(Béni Mellal, Maroc) depuis septembre 2016. Publications communes : [3, 5, 10, 11, 9]. Directeurs de thèse : Abdelilah Hakim, Saïd Raghay, professeur à l'université Cadi Ayyad (Marrakech, Maroc) et Mohammed El Rhabi,

2014-2017 : Idriss El Mourabit, LAMAI-FST-UCA. Sujet « Problèmes variationnels et EDP appliqués au traitement d'images ». Soutenue publiquement le 30 octobre 2017. Publications communes : [6, 8]. Directeurs de thèse : Abdelilah Hakim et Mohammed El Rhabi.

Thèses co-encadrées en cours :

2016-présent : Fatim Zehrae Aït Bella, deuxième année de thèse, LAMAI-FST-UCA. Sujet « restauration d'images de documents acquises dans des conditions difficiles (illumination variable, flou, bruit ....) ». Publication commune : [3]. Directeurs de thèse : Abdelilah Hakim, Amine Laghrib, professeur assistant à l'université Sultan Moulay Slimane (Béni Mellal, Maroc) et Mohammed El Rhabi,

2014-présent : Abdelghani Elmoussaoui, troisième année de thèse, LAMAI-FST-UCA. Sujet « Théorie cinétique discrète et application aux mouvements de foule ». Publications communes : [2, 4, EAEH]. Directeurs de thèse : Abdelilah Hakim, Pierre Argoul, directeur de recherche à l'Ifsttar (Champs sur Marne) et Mohammed El Rhabi,

2014-présent : Aïssam Jebrane, troisième année de thèse, LAMAI-FST-UCA-LVMT-École des Ponts. Thèse de doctorat en co-tutelle UCA-École des Ponts. Sujet « Modèles discrets de foule en mouvement. Foules denses et situation de panique. Systèmes Différentiels Stochastiques ». Publications communes : [1, AAHE, JAHE17]. Directeurs de thèse : Abdelilah Hakim, Pierre Argoul, directeur de recherche à l'Ifsttar (Champs sur Marne) et Mohammed El Rhabi.

Production depuis la thèse : 15 articles de revues référencées JCR ou Thomson-Reuters, 3 articles soumis dans des revues référencées, 11 conférences internationales avec comité de lecture dont 9 actes de conférences,

Projets de recherche : 1 projet Toubkal (franco-marocain), 1 projet de partenariat international « recherche et enseignement » (École des Ponts ParisTech, le groupe OCP (Office Chérifien des Phosphates, plus grand groupe marocain) et l'université Mohammed 6 Polytechnique-Maroc). Des projets industriels (startup). 1 projet Toubkal soumis (optimisation et réduction de dimension sur des problèmes de traitement de signaux multi-dimensionnels ou d'images en grande dimension, voir section 8.2),

Brevets et transfert de technologie : 2 brevets [ER08, EMHS14].

2 Autres activités

2.1 Organisation d'événements

Septembre 2015, 2016, 2017 et 2018 : Organisation des semaines « Initiation à la recherche et à l'innovation I, II, III et IV » (Casablanca, Marrakech et Palma de Majorque) : pendant une semaine, des élèves en deuxième année au département

IMI à l'École des Ponts ParisTech travaillent avec des étudiants en M2 d'une université partenaire autour de projets de recherche proposés par des chercheurs à l'École des Ponts ParisTech (ou des industriels partenaires). La première session a été organisée à Casablanca en collaboration avec l'École Hassania des travaux Public (EHTP, TOP 2 des écoles d'ingénieurs marocaines) et l'université Cadi Ayyad (Marrakech, Maroc), la deuxième à Marrakech avec le même partenariat, la troisième et la quatrième avec l'université des Îles Baléares [2],

28 au 31 mai 2012 : organisation de l'école « Mathématiques & Interactions II », proposée par les universités Cadi Ayyad (Marrakech, Maroc), Ibn Zohr (Agadir, Maroc), de Reims Champagne-Ardenne (France) [3],

16 au 19 mai 2011 : organisation de l'école « Mathématiques & Interactions I », proposée par les universités Ibn Zohr, de Toulon, de Versailles Saint-Quentin en Yvelines, de Reims Champagne-Ardenne (France) [4],

Jeudi 26 Novembre 2009 : organisation de la demi-journée industrielles « Traitement de l'image et application industrielles », proposée par les universités de Versailles Saint-Quentin en Yvelines (laboratoire de mathématiques de Versailles, LMV), Pierre et Marie Curie (laboratoire Jacques-Louis Lions, LJLL) et la startup Realeyes3d, SA,

2005-2006 : élaboration et administration de la page web de l'équipe « EDP & Matériaux » du laboratoire de mathématiques appliquées à l'École des Ponts ParisTech (Cermics),

1999-2002 : en tant que doctorant, j'ai eu la charge du secrétariat du groupe de travail hebdomadaire du laboratoire JACQUES-LOUIS LIONS, université Pierre et Marie Curie : « Groupe de Travail Méthodes Numériques » animé par Yves ACHDOU, Christine BERNARDI et Yvon MADAY.

Mars 2000 : animation du stand de l'I.U.T SÉNAT-FONTAINEBLEAU au salon de l'Étudiant.

2.2 Coopération internationale

De 2014 à 2017, j'ai pris en charge le développement des partenariats entre l'École des Ponts ParisTech et le Maroc. Tout d'abord, j'ai élaboré un nouveau processus de recrutement des élèves ingénieurs à l'École Hassania des Travaux Publics (EHTP, TOP 2 des Grandes Écoles d'ingénieurs marocaines). Ce recrutement est effectué dans le cadre d'une convention établie entre l'EHTP et l'École des Ponts ParisTech. Cette convention prévoit le recrutement de 5 élèves en « double diplôme » (format 2+2) et 5 élèves en échange académique (l'élève effectue sa 3 ème année à l'École des Ponts ParisTech).

En 2015, avec l'aide de mon réseau marocain, j'ai pu initier une collaboration avec l'École Mohammadia des Ingénieurs (EMI,TOP 2 des des « Grandes Écoles » d'ingénieurs marocaines). Il s'en est suivi une nouvelle convention signée par l'École des

2. pour voir le programme de cette année http://imi2018.enpc.fr
3. http://mathinfo.univ-reims.fr/EPMI2012/
4. http://mathinfo.univ-reims.fr/EEMIA/

Ponts ParisTech et l'EMI. Cette convention prévoit les mêmes termes que celle signée avec l'EHTP.

Enfin, l'École m'a confié le « démarrage » du projet de coopération entre l'École des Ponts ParisTech, le groupe OCP (Office Chérifien des Phosphates, le plus grand groupe marocain) et l'université Polytechnique Mohammed 6 (UM6P, une nouvelle université privée marocaine). Ce partenariat a été signé en juillet 2017. Il prévoit une collaboration « recherche & formation par la recherche » (une « graduate school » !) pour un montant d'1 million d'euros par an sur 5 ans (renouvelable). Actuellement, je suis chargé d'élaborer, avec nos partenaires, la construction d'un parcours de master mention mathématiques appliquées, 3 majeures sont visées, à savoir « la modélisation numérique », « l'optimisation, la recherche opérationnelle » et « les data sciences ». Nous espérons pouvoir ouvrir ce parcours aux étudiants en septembre 2019.

3 Activités d'enseignements

Le contenu de mes enseignements a évolué tout au long de ma carrière (public, privé). Mes activités d'enseignements concernent les mathématiques appliquées, le traitement du signal, le traitement d'images et plus récemment le management de projets de recherche visant l'innovation. Mes cours s'adressent à un large spectre d'étudiants, inscrits en licence, en master, ainsi qu'à des élèves ingénieurs.

3.1 Enseignements

Mes activités d'enseignement ont débuté à la faculté Jean Monnet de l'Université Paris XI-Orsay (FJM) où j'ai dispensé des travaux dirigés d'analyse et d'algèbre linéaire de première année de DEUG sciences économiques et sociales correspondant à un volume horaire de 96 heures (équiv. TD). Puis, l'année suivante, il m'a été confié la charge des travaux dirigés d'analyse et d'algèbre de première et deuxième année dans cette même section (96 heures).

Par la suite, j'ai été recruté en tant qu'Attaché Temporaire d'Enseignement et de Recherche (ATER) à l'I.U.T de Sénart-Fontainebleau au département d'Informatique à l'université Paris-Est Créteil (UPEC). Dans le cadre de mon contrat, j'ai enseigné de l'analyse et de l'algèbre à l'I.U.T de Sénart-Fontainebleau au département d'Informatique (Université Paris XII Créteil-Val de Marne). Ces enseignements se sont répartis sous la forme de travaux dirigés et de travaux pratiques en collaboration avec O. LOUKIANOV.

Les deux thèmes principaux du programme abordés sont présentés ci-dessous :

Bases mathématiques pour l'informatique théorique : L'objectif est de donner les notions de base permettant de décrire avec précision certains fondements de l'Informatique. A cet effet, on a exercé l'étudiant à la modélisation et on a montré comment certaines notions peuvent apporter une aide à la structuration de données, à la conception d'algorithmes. Cette partie se décompose comme suit :

— Vocabulaire de la théorie des ensembles. Nous avons veillé à expliciter et justifier les notions nécessaires de la théorie des ensembles (opérations binaires, produit cartésien, relations binaires entre autres).
— Logique - calcul propositionnel, tables de vérité (point de vue sémantique), introduction aux systèmes formels (point de vue syntaxique). Nous avons mis en évidence à partir d'exemples le lien entre le formalisme logique et l'écriture de démonstrations à l'aide, par exemple, du calcul booléen, de règles d'inférence. Initiation aux algorithmes de démonstration.
— Ensembles ordonnés - relation d'ordre, ordre total, ordre partiel. Nous avons donné à ce sujet les définitions de "plus grand élément", de "majorant", entre autres, des définitions et exemples de treillis, le treillis des parties d'un ensemble. Nous avons présenté des situations concrètes où intervient un formalisation à l'aide d'ensembles ordonnés, par exemple en traitement d'images (érosion, dilatation), entre autres.

Bases mathématiques pour l'informatique appliquée :

— Arithmétique, exemples de raisonnement par récurrence, division euclidienne, numération, nombres premiers, ppcm, pgcd, identité de Bézout, algorithme d'Euclide, calcul des coefficients de Bézout, congruences, $\mathbb{Z}/n\mathbb{Z}$.
Des résultats et théorèmes (Euler, th. chinois, polynômes sur $\mathbb{Z}/n\mathbb{Z}$, etc...) ont été présentés en fonction des choix faits sur les applications (cryptographie, codes détecteurs d'erreurs, résolutions d'équations en nombres entiers, ...).
— Algèbre linéaire, calcul matriciel, systèmes d'équations linéaires : aspects conceptuels et numériques (méthode du pivot de Gauss, ...), espaces vectoriels : exemples, indépendance linéaire, dimension, sous espace vectoriel ; on s'est limité aux espaces de dimension finie, applications linéaires ; nous avons succinctement abordé les notions de valeur propre et de vecteur propre, géométrie : produit scalaire, produit vectoriel, symétries, rotations, changement de repère, projections.
— Analyse, suites numériques, exemples de séries numériques ; les notations O(n) et o(n) ont été introduites (en vue de l'étude de la complexité d'algorithmes), fonctions réelles : limites, dérivation, intégration, formule de Taylor, fonctions réciproques. Ces notions ont été accompagnées d'exemples et de contre-exemples montrant la nécessité de certains développements théoriques, majorer, minorer, encadrer ; ces principes de gestion des approximations ont été mis en $\frac{1}{2}$uvre à l'occasion de calculs numériques sur les fonctions (valeurs, extrema, zéros, limites, intégration, dérivation, interpolation, etc ...). Des éléments de fonctions de plusieurs variables réelles ont été présentés en vue de problèmes d'optimisation.
— Décomposition des fractions rationnelles en éléments simples, compléments d'intégration (limités aux outils indispensables, études de problèmes de convergence (sur des exemples), convolution, séries de Fourier (étude en moyenne quadratique), transformées de Fourier et de Laplace (applications aux problèmes différentiels), transformée de Fourier discrète, principe de l'algorithme rapide (FFT).

Lors du renouvellement de mon poste, j'ai eu la charge des enseignements en

DEUG SV en collaboration avec J.P GEORGIN, maître de conférences à l'université Paris-Est-Créteil. Dans ces cours, nous avons abordé : la logique, des tables aux principaux raisonnements mathématiques. J'ai aussi pu collaborer avec G. ALLAIN, maître de conférences à l'université Paris-Est-Créteil, dans le cadre des enseignements en DEUG SM, deuxième année, dont voici les mots clefs du programme : en analyse : Suites réelles (convergence monotone et suites adjacentes). Intégrales généralisées (limité à la convergence absolue). Séries numériques ; séries entières : développement en séries entières, application aux équations différentielles ; et en algèbre : Déterminants. Systèmes linéaires.Diagonalisation de matrices.

Il m'a aussi été confié le cours et les travaux dirigés de l'unité d'enseignement « analyse numérique » en maîtrise de Physique. Cette unité d'enseignement introduit les principales méthodes numériques et leur mise en œuvre informatique. Nous avons abordé l'interpolation polynomiale, l'intégration numérique et la mise en œuvre de méthodes de discrétisation d'une EDO par différentes méthodes (différences finies, Runge Kutta entre autres). Puis une analyse de l'erreur a été abordée succinctement. Enfin des algorithmes numériques ont été proposés. Ceux-ci furent validés en T.P et par un projet encadré par S. NAILI, professeur à l'université Paris-Est Créteil. Le langage de programmation fut le Fortran 77. Cette unité d'enseignement fut santionnée par un examen que j'ai proposé.

Par ailleurs, j'ai participé activement à l'élaboration du cours, des feuilles de travaux dirigés et j'ai proposé des devoirs surveillés. J'ai notamment préparé les examens avec les responsables respectifs.

Durant mon stage postdoctoral au CReSTIC à l'université de Reims Champagne-Ardenne (URCA), il m'a été confié le cours et les travaux dirigés de l'option traitement du signal en DESS « Méthodes statistiques pour la finance et l'industrie » et les travaux pratiques de l'option « Méthodes et outils informatiques » en DEUG (MIAS, SM, SV).

Depuis 2008, j'ai été régulièrement invité à enseigner (cours de Master 2) les EDP appliquées au traitement d'images et des cours de traitement du signal à la faculté de sciences et techniques à l'université Cadi Ayyad (FST-UCA, Marrakech, Maroc), par le Professeur de l'enseignement supérieur [5] ABDELILAH HAKIM. Ces cours ont été ouverts aussi bien aux étudiants de Master 2 qu'aux enseignants-chercheurs et ingénieurs. Ces cours (intensifs) se sont déroulés sur une période d'une semaine chacun.

J'ai aussi enseigné l'option « EDP, traitement d'images et applications industrielles » dans le cadre du Master 2 « Modélisation Mathématique pour les Sciences de l'Ingénieur » (MMSI) proposé à l'université de Reims Champagne-Ardenne par le département de mathématiques.

Enfin depuis 2013, je m'occupe des projets de département pour le département du cycle master « Ingénierie Mathématique et Informatique » (IMI). Les projets de département sont souvent des problèmes que rencontrent des industriels partenaires d'IMI. Après un dialogue avec nos partenaires, nous établissons des sujets. Ces sujets laissent volontairement une certaine marge afin que les étudiants ayant choisi d'y travailler puissent l'affiner avec nous.

L'idée, ici, est de laisser une grande marge à la créativité des élèves, tout en restant

5. Le plus haut grade au niveau de l'enseignement supérieur marocain

« raisonnable » par la démarche scientifique !

Ainsi, une fois les sujets définis, les élèves se répartissent par groupe de 3 à 6 élèves et ils commencent à réfléchir à une solution au début du second semestre de leur deuxième année à l'École des Ponts ParisTech (niveau M1). Les élèves disposent d'une demi-journée par semaine pendant un semestre (4 mois). Les chercheurs ou partenaires sont des « consultants » : une fois les objectifs définis avec eux, les élèves travaillent en autonomie, ils les « consultent » en cas de besoin. Je fais néanmoins des points réguliers avec eux, avec une présentation à mi-parcours avec eux.

A la fin de leur projet, chaque groupe d'élèves est évalué deux fois. Un mois avant l'échéance du projet, un premier jury, composé d'un chercheur en économie, d'un dirigeant d'entreprise (idéalement une startup), un spécialiste des projets numériques et un spécialiste de la communication, se réunit. Ce jury évalue la capacité à valoriser leur réponse (viabilité économique, recul sur leur solution, les potentiels de l'équipe et éventuellement le retour sur investissement du projet). Puis, un second jury, composé de chercheurs ou ingénieurs seniors, évalue la démarche scientifique.

Le tableau suivant récapitule les différents enseignements que j'ai dispensés :

Année	Intitulé	Public	Lieu	Élaboration des sujets	Heures (en équivalent TD)
2017-2018	Mathematics for business and startup : in the shoes a CTO	M2	École des Ponts	100%	60
2013-2018	Projets de département	M1-M2	École des Ponts	100%	250
2008-2018	EDP, Traitement des images Traitement du signal	M2	FST-UCA	100%	200 100
2008-2011	EDP, traitement d'images et applications industrielles	M2	URCA	100%	66
2002-2005	Traitement du signal (Cours, TD,TP) Méthodes et outils informatiques	DESS DEUG MIAS, SM, SV	URCA	100%	90 108
2000-2001	Logique analyse et algèbre analyse numérique	DEUG SV DEUG SM Maîtrise de physique	UPEC	75%	32 32 32
1999-2000	Analyse et algèbre { cours TD TP	IUT 2nde année	IUT Sénart	100%	48 26 22
1997-1999	Analyse et algèbre	DEUG Sciences éco. Licence en économie	FJM	50%	96

3.2 Nouvel enseignement

Le cours, intitulé « Mathematics for business and startup : in the shoes of a CTO », a été conçu avec Abdelkader Slifi, maître de conférences à l'université de Paris-Dauphine.

L'idée principale de ce cours est de donner une idée du métier de directeur technique (CTO) aux élèves en dernière année a' l'École des Ponts ParisTech. Nous avons choisi l'approche par l'exemple et à partir de ma propre expérience en startup. Il a démarré en septembre 2017. Ce cours est proposé en partenariat avec l'incubateur de Télécom-Paris et ses startups.

3.2.1 Objectifs de ce module

— Comprendre l'environnement de concurrence imparfaite, l'univers incertain, l'organisation managériale, les contraintes technologiques et financières de la startup ; acquérir la culture du MVP (Minimum Viable Product)

— Mobiliser et appliquer des connaissances technologiques et des outils d'analyse pour résoudre un problème opérationnel en startup
— Développer des procédés innovants qui répondent à la fois aux critères de la fiabilité technique et de la viabilité économique. Découvrir le prototypage rapide.

3.2.2 Programme du module

L'idée fondamentale est de placer l'élève dans la situation d'un CTO (directeur technique), il a diverses sollicitations en « Recherche & Développement » (R&D) et doit faire des choix.

— Le client ne sait pas faire : mais la solution existe,
— Le client ne sait pas faire : la solution peut exister mais il existe des verrous techniques, (ce que l'état de l'art ne sait pas faire). Quelle stratégie adopter car il y a de l'IP potentielle (brevet) ?
— Pas de client mais une idée qui a un marché potentiel : il se pose alors la question des verrous, du coût, du moment où il faut s'arrêter en l'absence de solution stable.

Maintenant comment faire ? Avec l'aide de startups et sur des problèmes que l'on sait résoudre et qui ont un verrou (ou pas), les élèves sont mis au défi de bien choisir leur projet. Ce cours s'adresse à des élèves de 3 ème année à l'École des Ponts ParisTech dotés de prérequis en analyse numérique et en informatique, intéressés par le traitement du signal, les techniques de développement de logiciel, et motivés par l'innovation pour l'activité économique et la startup. Dans ce cours, on se place du point de vue de la R&D de la startup, ayant le souci de l'usage et de la viabilité économique des projets. Il est ouvert au maximum à 20 élèves (dossier et motivations). Les élèves choisissent parmi les startups intéressées. Une fois les groupes formés, nous nous entourons d'un pool de consultants. Ce pool évolue en fonction des difficultés rencontrées lors des différents projets. Ce pool est formé de chercheurs à l'École des Ponts ParisTech ou d'industriels partenaires au profil.

3.3 Responsabilités pédagogiques

Depuis 2013, je m'occupe du département Ingénierie Mathématique et Informatique à l'École des Ponts ParisTech. Je suis responsable des activités d'enseignements et de leur cohérence avec les cahiers des charges fixés par l'École des Ponts ParisTech. Je participe de manière générale à la politique de la direction de l'enseignement à l'École des Ponts ParisTech (nouveaux enseignements, recrutements des étudiants étrangers, construction de partenariats à l'international, principalement).

Une inspectrice des études et une assistante pédagogique sont rattachées à mon service. Elles ont pour mission principale l'organisation opérationnelle du suivi de mon département (suivi de la scolarité des élèves, alerte sur les situations atypiques, entre autres).

Je définis avec le président (un industriel d'un grand groupe, chargé de représenter les métiers liés au département) la politique et la stratégie pédagogique du département. Je suis chargé du pilotage, de la mise en œuvre et de l'évaluation de l'offre de

formation. Cette évaluation est effectuée avec le concours de spécialistes de l'ingénierie pédagogique à l'École des Ponts ParisTech. En ce qui concerne les nouveaux cours, je sollicite un enseignant au profil afin d'établir le cahier des charges, ses objectifs et son mode d'évaluation. Un appel à candidatures est ensuite envoyé sur les sites dédiés par l'inspectrice des études ou l'assistante pédagogique. J'organise le recrutement de l'enseignant en programmant des auditions, après avoir défini un jury.

En outre, j'anime et je coordonne les équipes enseignantes avec l'aide de l'inspecteur des études. Je présente l'offre de cours et je supervise le suivi des élèves. Sur rendez-vous, j'assure un « office hours » pendant lequel j'apporte un cadre et des conseils aux élèves (projet professionnel, choix de cours et de stage). Pour les élèves s'interrogeant sur la pertinence d'effectuer une thèse de doctorat à l'issue de leur formation d'ingénieur, je planifie des rencontres progressives afin de les aider à clarifier leur projet professionnel, au bénéfice d'une meilleure visibilité du domaine de la recherche visé et des métiers associés. Par ailleurs, j'introduis dès la deuxième année des projets orientés vers la recherche en mathématiques appliquées ou en informatique, afin de sensibiliser davantage les élèves ingénieurs (semaine pédagogique « initiation à la recherche et à l'innovation », projets de département, voir les sections 2.1, 3.1 et 3.2).

J'ai la responsabilité de la validation de l'ensemble des stages du cycle master en mathématiques appliquées et informatique à l'École des Ponts ParisTech (en fonction du sujet, je sollicite éventuellement un enseignant).

Je suis aussi chargé du développement d'actions partenariales avec les autres départements du cycle master à l'École des Ponts, ainsi qu'avec des industriels.

Le département des relations internationales me sollicite régulièrement pour la construction de partenariats académiques internationaux (programme d'échanges ou de double diplômes, entre autres).

J'assure la gestion de l'enveloppe budgétaire de mon département.

4 Travaux de thèse

Mon travail doctoral, préparé sous la direction de CHRISTINE BERNARDI, directrice de recherche CNRS, a eu pour objet l'*analyse numérique et discrétisation par éléments spectraux avec joints des équations tridimensionnelles de l'électromagnétisme*. Ces travaux ont débuté par une étude théorique approfondie, suivie d'une analyse numérique non classique et d'une discrétisation tridimensionnelle. Ces résultats ont été confirmés et illustrés par des codes de calcul 3D que j'ai développés *ex nihilo* en Fortran.

4.1 Analyse du modèle en géométrie complexe

Nous avons donc commencé par une étude théorique du problème en considérant les équations de Maxwell (en régime harmonique) dans un domaine à géométrie complexe, à savoir un domaine borné multiplement connexe et ayant une régularité minimale. Nous nous sommes alors intéressés aux questions d'existence et d'unicité à travers une *nouvelle approche* théorique pour le traitement des équations de Maxwell en régime harmonique. Cette approche est basée sur la définition d'un nouvel

opérateur permettant de traiter le problème d'une manière unifiée, tenant compte des complexités géométriques du domaine et de son manque de régularité. En outre, il nous a permis d'établir un résultat d'existence global (obtenu à l'aide d'une alternative de Fredholm). De surcroît, nous avons montré que les fréquences de résonance du problème de Maxwell sont les *valeurs singulières* de cet opérateur. Enfin, nous avons proposé une formulation faible (dans le cas des basses fréquences) en introduisant une forme bilinéaire adaptée sans contraintes [TE03, 14].

Or, lorsque l'on s'intéresse à la discrétisation de ces équations, de sérieux problèmes techniques peuvent apparaître si l'on considère le caractère vectoriel (trop coûteux si l'on veut la même précision dans tous le domaine) ou si les paramètres physiques sont variables (domaine hétérogène). Ces raisons nous ont conduit naturellement à une décomposition de domaine.

4.1.1 Analyse numérique du problème : le cas conforme

Tout d'abord en nous plaçant dans un domaine (homogène) plus régulier ou un parallélépipède, nous avons donné une approximation de la solution par une méthode spectrale. Après avoir établi un résultat de régularité, nous avons developpé d'une part une analyse numérique du problème dans le cas d'une décomposition conforme du domaine et d'autre part la solution numérique calculée par un code $3D$ spectral écrit en Fortran 77. Les tests numériques ont confirmé les prévisions théoriques puis ont été comparés à ceux obtenus par une méthode d'éléments finis (P^1). Les quatre premières valeurs propres du problème discret ont été aussi calculées pour les deux méthodes puis comparées à celles du spectre exact [14].

4.1.2 Analyse numérique du problème : le cas non conforme

Cependant, si l'on veut considérer un domaine hétérogène ou si l'on veut choisir des paramètres de discrétisations différents dans chacun des sous-domaines ou même utiliser des méthodes de discrétisation différentes, une décomposition non conforme semble plus adéquate. Ce dernier choix induit de sérieuses complications à la fois pour l'analyse numérique du problème mais aussi dans son implantation. Ainsi la définition du repère pour la formulation du problème initial et l'imposition de conditions de transmission, concernant la continuité des composantes tangentielles ou normales des champs aux interfaces communes entre sous-domaines adjacents, ont joué un rôle crucial dans le choix de la méthode numérique à appliquer.

Au final, le passage en revue des méthodes existantes pour la prise en compte du raccordement aux interfaces de la décomposition avait montré que l'approche du problème et l'imposition « ponctuelle » des conditions de transmission ne sont satisfaisantes ni du point de vue formel ni du point de vue numérique. En revanche, la méthode des éléments avec joints appliquée à une formulation du problème initial s'est révélée plus adaptée pour la modélisation du problème. Elle a permis de traiter élégamment des discrétisations spatiales non conformes grâce au caractère « faible » avec lequel les conditions de transmission sont imposées aux interfaces entre sous-domaines. C'est pourquoi, il nous avait paru judicieux d'appliquer cette méthode discrétisée par des éléments spectraux au problème de Maxwell.

Par ailleurs, si l'on pose le problème d'évolution et le problème en régime harmonique, on se rend vite compte que les deux problèmes ne se formulent pas convenablement dans le même espace (une fois que le problème d'évolution est ramené à un problème stationnaire par une discrétisation en temps). J'avais donc introduit une approche commune (par un schéma de relaxation pour le problème en régime harmonique) permettant ainsi de résoudre les deux problèmes dans le même espace. Seules les conditions initiales et les paramètres (de discrétisation en temps pour le problème d'évolution et le nombre d'onde pour le problème en régime harmonique) changent. On se ramène alors à la résolution d'un problème-type, auquel la méthode spectrale avec joints a été appliquée. Par la suite, nous avons établi une analyse numérique où une estimation *a priori* de l'erreur (pour des paramètres physiques variables est donnée. Cette estimation originale a necessité l'introduction de différentes étapes ainsi que la définition de plusieurs opérateurs discrets (voir [17] pour plus de détails). Enfin, un code $3D$ spectral avec joint élaboré en Fortran 77, conclut cette étude. Ces résultats prouvent d'une part la précision de la méthode et d'autre part, sa flexibilité à modéliser les différents phénomènes physiques. En outre, les codes de calcul ont été testés avec des paramètres physiques constants par morceaux (ce qui numériquement constituait, à ma connaissance, une innovation intéressante).

Ces travaux ont fait l'objet de publications dans des revues internationales, dans les actes d'une conférence internationale et de conférences nationales (toutes bien entendu avec comité de lecture) [14, TE03, 17, TE02, TE01]. Une version papier, ainsi que quelques résultats numériques obtenus à partir des codes $3D$ développés illustrant ses performances, sont disponibles sur ma page web [6].

Enfin, dans l'article [15], nous avons généralisé l'étude proposée au cours de ma thèse. En fait, nous avons étendu cette méthode à certains problèmes variationnels formulés dans des espaces de champs de vecteurs de carré intégrable, à rotationnel de carré intégrable. Nous avons considéré le problème consistant à calculer le potentiel vecteur associé à une fonction à divergence nulle en dimension 3 et nous en avons proposé une discrétisation. Nous avons effectué l'analyse numérique du problème discret et nous avons présenté des résultats numériques cohérents avec ceux de l'analyse.

Ces travaux de doctorat ont fait l'objet des publications suivantes :

— articles de revues : [14, 17, 15],
— acte de conférence internationale : [TE03],
— conférences nationales : [TE02, TE01].

5 Travaux post thèse

Après ma thèse de doctorat, je me suis intéressé à deux grandes thématiques du traitement du signal et d'images, respectivement la séparation aveugle de source et la restauration d'images et plus récemment à la dynamique des foules (cette dernière thématique ne sera pas détaillée dans ce document). Mes travaux en traitement du signal ont été effectués en collaboration avec les universités de Reims Champagne-Ardenne, de Toulon et l'université Cadi Ayyad (Marrakech, Maroc). Ceux en traitement d'images avec la startup Realeyes3d, l'université de Toulon et l'université Cadi

6. http://elrhabi.free.fr/

Ayyad à Marrakech au Maroc. Ces travaux ont conduit à co-encadrer trois thèses soutenues en traitement d'images et du signal et une thèse en cours. Je co-encadre également deux thèses sur le thématique de la dynamique des foules.

5.1 Recherche en traitement du signal et d'images à l'École des Ponts ParisTech

Dans cette section, sont présentés mes travaux de recherche à l'École des Ponts ParisTech. Ces travaux se sont principalement focalisés sur la restauration d'images (traitement d'images) et la séparation aveugle de source (traitement du signal) et j'ai aussi abordé une nouvelle thématique, à savoir la dynamique des foules (voir section 7, non détaillé dans ce document) . Ces travaux sont bien évidemment dans la continuité de mes travaux antérieurs en traitement du signal et en traitement d'images.

Tout d'abord, nous avons proposé une nouvelle stratégie permettant de tenir compte du bruit dans le modèle classique de la séparation aveugle de sources. Pour résumer, la séparation aveugle de source consiste en l'extraction de signaux sources (inconnus !) à partir de la seule connaissance de leurs « mélanges » (lors de l'acquisition, les signaux se sont combinés linéairement). Ces mélanges sont définis par l'action d'un opérateur de mélange (lui aussi inconnu !) sur les sources, d'où le caractère *aveugle* de la séparation. Ce problème est bien entendu fortement mal posé. Il nécessite donc certaines hypothèses afin de le régulariser.

Sous les hypothèses de linéarité et d'inversibilité (au moins à gauche) de l'opérateur de mélange et d'indépendance statistique des signaux sources (avec au plus un d'entre eux gaussien), cette séparation revient alors à rendre statistiquement indépendantes les observations et il est assez simple de reposer/reformuler cette méthode en termes d'un problème d'optimisation stochastique sous contraintes.

Très souvent dans la littérature associée à cette méthode, le bruit dû aux capteurs de mesure est négligé. Dans notre méthode, nous tenons compte d'un bruit additif gaussien dans le modèle et nous proposons la minimisation d'un critère de séparation tenant compte de la variation totale du signal. Les résultats préliminaires ont été présentés lors de la conférence The 11th IEEE International Workshop on SPAWC (voir [BEFM10]). Puis les résultats définitifs ont été publiés dans [13].

Motivé par ce résultat et mes travaux en traitement d'images, j'ai initié un projet synthétisant ces deux thématiques (restauration d'images et séparation aveugle de source). Nous l'avons soumis, en 2011, dans le cadre de l'appel d'offre franco-marocain Toubkal (Campus France-France, ex Volubilis-EGIDE)[7].

5.2 Projet Toubkal (ex PHC-Volubilis)

Ce projet d'une durée de 3 ans a été accepté en octobre 2011 par le comité mixte inter-universitaire Franco-Marocain. Son objectif était de développer de nouvelles méthodes pour restaurer des images très dégradées de textes afin de pouvoir les reconnaître par un logiciel de reconnaissance de caractères « OCR » (Optical Character Re-

7. https://www.campusfrance.org/fr/toubkal

cognition)[8].

Ce projet s'appuie sur trois domaines de recherche distincts, particulièrement bien maîtrisés par les participants de ce projet :

— Traitement des images par Équations aux Dérivées Partielles (EDP) : intervenants : Simon Masnou (responsable scientifique, Pr, ICJ, université Claude Bernard, Lyon 1), Jérome Pousin (Pr, ICJ, Insa de Lyon), Frank Lebourgeois (MCF,LIRIS, Insa de Lyon), Mohammed El Rhabi, Abdelilah Hakim (PR, LAMAI-univ Cadi Ayyad, Marrakech, Maroc), Sahar Saoud et Zouhir Mahani (MCF, univ. Ibn Zohr, Agadir, Maroc)

— Séparation Aveugle de Sources (SAS) : intervenants : Eric Moreau (responsable scientifique, Pr, université de Toulon), Hassan Fenniri (MCF, CReSTIC, URCA), Amor Keziou (MCF, Lab. Maths-URCA), Mohammed El Rhabi, Abdelilah Hakim, Zouhir Mahani et Abdelkarim Zatni (Pr, Univ. Ibn Zohr, Agadir, Maroc).

— Approches connexionnistes et Réseaux de Neurones Convolutifs (CNN) : intervenants : Christophe Garcia (responsable scientifique, Pr, LIRIS, Insa Lyon), Khalid Idrissi (MCF,LIRIS, Insa de Lyon), Frank Lebourgeois, Zouhir Mahani et Abdelkarim Zatni.

Les travaux concernant la séparation aveugle de sources ont été valorisés dans le cadre de la thèse d'Abdelghani Ghazdali (voir la section 5.3.1), ceux en traitement d'images dans le cadre des thèses d'Amine Laghrib (voir la section 5.3.2) et d'Idriss El Mourabit (voir la section 5.3.3). Ces thèses ont été évidemment abordées dans un cadre plus général.

Nos résultats ont été publiés dans [9, 10, 11, 13, 12, EHM+12, MZS+12, EFHM13, BEFM10].

5.3 Encadrements de thèses

Grâce à ce projet pour les trois premières thèses et à ma collaboration avec l'université Cadi Ayyad à Marrakech au Maroc (où je suis chercheur associé), j'ai pu participer à l'encadrement de :

— thèses soutenues : 3 ;

— thèses en-cours : 3 ; (voir partie 7)

Dans cette partie, nous décrirons les thèses en liens direct avec le projet Toubkal ou au moins en lien avec le traitement du signal et le traitement d'images. Dans la section 7, nous présenterons les thèses en lien avec la modélisation numérique des mouvements de foules.

5.3.1 La thèse d'Abdelghani Ghazdali

Cette thèse porte sur la séparation aveugle de sources (SAS). Cette méthode consiste à restituer des signaux « sources » à partir de la seule connaissance d'un ensemble de signaux mesurés, les « observations ». Chacune de ces observations est constituée d'un mélange linéaire de ces sources où les coefficients du « mélange » sont inconnus. Ce problème peut se résoudre moyennant trois grandes hypothèses :

8. voir http://elrhabi.free.fr/Projet

1. l'indépendance statistique des signaux « sources »
2. l'opérateur de mélange est inversible (au moins à gauche)
3. au plus une des sources est un signal à distribution gaussienne

Sous ces hypothèses, le séparation aveugle de sources peut alors se ramener à un problème d'optimisation stochastique avec la définition d'un critère d'indépendance. Les « sources estimées » approchant les « sources » originales sont un extremum de ce critère.

Tout d'abord, nous avons proposé un critère d'indépendance basé sur les copules, ce qui a conduit à un nouvel algorithme de séparation aveugle de sources pour des signaux indépendants. Ce dernier critère a l'avantage d'être plus général et nous avons pu l'étendre à la séparation de sources dépendantes à condition d'avoir une *a priori* connaissance de la structure de la dépendance. À partir des travaux proposés dans [13], nous reprenons le modèle tenant compte du bruit additif et dans [7], nous proposons une méthode en deux temps combinant une réduction du bruit au niveau des signaux observés et une régularisation du critère de séparation afin d'éviter d'amplifier le bruit résiduel. Cette régularisation est basée sur le contrôle de la variation totale du signal. Les résultats de simulations montrent la consistance des estimateurs des sources obtenues dans les cas de sources indépendantes, dépendantes, avec ou sans bruit.

Cette thèse a été soutenue publiquement, à l'université Cadi Ayyad à Marrakech, le 12 décembre 2015. Abdelghani est à présent enseignant-chercheur à l'Ensa (eq marocain d'Insa) de Khouribga au Maroc, depuis septembre 2016.

5.3.2 La thèse d'Amine Laghrib

Cette thèse porte sur le problème de la super résolution pour des images sous résolues. Ici, l'idée est d'obtenir une image haute définition à partir d'une suite d'images sous-résolues de la même scène. Ce problème est mal posé et une approche par des modèles variationnels nous a semblé pertinente. Afin de résoudre ce problème, deux étapes sont cruciales : le recalage (en choisissant une image, recaler les autres dans le référentiel de cette image pour une mise en correspondance cohérente des pixels) et l'interpolation (une fois les images recalées, il s'agit de reconstruire l'image « mieux » résolue). Dans un premier temps, nous avons proposé deux approches, ayant conduit à deux algorithmes, afin de traiter la partie proposant le recalage, à savoir : un recalage par « diffusion » et un recalage « élastique ». Pour ces deux approches, nous avons supposé que les images n'ont subi que de « petites » déformations, nous proposons alors un critère régularisé. La régularisation par diffusion consiste à privilégier les déformations lisses. L'idée derrière le recalage élastique est de traiter la déformation de l'image de référence par rapport à l'image modèle comme étant un processus physique qui ressemble à l'étirage d'un matériau élastique. Il se résout grâce aux équations de l'élasticité linéaire. Pour chacune des deux régularisations, le cadre fonctionnel est posé et un résultat d'existence et d'unicité est montré [9, 11]. Dans une seconde partie, nous nous sommes intéressés à la dernière étape de la super résolution qui consiste à réduire, à la fois, le flou dû à l'interpolation et le bruit dû à la caméra. Pour cela, nous avons proposé deux méthodes variationnelles permettant de résoudre le problème de

l'effet « staircasing » (effet d'escalier) engendré par la désormais classique régularisation par la variation totale. Tout d'abord, nous avons proposé une combinaison entre deux régularisations , l'une d'ordre 1 (en fait la variation totale) et l'autre d'ordre 2 (cette dernière doit atténuer l'effet de staircasing tout en préservant les contours). Encore un fois, le problème est écrit dans un cadre mathématique bien posé. Au final, de tests numériques ont illustré les performances des méthodes proposées en les comparant avec d'autres approches de la littérature proposée dans cette thématique (la super résolution) [7, 5]. Amine a soutenu sa thèse le 19 décembre 2015 à l'université Cadi Ayyad à Marrakech. Il est à actuellement enseignant-chercheur à l'université Sultan Moulay Slimane à Béni-Mellal au Maroc, depuis septembre 2016.

5.3.3 La thèse de Idriss El Mourabit

Cette thèse porte sur la restauration d'images par des méthodes variationnelles dans un cadre mathématique rigoureux. Plus précisément, elle concerne des problèmes classiques de traitement d'images, en l'occurrence le débruitage, la super-résolution et la déconvolution aveugle. Le point commun des modèles est qu'ils font intervenir un terme de régularisation portant sur la variation totale des images, ou bien des variantes, en particulier la variation totale bilatérale (BTV). Tout d'abord, nous proposons deux nouveaux algorithmes de débruitages s'inspirant de la première partie de l'algorithme dit de projection proposé dans [13]. Cela offre une alternative originale à l'algorithme de référence « Chambolle et Pock » ([CP11]). Nous montrons que ce problème est bien posé. Puis, nous proposons une variante à l'ordre 2 (reposant sur les dérivées à l'ordre 2) est aussi proposée, et nous montrons à nouveau son caractère bien posé. Les expériences montrent que la décroissance de l'énergie est bien plus rapide dans les premières itérations pour l'algorithme de projection proposé, aussi bien à l'ordre 2 qu'à l'ordre 1 [8]. Ensuite, nous avons traité deux problématiques importantes et difficiles en traitement d'images : la super-résolution et la déconvolution du flou. Cette fois encore, ces sujets ont été abordés d'un point de vue mathématique et une analyse rigoureuses des deux problèmes a été établies. Pour chacun des deux problèmes nous proposons de nouvelles méthodes comparées à celles de la littérature. Pour le problème de la super-résolution, c'est surtout l'étape de restauration qui est abordé ! Nous introduisons une nouvelle équation [6] permettant de réduire le bruit et le flou de l'image Haute Résolution. Il s'en suivra une étude théorique de l'équation proposée basée sur le théorème du point fixe de Schauder. En outre, des simulations numériques du problème seront présentées, puis une comparaison des performances avec des résultats de la littérature récente sera proposée en termes de PSNR et SSIM.En ce qui concerne la déconvolution du flou, nous reprenons les travaux de T. Chan *et al* [CW98] en changeant le terme de régularisation. Puis en considérant les travaux de Peronne *et al* [PF16], nous reformulons le problème à l'aide d'une fonctionnelle relaxée. Nous prouvons l'existence de solution mais pas l'unicité (la fonctionnelle n'est pas strictement convexe).

La soutenance d'Idriss a eu lieu à l'université Cadi Ayyad à Marrakech, le 30 octobre 2017.

5.4 Thèse en cours : la thèse de Fatim Zehrae Aït Bella

Cette thèse fait suite à celle d'Idriss El Mourabit (voir la section 5.3.3). Elle a pour objectif d'étendre les résultats d'Amine Laghrib (section 5.3.2) et ceux d'Idriss El Mourabit (section 5.3.3) par l'introduction d'opérateurs non locaux. Autrement dit proposer des méthodes de restaurations d'images par des opérateurs non locaux.

Cette étude contient une partie théorique non négligeable et une partie de tests de validation assez difficile à mettre en œuvre.

Dans un premier temps, nous nous intéressons à la restauration de document images, où la principale information contenue dans l'image est du texte. Nous supposons ces images acquises dans des conditions difficiles (faible luminosité et illumination variable, bruit, flou entre autres).

Fatim s'intéresse à la réduction de l'illumination variable pour une image de document (le document ne contient que du texte) [3]. Elle reprend une première approche que nous avons proposée dans [EHM+12, 12, MZS+12]. Ces premiers essais furent un succès. En effet, plus de 150 millions de téléphones portables ont été vendus dans le monde, intégrant les applications d'envoi de messages "manuscrits" (Digitizer, w-Postcard, Clipper, MagicWanda) utilisant cette approche en prétraitement.

Fatim a repris l'étude complète de cette approche basée sur des opérateurs locaux et elle en a proposé une extension en terme d'opérateurs non locaux. Tout comme dans [MZS+12], Fatim se base sur une classique décomposition de l'image en un produit d'une réflectance et d'une illumination variable, elle propose une méthode où l'objectif principal est l'utilisation des caractéristiques répétitives de la réflectance pour réduire de manière efficace l'illumination variable. Autrement dit, elle propose une extension de [MZS+12] à des opérateurs non locaux.

Elle conclut en proposant des résultats numériques illustrant les performances et la pertinence de cette approche pour un tel problème. Elle propose aussi une étude théorique, d'abord elle montre l'existence et l'unicité de solution du problème défini dans [MZS+12]. Puis elle s'intéresse à l'analyse du problème qu'elle propose. Une étude de l'existence et de l'unicité est aussi examinée.

Nous détaillerons les suites de cette thèse dans la section 8.1.

5.5 Recherche au sein d'une startup. Optimisation, EDP & traitement d'images

Ces travaux ont concerné la période 2006-2011, lors de mon recrutement au sein de la société Realeyes3d [9]. Cette startup était spécialisée dans l'imagerie numérique embarquée (typiquement smartphone) et comptait un peu plus de 50 employés (à l'international). Elle avait son siège social à Saint-Cloud, une filiale à San Francisco (Qipit, inc) et à Hong Kong, ainsi qu'un bureau à Tokyo et à Pékin. Elle a reçu de nombreux prix récompensant l'innovation en particulier [Rea07b, Rea07a]. En 2011, nous avons vendu nos brevets et logiciels à une société japonaise. Les applications de Realeyes3d ont été commercialisées et déployées à plus de **150 millions** d'exemplaires sur les terminaux de plusieurs constructeurs de premier plan, ainsi que chez plusieurs grands opérateurs mobiles.

9. https://www.crunchbase.com/organization/realeyes-3d

En 2008, j'ai été nommé directeur scientifique et j'ai organisé le pôle « Recherche & Développement » de la startup Realeyes3d, ainsi :

Développement pur : Le client ne sait pas faire : mais la solution existe.

Recherche & Développement : Le client ne sait pas faire : la solution peut exister mais il existe des verrous techniques (ce qui n'est pas faisable selon l'état de l'art) avec de l'IP potentielle (brevet) ?

Recherche : Pas de client mais une idée qui a un marché potentiel mais pas encore de solution connue : il se pose alors la question des verrous, du coût, du moment où il faut s'arrêter en l'absence de solution stable.

La startup Realeyes3d avait pour ambition de « rendre intelligente » (ou au moins utile) la prise d'images à l'aide de son smartphone. Elle s'appuie donc fortement sur son pôle « Recherche & Développement ».

J'avais donc proposé deux grands axes focalisés sur la caméra du smartphone :

Restauration d'images : Les caméras des smartphones dégradaient très souvent le contenu des images. Lorsque ce contenu est un « code » (caractères ou code à barres). Par exemple lorsque l'utilisateur souhaite envoyer à l'aide de son smartphone un document similaire celui qui serait produit par un scanner à plat ou si cet utilisateur souhaite décoder un code à barre pour avoir de l'information sur un produit.

Estimation du mouvement relatif de la caméra : A l'aide la caméra et d'une estimation de mouvement relatif de la caméra, le contrôle d'interface devient possible. Par exemple naviguer dans une image par le mouvement ou faire défiler les noms d'un agenda (ou tout autre application).

5.5.1 La numérisation d'un document à l'aide d'une caméra (par exemple un smartphone)

Nous avons proposé une méthode originale de numérisation de documents basée sur une optimisation non convexe et nous avons établi un résultat nous permettant de ramener le problème à la résolution d'une équation aux dérivées partielles d'évolution [3] :

— lorsque le document ne contient que du texte (par exemple : facture, ticket de caisse, voir [MZS+12]) (voir figure 1 [10])

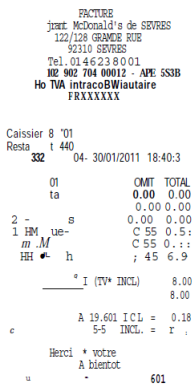
— lorsque le document est plus complexe, il contient à la fois du texte et des images ou logos etc (exemple magazine) ([12, EHM+12]). Ici une étape de segmentation est nécessaire afin de séparer les zones de textes et celle des images contenues dans le document (segmentation basée sur la ligne de partage des eaux, voir [Bea06]). Le texte est traité par la méthode précédente, l'image par un retinex (voir [RW04]). Le document final est la fusion des deux solutions issues des deux méthodes (précédemment citées [MZS+12] et [RW04]). Cette fusion est réalisée à l'aide de la résolution d'une équation de poisson (voir [PGB03]).

10. https://www.abbyy.com/fr-fr/

Le cas échéant, la « correction » de perspective du document (dewarp, deskew) est aussi traitée. La méthode (désormais classique) est basée sur une transformée de Hough.

Ces travaux ont été repris par Fatim Zehrae Aït Bella, étudiante en deuxième année de thèse à l'université Cadi Ayyad (Marrakech, Maroc), ses premiers résultats ont été soumis dans [3], voir aussi la partie 5.4.

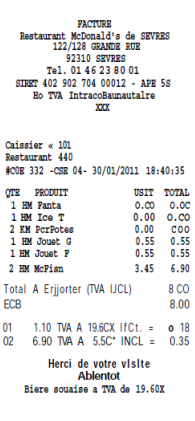
(a) Example 3, original image



(b) Recognition by a classical OCR software ; original image

(a) L'image originale et sa reconnaissance par l'OCR ABBYY Fine Reader

(e) Text enhancement : estimated reflectance



(f) Recognition by a classical OCR software ; estimated reflectance

(b) L'image restaurée et sa reconnaissance par l'OCR ABBYY Fine Reader

FIGURE 1 – Restauration et numérisation d'un document image.

5.5.2 Réduction du flou dans une image

Cette étude a été réalisée dans le cadre d'un projet avec Orange Labs en 2007.

Les smartphones avaient très souvent des systèmes optiques d'assez basse qualité. Ainsi, lors de la prise de vue, l'image subissait diverses transformations, par exemple, des distorsions optiques ou un flou dû à une distance focale mal ajustée. C'est à ce dernier type de transformation que l'on s'est intéressé.

Le flou dans les images est un phénomène essentiellement convolutif. Il est, en grande partie, dû au fait que la profondeur de champ d'un appareil photographique ne peut être infinie. De plus, les cameraphones (hormis les smartphones) sont souvent fabriqués avec des focales fixes (i.e. sans auto-focus ni de mode macro) et/ou une ouverture fixe, seule la vitesse d'obturateur est commandée. En outre, les capteurs utilisés dans les caméraphones sont très limités dans des conditions de luminosité faible. Dans ces conditions, on récupère souvent des images bruitées et floues.

Il existe plusieurs sortes de flou. Souvent, lorsque le plan de l'image ne coïncide pas avec le plan du capteur, on dit que le plan objet est défocalisé : l'image apparaît floue. On parle alors de flou d'*out-of-focus* ou *défocalisation*. Si c'est le mouvement de l'utilisateur qui crée le flou dans l'image, on parle alors de flou de bougé. Enfin il se peut aussi que le flou soit causé par des turbulences atmosphériques (température, du vent, humidité entre autres).

Dans le cas où le système optique peut être assimilé à un SLIT [11], il est possible de lui associer une fonction caractérisée par une réponse impulsionnelle (PSF en anglais).

Le flou dans une image est complètement caractérisé par sa PSF. Ces PSF ou noyaux de flou sont décrits par une ligne pour les flous de bougé, un disque pour les flous de défocalisation et enfin par un noyau gaussien pour les flous atmosphériques.

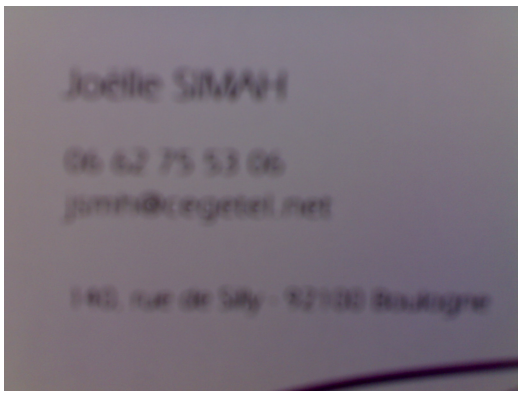


FIGURE 2 – Exemple d'une image ayant subi un flou d'out of focus acquise à l'aide du Nokia N70 dont l'appareil photographique est à focale fixe.

Le problème peut se modéliser par une déconvolution aveugle, aveugle car il n'est pas possible de connaître le flou *a priori* à cause de la variabilité des caméra des smartphones.

L'objectif de cette partie est d'introduire brièvement le modèle de déconvolution aveugle appliqué à l'image.

11. Système linéaire Invariant par Translation

La déconvolution aveugle est le processus d'estimation d'une image idéale et du noyau de flou à partir de la seule connaissance de sa version floue (dégradée) (pour plus de détails voir le tutorial [KH96]).

Autrement dit, si nous considérons l'équation :

$$u(x, y) = h * v(x, y) + b(x, y), \tag{1}$$

où $*$ désigne le produit de convolution, le couple (x, y) est un pixel, $h \in L(\mathbb{R}^2)$ est un noyau de convolution et b est une variable aléatoire supposée gaussienne caractérisant le bruit.

Une analyse rapide de ce problème montre assez facilement que :

— En présence de bruit, une déconvolution exacte est impossible, seule une solution approchée est accessible.

— Le plus souvent, cette solution n'est pas unique, il est donc nécessaire d'imposer des contraintes pour réduire le champ de solutions (et/ou) d'ajouter des hypothèses et un *a priori* sur l'image à restaurer.

— Nous avons à estimer à la fois le noyau de flou et l'image« idéale ». Un algorithme permettant d'estimer la solution approchée du problème 1 aura forcément un coût de calcul élevé. Notre challenge a été d'établir un procédé ayant le meilleur compromis coût, robustesse et portabilité.

Pour résoudre ce problème, nous avons proposé un nouvel algorithme de déconvolution du flou. Nous nous sommes placés dans le cas où le noyau est un flou de défocalisation, nous avons supposé que ce noyau est de la forme

$$h(x, y) = \frac{1}{\pi r^2} \mathbb{1}_{B_r}(x, y), \tag{2}$$

où $\mathbb{1}_{B_r}(x, y)$, est la fonction indicatrice de la boule de rayon r, où $r > 0$.

Ainsi, ce flou est complètement caractérisé par le rayon r, nous avons alors proposé une nouvelle méthode d'optimisation rapide d'estimation conjointe du flou (ici le rayon !) et de l'image dans deux cas particuliers, à savoir :

— l'image est un code à barre

— l'image est une image ne contenant que du « texte »

Nous avons alors construit un critère tenant compte de cet *a priori*. Á l'aide d'un OCR (reconnaissance automatique de caractère), nous avons établi une méthode rapide basée sur l'optimisation de ce critère. Cette étude a fait l'objet du dépôt de deux brevets à l'INPI (France) et à l'OMPIC (équivalent de l'INPI au Maroc) (voir [ER08, EMHS14]).

Nous illustrons les performances de notre algorithme par la restauration de deux types d'images floues :

— des images floues de codes à barres,

— des images floues contenant du texte.

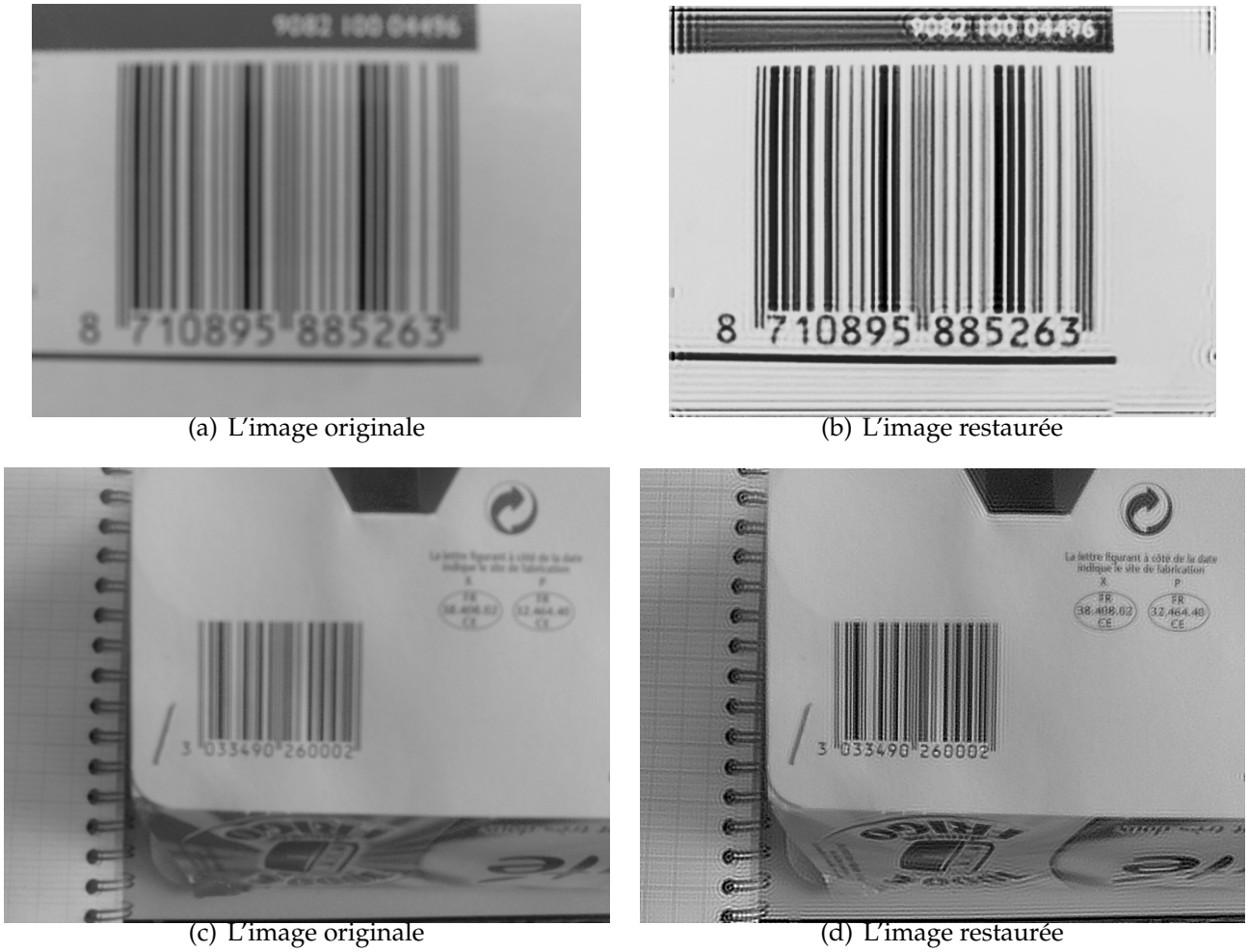
Ce choix est motivé par le fait que pour ce type d'images, il ne s'agit pas de restaurer leurs esthétiques mais bien de retrouver une information qui y est contenue.

Dans la figure 3, les images ont été prises à l'aide de la caméra d'un smartphone, à savoir le Nokia N80 [12]. Dans le premier test, le mode macro a été activé, l'image

12. voir https://www.gsmarena.com/nokia_n80-1347.php pour les spécifications techniques

obtenue est malgré tout floue. Ici, le flou est en majorité un flou de défocalisation. Afin de réduire le temps de traitement, les images ont été sous-échantillonnées de 2000×1504 (3 mégapixels) à 494×371 (splines cubiques). Le rayon du flou estimé par notre algorithme est $r = 4.92$ (cela correspond à un filtre de taille 5×5 pour une image de taille 494×371).

Dans le second test, le mode macro a été désactivé, ici le rayon estimé est $r = 5.14$.



(a) L'image originale (b) L'image restaurée

(c) L'image originale (d) L'image restaurée

FIGURE 3 – Restauration d'images de code barres floues.

Nous pouvons constater que non seulement les images ont été restaurées mais que l'information a suffisamment été reconstituée pour être décodée par un lecteur de code à barres commercial comme le montre la figure 4. Bien évidemment, le code à barre n'est pas décodé lorsque l'image est floue alors qu'il est décodé après prétraitement par notre algorithme.

Au bilan, en ce qui concerne la restauration des codes à barres et sur une base de plus de 500 codes à barres acquis par différents smartphones (à focale fixe), nous avons obtenu le résultat suivant ([13]) :

13. http://www.youtube.com/watch?v=i4ewsSJvkds

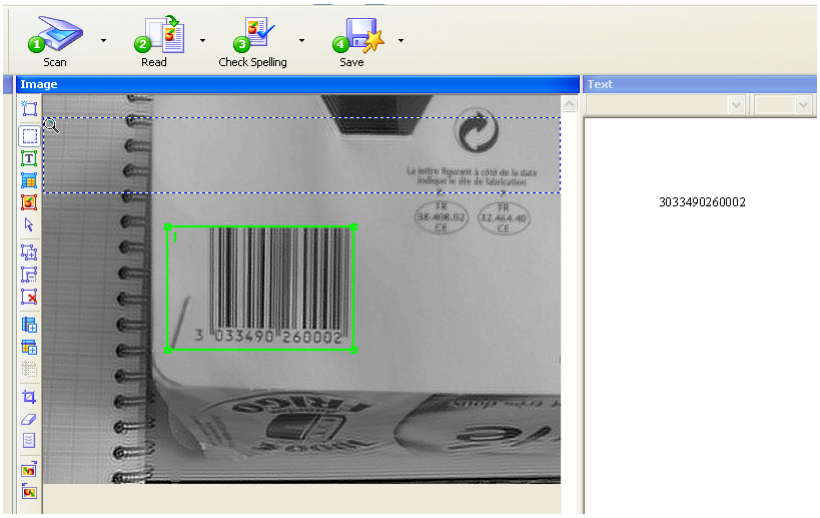


FIGURE 4 – L'image restaurée décodée par le lecteur de code à barres.

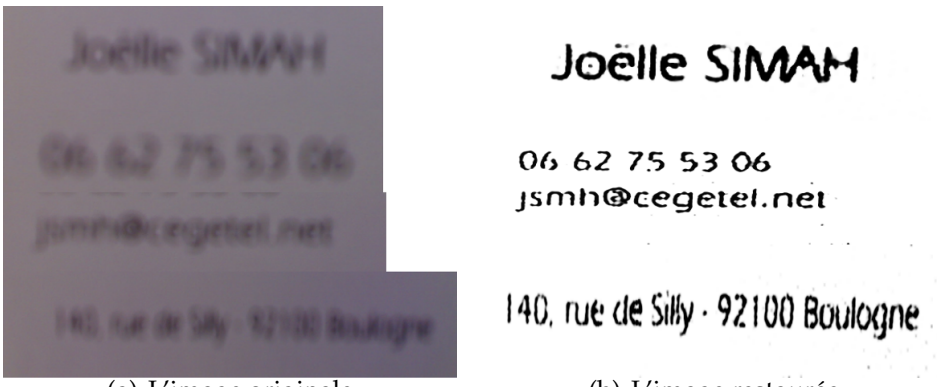
On a significant bench	Decoded	False positive	Not decoded
Nokia phone : N70 ; original	0 %	0 %	100 %
Nokia phone : N70 ; deblurred	85 %	4 %	11 %

Dans le test de la figure 5, nous nous sommes proposés de restaurer 4 crops (parties) de l'image de la figure 2 (pour des raisons de temps de calcul et aussi fin d'éviter de restaurer des zones inutiles) contenant un texte (carte de visite). Le rayon estimé est $r = 8.2$ pour des images de taille 467×121, 467×121 et 565×133. Ici, il s'agit d'un flou vraiment très important mais encore une fois, ce test illustre la robustesse de notre approche. Nous avons testé la reconnaissance des caractères contenu dans ces images par deux OCR commerciaux, ABBYY [14] et CardIris [15]). Les image floues n'ont pas été reconnues (aucun caractère). En revanche, tous les caractères contenus dans les images restaurées ont été reconnus.

Ce résultat avait été jugé limité par le pôle commercial de Realeyes3d. C'est pourquoi, j'avais proposé à des collègues-chercheurs universitaires des contrats sur ce sujet, à savoir :

— une collaboration d'une durée de 6 mois avec M. Tahar Boulmezaoud, MCF HdR au laboratoire de mathématiques de Versailles, où nous avions « revisité »le modèle de la déconvolution du flou. Nous avions relié ce problème au problème des moments (voir [ST43, Las09] par exemple). Nos premiers résultats avaient été soumis et présentés à la conférence : SIAM Conference on Imaging Science (IS10). Cette conférence a eu lieu du 14 au 15 avril 2010 à Chicago aux Etats-Unis (voir [BER10]). J'ai participé activement à l'effort de recherche et développement. Albert Cohen, professeur au laboratoire Jacques-Louis Lions, m'avait invité à présenter mes travaux dans le cadre des « Journées 40 ans du LJLL » [16]. J'avais choisi de présenter ce travail.

14. https://www.abbyy.com/fr-fr/
15. http://www.irislink.com/
16. http://www.ljll.math.upmc.fr/40ans/programme-theme1-2.php?Choix=programme



(a) L'image originale (b) L'image restaurée

FIGURE 5 – Restauration d'images contenant du texte (carte de visite).

— une collaboration d'une durée de 3 mois avec Laurent Dumas, MCF HdR au laboratoire Jacques-Louis Lions. Laurent Dumas m'avait proposé de recruter un stagiaire du Master MPE (Mathématiques Pour l'entreprise). Ils avaient tous les deux proposé une méthode d'optimisation stochastique (basée sur des algorithmes génétiques) originale pour la restauration de codes à barres flous. Les premiers résultats ont été présentés à la conférence SIAM Conference on Imaging Science (IS10) (voir [DER10]).
— Avec Laurent Dumas et Gilles Rochefort (ingénieur de recherche à Realeyes3d), nous avions retravaillé l'algorithme précédent et nous l'avions appliqué avec succès sur des données réelles. Nous avions soumis ce nouveau résultat à la conférence Congress on Evolutionary Computation (voir [DER11]).

Ces travaux ont été poursuivis dans la deuxième partie de la thèse d'Idriss El Mourabit, mon doctorant, en collaboration avec Amine Laghrib, un autre de mes doctorants actuellement professeur assistant. Nous avons détaillé cela dans la partie 5.3 de ce document.

5.5.3 Réduction du bruit

Lors de la mise en production d'un smartphone « Samsung », les ingénieurs du contrôle qualité s'étaient rendu compte du fait que les images produites par la caméra de ce téléphone portable étaient trop bruitées et surtout que la taille (en kilo-octet : ko) de ces images était trop grande (problème de stockage, d'envoi par e-mail entre autres).

Il nous avait alors été demandé de construire un algorithme répondant au cahier des charges suivant : améliorer l'image et réduire sa taille (en Ko) et surtout que cet algorithme puisse produire un résultat en quasi-temps réel (moins de 2 secondes).

Je m'étais alors penché sur le problème difficile de la réduction du bruit dans le contexte des smartphones (variabilités très fortes dues différentes conditions acquisition, aux différentes caméras dont la qualité varie selon les marques). J'avais alors recruté et encadré pour une durée de six mois deux stagiaires en Master 2 de la faculté des sciences et techniques à l'université Cadi Ayyad (Marrakech, Maroc). Après une étude bibliographique, nous avions proposé un nouvel algorithme de débruitage.

Puis avec l'un des stagiaires, nous avions travaillé sur l'amélioration de notre technologie d'estimation de mouvement par la caméra (*motion cortex*). En particulier, nous avions renforcé la prise en compte du bruit dans le flux d'images. Nous avions alors proposé une méthode robuste d'estimation du mouvement pour :

— la navigation dans une image grâce au mouvement et à la caméra (images, documents ou navigateurs web [17]
— le jeux où le mouvement et la caméra permettaient de remplacer le joystick et ainsi permettre la contrôle du jeu (console Nintendo DSI [18]

Cette technologie avait en particulier reçu le prix de la meilleure innovation lors du congrès « MobileMonday Peer Awards Barcelona 2007 » ([Rea07a]). Le Mobile Monday a été créé en Finlande dans les années 2000. Il est arrivé en France en 2005. Le Mobile Monday est un événement dédié à la mobilité. Des entreprises présentent leurs application mobile et des thématiques communes sont abordées pour partager les dernières innovations du mobiles. Ces travaux ont été poursuivis par Idriss El Mourabit, étudiant en thèse que j'ai co-encadré (50%). Idriss a soutenu sa thèse le 30 octobre 2017, il a obtenu sa thèse à l'université Cadi Ayyad à Marrakech au Maroc avec la mention très honorable. Il a eu 3 rapporteurs, 2 rapporteurs marocains et 1 rapporteur français Pascal Monasse (chercheur hdr, professeur à l'École des Ponts). En effet, dans la première partie de sa thèse, Idriss propose deux nouveaux algorithmes de débruitages (voir [8]) qu'il compare à ceux de la littérature.

5.5.4 Suppression d'objets dans une image, inpainting

Dans un cadre plus général et de reconstruction d'images, nous nous étions proposés de travailler sur diverses techniques de restauration d'images à savoir : « l'inpainting » ; le poisson image editing et dans un cadre plus ludique de colorisation.

Nous avions donc recruté et encadré une stagiaire (niveau en master) pour travailler sur cette thématique.

Le sujet avait porté sur l'évaluation d'outils de restauration dans les cas de modification de couleur (« color transfer »), et suppression de défauts (« inpainting ») et d'objets (« region filling »). Les problèmes sélectionnés avaient en commun d'exploiter des informations de référence (dans l'image ou dans une autre image) et des outils mathématiques pour la modélisation et la modification de l'image.

Le projet avait comporté les phases suivantes :

1. Analyse du problème, étude bibliographique
2. Sélection d'algorithmes et programmation
3. Évaluation, amélioration

Ce travail avait concerné le développement et l'évaluation d'algorithmes de reconstruction d'images. L'objectif avait été d'analyser, d'évaluer et de démontrer les solutions les plus pertinentes de la littérature. Ce travail nous avait permis de constituer une base interne de programmes et de démonstrations.

Entre autres, ce projet avait permis l'élaboration d'une nouvelle méthode d' « inpainting » (voir fig. 6) non basée sur la résolution d'une EDP mais sur la notion et la

17. voir http://www.youtube.com/watch?v=kNoTTXOyDjg et https://www.youtube.com/watch?v=UOE9nl6OU9M
18. voir http://www.youtube.com/watch?v=vYqjKttinl0

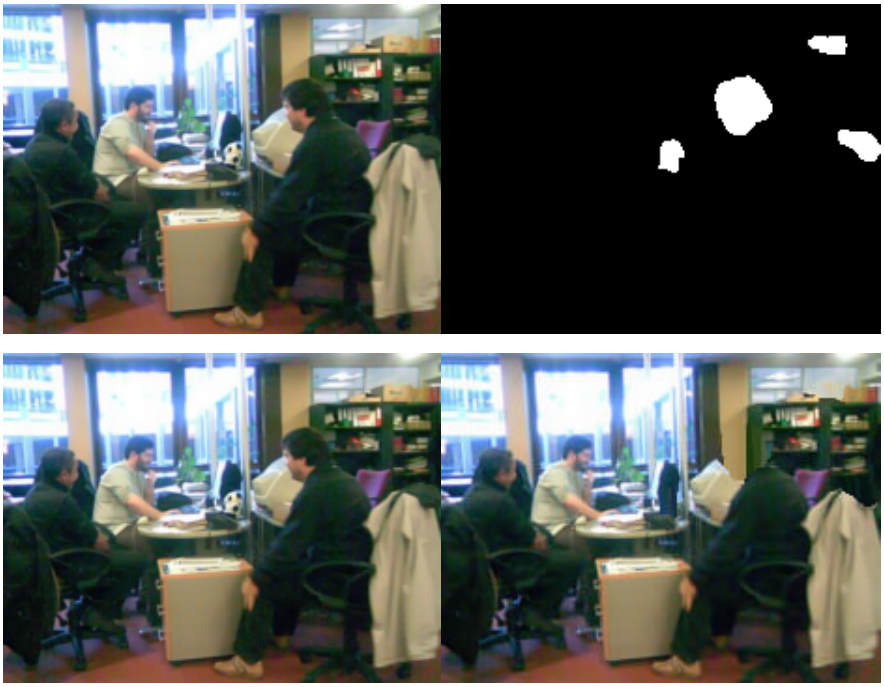


FIGURE 6 – En haut de gauche à droite : l'image d'original et le masque indiquant les zones à supprimer ; en bas de gauche à droite : l'image d'origine et l'image traitée par notre méthode d'inpainting

définition de distances discrètes et sur une optimisation discrète imitant le comportement de l'équation eikonale. Celle-ci reconstruisait à la fois la géométrie manquante de la zone à inpainter, mais aussi les textures, et ce en une seule passe. Le but final étai la proposition d'une méthode en quasi-temps réel sur un smartphone (le masque indiquant les zones à reconstruire était supposé donné).

5.6 Travaux postdoctoraux : optimisation stochastique ; séparation aveugle de sources ; traitement du signal

J'ai travaillé dans le cadre d'un stage postdoctoral, proposé par le laboratoire DÉCOM-CRESTIC (Centre de Recherche en STIC) à l'université de Reims Champagne-Ardenne, *sur la séparation aveugle de source en mélange convolutif*. Il m'a permis, d'une part, d'élargir mon thème de recherche et d'autre part, de travailler dans un environnement pluridisciplinaire.

En effet, une des thématiques de cette équipe de recherche était le diagnostic des machines tournantes par analyse vibro-acoustique.

Dans le cas des signaux issus de machines tournantes - qu'il s'agisse d'applications à des fins militaires ou industrielles- l'effort est mis sur la maintenance conditionnelle et sur la recherche d'éventuels dysfonctionnements. Le but est de diagnostiquer le type de panne et la machine à incriminer, sans que cela nécessite pour autant l'arrêt de toutes les autres unités en état de marche car cela serait trop pénalisant du point de

vue des coûts (immobilisation du processus de fabrication entre autres).

Notre projet de recherche s'insérait dans les objectifs de l'action incitative inter-Groupes de Recherches en Automatique, Electrotechnique et ISIS (Information, Signal, Image et viSion). Cette action, intitulée "Commande et Diagnostic de systèmes à entraînement électrique", était financée conjointement par le CNRS et le Ministère de la recherche. Par ailleurs, ce stage postdoctoral entrait aussi dans le cadre du projet *SurVib* (Surveillance de machines industrielles par diagnostic Vibratoire) dans une stratégie de maintenance prédictive.

Le but de nos travaux était d'utiliser la *déconvolution aveugle multicapteur* dans un but de pré-traitement des signaux afin de faciliter le diagnostic des systèmes mécaniques étudiés. Pour cela nous avions proposé un modèle très général permettant de dégager une étude originale des mélanges convolutifs spécifiques aux signaux à spectre harmonique. Afin de proposer de nouveaux algorithmes, la construction mathématique du modèle pour décrire les phénomènes observés et l'optimisation des fonctions séparatrices avait été examinée. Ceci nous avait amené à approfondir les conditions d'équilibre de la solution séparante et d'optimiser le comportement de ces algorithmes dans son voisinage. Puis, nous avions défini le cadre expérimental des simulations à mener, et présenter une validation expérimentale de l'étude théorique développée. Enfin, nous avions testé par la suite la robustesse des approches élaborées.

La clé pour résoudre le problème était le choix d'un critère d'indépendance. Le passage en revue des critères d'indépendance existants avait montré que l'approche considérant l'information mutuelle comme critère séparant s'était révélée pertinente. Nous nous ramenions donc à résoudre un problème d'optimisation stochastique.

Toutefois, les indétermination, inhérentes au problème, nous avaient conduit à l'imposition de contraintes supplémentaires. Une contrainte naturelle était de normaliser les sources estimées. Elle permettait alors de lever l'indétermination d'échelle et d'éviter les convergences triviales (vers $\mathbf{0}$).

Dans un premier temps, j'avais proposé une critère pénalisé basé sur l'information mutuelle et sur un critère de normalisation. Ce nouveau critère atteignait son minimum en un vecteur ayant des composantes statistiquement indépendantes et normalisées. Le problème de minimisation avait été résolu par une méthode de gradient stochastique.

À l'issue du calcul du gradient de ce critère, il fut aisé d'élaborer un nouvel algorithme de séparation aveugle de sources que l'on avait comparé avec un algorithme existant et faisant référence. Les résultats de séparation furent été améliorés dans la plupart des cas de figure (au sens du rapport signal sur bruit, voir [EGFD04, 18] pour plus de détails). Le code de calcul avait été écrit *ex nihilo* en langage C et en utilisant l'interface graphique du logiciel Matlab (section 6).

Puis, nous avions appliqué cet algorithme à un mélange de signaux vibratoires, obtenu à l'aide du banc expérimental appartenant au laboratoire (cf figure 7). Ce banc est composé de deux moteurs (ayant des fréquences de rotations différentes) dont l'un des deux (le moteur 1 de la figure 7) avait un défaut de roulement. Deux capteurs-accéléromètres avaient été collés respectivement sur chacun des deux moteurs. Encore une fois, les résultats obtenus avaient montré les performances de ce nouvel algorithme. En effet, il avait non seulement restitué les signatures vibratoires de chacune

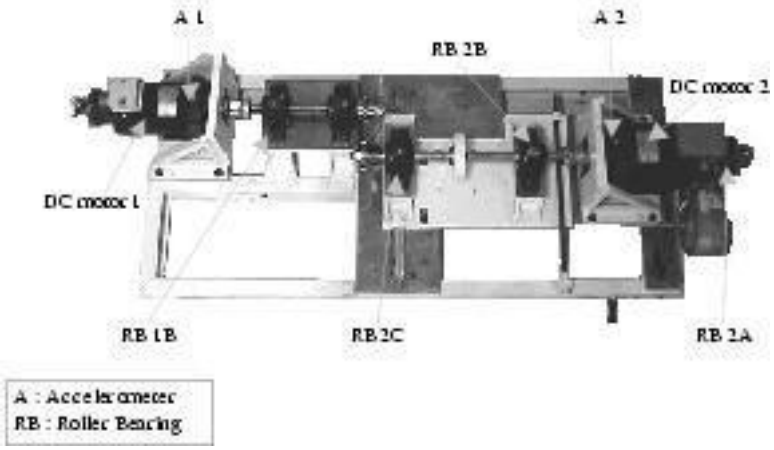


FIGURE 7 – Le banc expérimental.

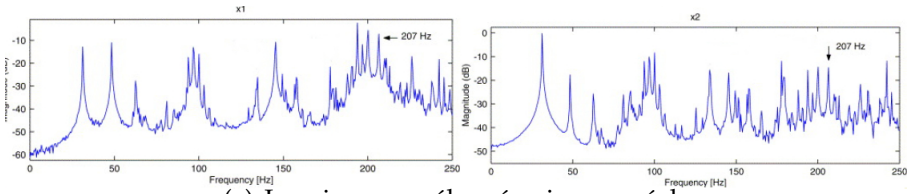
des deux machines mais aussi la fréquence correspondant au *défaut* du moteur 1 sachant que cette fréquence était présente dans les deux signaux-mélanges mesurés par les capteurs respectifs.

Cette étude avait été présentée lors de la 5ème conférence internationale « Acoustical and Vibratory Surveillance Methods and Diagnostic Techniques » ([EFGD04]). Par la suite et au vu des résultats obtenus, les membres du comité de lecture de cette conférence nous avaient recommandé de rédiger cette étude sous le format de la revue pour une soumission à la revue *Mechanical System and Signal Processing* (MSSP).

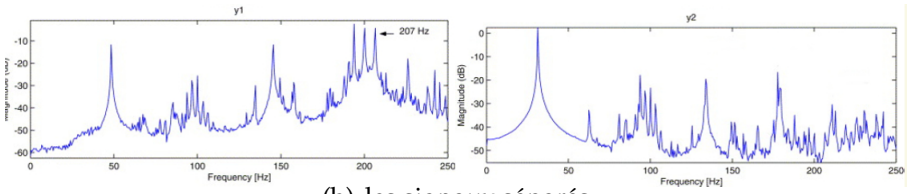
En dimension supérieure (dans le cas N sources / N capteurs avec $N > 2$) la distorsion causée par l'effet de l'opérateur filtrant sur la qualité de la séparation devenait de plus en plus inacceptable.

Pour être applicable, il était alors nécessaire de réduire l'effet de cet opérateur filtrant sur les sources estimées, surtout lorsqu'il s'agissait de proposer un diagnostic à partir d'un signal ou de sa représentation spectrale.

C'est pourquoi, j'avais couplé notre approche avec le principe de distorsion minimale. Cette approche appliquée aux signaux vibratoires s'était montrée dans un contexte multidimensionnel très innovante et avait permis une séparation des plus satisfaisantes. Ce fut cette version que nous avions soumis à la revue MSSP (voir [16] pour plus de détails). Un exemple de résultat, illustrant les performances de cette méthode, fut présenté dans la figure 8. En effet, nous avions inséré un roulement défectueux sur le moteur 1 du banc (voir figure 7. Ce défaut fut placé de manière à ce qu'il apparaissait au niveau du spectre à la fréquence 207Hz. Après avoir mesuré les signaux à l'aide de deux accéléromètres collés sur chacun des deux moteurs du banc, ces signaux furent été traités par notre algorithme en « aveugle ». L'algorithme séparait bien les signaux en resituant et associant les signatures à chaque machine mais aussi et surtout attribuait bien le défaut la bonne machine !



(a) Les signaux mélangés, signaux réels



(b) les signaux séparés

FIGURE 8 – Séparation aveugle de source, signaux vibratoires

5.7 Recherche au laboratoire de mathématiques appliquées à l'École des Ponts ParisTech

Lors de mon recrutement en tant que chargé de recherche au laboratoire de mathématiques appliquées à l'École des Ponts ParisTech (Cermics) en septembre 2006 (10 mois), la compréhension de la plasticité des métaux constituait un des défis majeurs de la science des matériaux, et une branche très active de la recherche. La cause principale de la plasticité d'un monocristal était l'existence en son sein de défauts microscopiques appelés dislocations. Sous l'action de contraintes de cisaillement, ces dislocations pouvaient se mettre à se déplacer et ainsi changer les propriétés mécaniques du matériau.

Au niveau microscopique, la dynamique des dislocations, bien que très complexe, était relativement bien comprise et simulée. À l'échelle mésoscopique du monocristal, la théorie de l'élasto-visco-plasticité, bien que très bien établie, restait généralement basée sur des lois de comportement phénoménologiques calées sur des données expérimentales. Ceci créait en particulier des difficultés pour décrire des cristaux de petite taille dans lesquels les effets d'échelle étaient importants et pour lesquels les données expérimentales n'étaient pas nécessairement accessibles.

D'un point de vue industriel, les enjeux étaient très importants. Il s'agissait de permettre la modélisation de matériaux cristallins de très petite taille (les nano-matériaux) et pour lesquels les modèles continus étaient peu satisfaisants. Ce fut le cas par exemple de la plasticité des couches minces utilisées pour les revêtements, les microstructures d'alliages multiphasés et les multicouches présentes dans les systèmes micro et nano électromécaniques (MEMS). L'enjeu était la modélisation des effets d'échelle observés dans la déformation plastique de ces structures. Il était en particulier important de rapprocher les méthodes numériques adaptées aux modèles micromécaniques développées à celles qui étaient classiquement utilisées en mécanique des structures et des microstructures pour permettre à terme un véritable transfert technologique.

En outre ce projet avait permis la coopération entre deux laboratoires, le Cermics à l'École de Ponts et Chaussées et le centre des matériaux à l'École des Mines de Paris. Il avait pour objectif de qualifier les modèles obtenus par les schémas d'homogénéisation mathématique dans le contexte des lois de comportement mécanique des

matériaux cristallins.

5.7.1 Projet proposé

Notre travail était centré sur la simulation et l'analyse numérique de cette classe de modèles qui couple les équations de l'élasticité linéaire avec des équations de type Hamilton-Jacobi vérifiées les variables de nature plastique. Les équations de l'élasticité avaient été calculées par deux méthodes : la méthode des éléments finis et une méthode basée sur une décomposition en série de Fourier. L'implantation des équations d'Hamilton-Jacobi posait des difficultés spécifiques.

Le point de départ de nos recherches, qui avaient débuté à la mi-septembre 2005 (jusque mi-juillet 2006) , furent les travaux de I. Groma et P. Balogh, en particulier [GB99] et les travaux proposés par Régis Monneau et deux de ses élèves : A. Ghorbel et A. El Hajj.

Nous avions tout d'abord proposé un modèle simple de dynamique de murs de dislocations de type « coins » de vecteurs de Burgers $\pm\vec{b}$. La dynamique des dislocations pour ce modèle était décrite par un système à deux inconnues (les dislocations $+$ ou $-$) de deux équations de transport dont la vitesse était donnée par la force de Peach-Koehler contenant les contributions de toutes les dislocations, même à longue distance.

Il s'agissait d'un système unidimensionnel d'équations Hamilton-Jacobi non locales. Cette étude fut présentée au cours du colloque National MECAMAT - Aussois 2006, École de Mécanique des Matériaux, qui s'était déroulé à Aussois, du 23 au 27 janvier 2006 (voir [AE06]).

Enfin, nous nous étions aussi intéressés à la simulation numérique du problème introduit par I. Groma et P. Balogh. Ces problèmes étaient aussi décrits par des équations de type transport non local mais cette fois dans le cas bidimensionnel.

5.7.2 L'homogénéisation numérique

Dans cette étude, nous nous proposions de calculer numériquement les vitesses moyennes des dislocations $+$ ou $-$ en fonction des contraintes imposées. Ceci nous permettait de retrouver une loi de comportement du type $\dot{\epsilon}^p = f(\sigma)$, où ϵ^p désigne le tenseur des déformations et σ le tenseur des contraintes.

Ici, nous présentions des résultats de simulation pour estimer l'hamiltonien effectif, noté $\tilde{H}$. Les équations du problème étaient résolues numériquement par un schéma aux différences finies de type « upwind » (voir par exemple [RT92]). L'hamiltonien numérique effectif $\tilde{H}^{num}$ était alors calculé en regardant l'évolution en temps long des solutions du problème.

Dans ce test, nous avions représenté le graphe de l'hamiltonien effectif $\tilde{H}$ en fonction de σ^{ext}, la contrainte extérieure, pour différentes données initiales $u_0^{\pm}$. Pour les simulations, la densité de dislocations type $\pm$ était fixée à $\rho = 4$. Nous obtenions alors des hamiltoniens numériques antisymétriques. De plus, il apparaîssait clairement deux valeurs seuils (opposées) $\sigma_0 = 0.09$ et $-\sigma_0$ au delà desquelles l'hamiltonien n'était pas nul. Pour des valeurs de σ^{ext} du même ordre de grandeur que celles du

terme décrivant les interactions entre dislocations, l'hamiltonien passait par un régime "transitoire" (de l'état de "piégeage" au "mouvement") pour atteindre lorsque $|\sigma^{ext}| >> \sigma_0$, un état où l'hamitonien devenait alors proportionnel à σ^{ext}, ce qui correspondait à un mouvement conservatif de dislocations (dans ce test on a supposé qu'il n'y a pas de sources de dislocations).

En outre, au vu de ces simulations, il semblait que les résultats concernant l'hamiltonien effectif numérique $\tilde{H}$ étaient peu dépendants des conditions initiales choisies.

5.7.3 Le problème de Groma-Balogh 2D

Dans cette partie, nous présentions nos essais de simulation numérique du problème de Groma-Balogh 2D. Il s'agit d'un système de transport non linéaire couplé avec les équations de l'élasticité linéaire.

La discrétisation s'effectuait à l'aide d'un schéma itératif basé sur un schéma de type « upwind » pour la partie hyperbolique et une décomposition en série de Fourier pour la partie elliptique.

Ici, nous considérions un matériau cristallin sachant que l'état cristallin était défini par un caractère périodique et ordonné à l'échelle atomique. Nous nous intéressions alors à la dynamique des densités de dislocations sous l'effet d'une contrainte extérieure.

Dans la figure 9, nous nous placions dans un pavé de ce matériau, où les dislocations étaient concentrées par paquets et nous nous proposions de regarder leurs évolutions au cours du temps. Nous constations , comme on aurait pu le prévoir intuitivement, que les dislocations se distribuaient uniformément (homogénéisation) dans tout le pavé. En outre, le code de calcul, qui a permis cette simulation, fut écrit en C++ en collaboration avec Philippe Hoch, ingénieur de recherche au CEA.

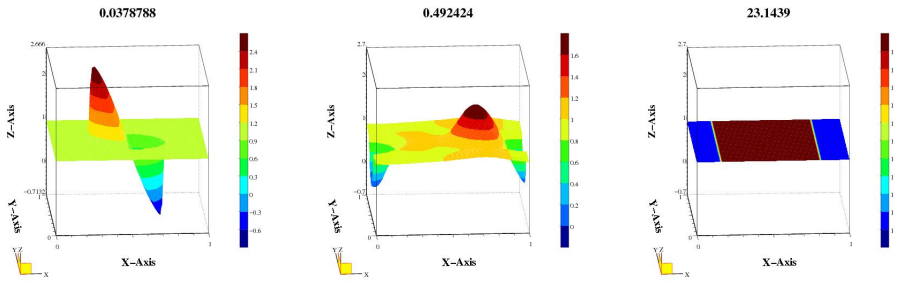


FIGURE 9 – Évolution des densités de dislocations soumises à une contrainte extérieure (au temps initial à gauche ; en un temps intermédiaire au centre ; au temps final à droite)

5.8 Travaux post-thèse : bilan

Les travaux de cet axe ont donné lieu aux publications suivantes :
— articles de revues : [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 16, 18],
— actes de conférences internationales : [EFHM13, EHM+12, MZS+12, DER11, DER10, BER10, BEFM10, EFD06, EGFD04],
— conférences nationales : [EFGD04, AE06, JAHE17],
— articles en cours : [EFM, AAHE, EAEH].

6 Mes codes de calcul

Mes différents travaux de recherche ont conduit à de nombreux algorithmes proposant des solutions approchées aux différents problèmes adressés. Je propose dans cette partie une liste de mes codes de calcul.

Les codes développés (*ex nihilo*) durant ma thèse (section 4) :

MSM_MAX est un solveur $3D$ discrétisant le problème de Maxwell dans le cadre d'une décomposition non conforme de domaine par une méthode d'éléments spectraux avec joints. Il traite de la simulation numérique du champ magnétique dans le cas tridimensionnel. Ce code de calcul représente environ $5\,000$ lignes de code Fortran (section 4.1.2 et [17, 15]).

MS_MAX est une version « conforme » du solveur MSM_MAX (dans le cadre d'une décomposition conforme du domaine). Son développement, bien que plus simple, est néanmoins différent. Il a représenté environ $2\,000$ lignes de code Fortran (section 4.1 et [14]).

Les codes en traitement d'images (section 5.5) :

DOC_SCAN est un solveur proposant la numérisation d'un document complexe (texte+image, magazine) acquis dans des conditions d'illumination variable (section 5.5.1 et [MZS+12, 12, EHM+12, 3]).

BLIND_DEB est un solveur proposant une nouvelle méthode de réduction du flou dans une image (de texte ou un code à barres ou plus généralement lorsque l'on connaît *a priori* le nombre de niveau de gris dans l'image). Il estime conjointement le noyau de flou (dans le cas du flou de défocalisation ou le flou atmosphérique), l'illumination variable et une approximation de l'image dite idéale (voir la section 5.5.2 et [ER08, EMHS14, 6]).

MOTION_EST est un solveur temps-réel et conçu pour être embarqué dans un smartphone. Il estime le mouvement apparent de la caméra. Ce mouvement est alors traduit en commande contrôlant une interface, permettant de surfer sur le web, naviguer dans une image (un plan), de se déplacer par un simple mouvement de gauche à droite et de zoomer par un simple mouvement de haut en en bas (section 5.5.3).

FAST_DEN est un solveur quasi temps-réel, il a été conçu pour répondre au compromis : réduire la taille (en ko) d'une image bruitée et améliorer l'interprétation de cette image en quasi temps réel (section 5.5.3 et [**?**, 8]).

FAST_INPAINT est un solveur permettant la suppression d'objets dans une image (il permet aussi de reconstruire au possible des zones détériorées ou perdues dans une image). Ce solveur imite le comportement de l'équation eikonale (section 5.5.4).

Les codes développés (*ex nihilo*) en traitement du signal (section 5.6)

NQ_BSS est un solveur basé sur une méthode de type « Quasi-Newton » pour la résolution du problème d'optimisation stochastique. Il fera l'objet d'une publication à soumettre dans les mois à venir [EFM].

TV_BSS est un solveur proposant la solution d'un problème d'optimisation stochastique liée à la séparation aveugle de source dans le cas où les données observées sont bruitées. Il permet de tenir compte de la variation totale du signal. Il fournit une approximation de la solution au problème la moins bruitée possible ([13]).

PMI_BSS est un des solveurs que j'ai développés au cours de mon premier stage posdoctoral. C'est un code général écrit en C et utilisant l'interface graphique de Matlab pour le traitement des données. Il a été appliqué à la séparation aveugle de signaux synthétiques, signaux vibratoires (réels) mais aussi à la restauration d'images mélangées . Ce code de calcul représente environ $1\,000$ lignes de code C ([18]).

MDPMI_BSS est une extension du solveur PMI_BSS, il en améliore les performances dans le cas multidimensionnel (N sources/N capteurs, $N > 2$) puisqu'il permet de minimiser la distorsion due aux indéterminations (en fait à celle due à l'opérateur filtrant !), inhérentes à la séparation aveugle de source ([16]).

Les codes de calcul en dynamique des dislocations (section 5.7) ont été développés à partir du code de calcul « GO++ » élaboré par Jean-David Benamou (INRIA) et Philippe Hoch (CEA).

Le code SPAR_PRE, développé durant mon DEA, est un solveur développé à partir de celui fourni par mon directeur de stage, il est essentiellement conçu autour de programmes, en Fortran 90, proposant des matrices de préconditionnement par une approximation « creuse » de l'inverse de matrices assemblées par une méthode d'éléments finis.

7 Autres travaux et encadrements : thèse en cours

Pierre Argoul, directeur de recherche à l'Ifsttar (mécanique des structures), travaille depuis plus de dix ans sur la dynamique des foules, en particulier sur l'effet de cette dynamique sur les structures. Après une discussion des problématiques au niveau mathématiques que lui et ses collaborateurs rencontrent, nous avons décidé de travailler ensemble. Cela a conduit à l'élaboration de deux sujets de thèse en partenariat avec Abdelilah Hakim, professeur au laboratoire de mathématiques appliquées et informatique à l'université Cadi Ayyad (Marrakech, Maroc).

7.1 La thèse d'Aïssam Jebrane

Le thème étudié concerne l'analyse des mouvements de la foule. Cette thèse est en co-tutelle entre l'École des Ponts-université Paris-Est et l'université Cadi Ayyad (Marrakech, Maroc). Elle est consacrée à l'amélioration (extension au cas des situations de panique) de l'approche discrète développer par l'équipe de Pierre Argoul, directeur de recherche à l'Ifsttar. Cela permet l'application du modèle à des situations de grand intérêt dans le domaine de la sécurité et du confort pour les usagers des ouvrages du génie civil et urbain.

Il s'agit d'une thèse essentiellement « théorique et numérique » dans le domaine des mathématiques appliquées aux sciences de l'ingénieur.

Le développement d'une approche discrète ou moléculaire pour la foule se fonde sur l'hypothèse que le mouvement de chaque piéton peut être défini par l'équation de Newton (piétons rigides) ou l'équation de la mécanique des milieux continus (piétons déformables) et par la définition des forces que le piéton échange avec son environnement (par exemple forces d'évitement).

La première phase est celle du développement et l'implémentation numérique d'une loi de contact qui permet de modéliser correctement les forces échangées entre piétons lorsque la foule est dense.

En perspective, nous espérons une mise en équations du modèle stochastique avec les hypothèses retenues (par exemple l'écart type des paramètres varie fortement) et la mise au point d'algorithmes adaptés pour sa modélisation numérique.

Les résultats attendus peuvent se résumer comme suit : obtention des forces de contact dans une foule dense voire très dense, prise en compte des aléas liés aux paramètres de la modélisation et application au cas de situations de panique.

Nos travaux ont été présentés ou publiés dans [1, AAHE, JAHE17].

7.2 La thèse d'Abdelghani Elmoussaoui

Le sujet de la thèse d'Abdelghani Elmoussaoui porte sur la dynamique des foules. Dans le cadre de sa thèse, il travaille sur les mouvements de groupes de piétons en passant par la théorie cinétique, qui permet de passer à la limite en remplaçant un groupe contenant N individus (avec N très grand) par une densité. Cette densité vérifie alors une équation aux dérivées partielles, similairement à l'équation de Boltzmann pour les gaz. Après une étude bibliographique, il a analysé une méthode qu'il propose, en se plaçant dans des domaines bornés non connexes. Les résultats ont été publiés dans le papier de revue [4].

Abdelghani a aussi travaillé sur l'influence d'obstacles sur le temps d'évacuation d'une salle (le cas de la « panique » est aussi envisagé) [2].

Actuellement, il travaille sur la caractérisation d'un état d'équilibre dans la théorie cinétique pour la modélisation des mouvement de foules. Dans cette partie, il pose mathématiquement le problème (existence de solution et conditions pour l'unicité) et il illustre son étude par des (premiers) résultats numériques [EAEH].

8 Perspectives et programme de recherche

Directement lié à mes travaux antérieurs, j'aimerais poursuivre l'effort de recherche en traitement d'images, en particulier des images de document et leur restauration. L'idée est la proposition d'un outil de numérisation pour smartphone capable de réduire l'illumination variable, le bruit et les distorsions (flous, déformations géométriques, dans la mesure du possible), causés par la caméra du smartphone. L'objectif est d'approcher autant que possible le résultat d'un scanner à plat. Bien évidemment, il existe de nombreuses applications offrant cette numérisation mais aucune ne traite vraiment complètement les documents complexes (photographie de magazine par exemple). Par ailleurs, cet outil pourrait aussi être un prétraitement très utile pour de nombreux logiciels de reconnaissance de caractères (OCR) (section 8.1).

En parallèle et plutôt lié à mes travaux en traitement du signal, je m'intéresserais, dans le cadre d'un projet franco-marocain, à des problèmes d'analyse de données multi-dimensionnelles, plus précisément ce projet porte sur démélange d'images (de spectroscopie de fluorescence, hyperspectrale) ou encore de signaux multidimensionnels en grande dimension. Ces données sont très souvent massives où la réduction de la dimension semble pertinente. Une grande partie du projet va adresser des problèmes mal posés et la recherche de régularisation pertinente sera une issue cruciale (section 8.2).

Enfin, je compte aussi continuer ma collaboration avec Abdelghani Elmoussaoui (section 7.2) et soutenir son effort de recherche sur la dynamique des foules, en particulier sur les modèles cinétiques et la proposition de méthodes résistant à la grande dimension (foules denses, section 8.3).

8.1 Traitement d'images et EDP

Notre ambition est de reprendre les travaux d'Amine Laghrib (section 5.3.2) et ceux d'Idriss El Mourabit (section 5.3.3) afin d'en améliorer dans un premier temps les aspects de modélisation par une nouvelle écriture du problème, à l'aide d'opérateurs « non locaux ».

Nous souhaitons cette approche la plus générale possible, nous nous intéresserons néanmoins aux images de document en particulier.

Dans un premier temps, nous étudierons l'apport des filtres non locaux (EDP sur des informations non locales et sur des patches d'images similaires voisines pour le filtrage), la restauration des formes des caractères et de la magnification spatiale des images de textes, en comparaison avec les filtres locaux par des filtres locaux (EDP) [3].

Cette approche non locale apporte une quantité d'informations plus importante pour restaurer et magnifier les images de texte car celles-ci contiennent une plus forte redondance d'informations spatiales que les images naturelles aux contenus plus aléatoires. Pour les vidéos, nous étudierons aussi l'apport des EDP spatiales et spatio-temporelles utilisant l'information des images successives temporellement pour, à la fois, segmenter le fond, magnifier les images en sous résolution, réhausser les images floues, filtrer le bruit et reconstruire les formes des caractères dégradés.

La seconde partie consistera à développer des algorithmes permettant de résoudre

les équations qui découlent de la partie théorique. Nous aborderons une analyse numérique qui définira l'algorithme et fournira une approximation de la solution en fonction de la consistance, la stabilité du schéma numérique. Puis il sera abordé la complexité de l'algorithme.

Les performances des algorithmes de prétraitement et de restauration des images de textes seront mesurées en comparant les résultats d'OCR sur des images prétraitées par rapport aux résultats obtenus sur des images non traitées. Une évaluation sera aussi conduite sur les performances de l'OCR à base de réseaux de neurones convolutifs (CNN) avec et sans prétraitement d'images et les performances de CNN robustes entraînés aux dégradations.

8.2 Traitement du signal et optimisation

Dans la continuité de mes travaux en traitement du signal et en particulier en séparation aveugle de source, nous avons répondu à un appel à projet Toubkal (voir section 5.2). Ce projet porte sur l'« Optimisation et réduction de dimension sur des problèmes de traitement de signaux multidimensionnel ou d'images en grande dimension ». Il aura pour objet le démélange d'images (de spectroscopie de fluorescence, hyperspectrale) ou encore de signaux multidimensionnels en grande dimension (tenseurs Big Data) par des décompositions matricielles et tensorielles (non-négatives) couplées (tenseurs d'ordre deux, trois et/ou quatre). Ce projet a été soumis en partenariat avec :

Nadège Thirion-Moreau : professeur au Laboratoire d'Informatique et des Systèmes (LIS), à l'université de Toulon (porteur français).

Abdelghani Ghazdali : professeur assistant Laboratoire de l'Ingénierie des procédés et optimisation des systèmes industriels, à l'École Nationale des Sciences Appliquées [19] (voir aussi le section 5.3.1).

Abdelilah Hakim : professeur de l'enseignement supérieur, laboratoire de mathématiques appliquées et informatique (LAMAI), à l'université Cadi Ayyad à Marrakech au Maroc.

Karima El Qate : étudiante en deuxième année de thèse, thèse en co-tutelle (université Cadi Ayyad/université de Toulon). Encadrants : Nadège Thirion-Moreau, Abdelilah Hakim

Mohammed El Rhabi.

Avec l'essor des systèmes d'acquisition et l'utilisation massive de capteurs voire de réseaux de capteurs, le volume des données acquise a explosé. Ces grands ensembles de données expérimentales peuvent bien souvent être organisés sous la forme de tableaux multidimensionnels. Les approches tensorielles et l'algèbre multilinéaire offrent, par conséquent, un cadre mathématique naturel et rigoureux pour la formulation de nombreux problèmes actuels et pour l'écriture des modèles qui leur sont associés. Il est également de plus en plus fréquent qu'une même quantité/scène soit observée au moyen de plusieurs capteurs aux caractéristiques complémentaires (par ex.

19. école d' ingénieurs universitaire marocaine, université Hassan I-Settat

en télédétection satellitaire : faible résolution spatiale mais très haute résolution spectrale pour des capteurs hyperspectraux et faible résolution spectrale mais très bonne résolution spatiale pour les capteurs multi-spectraux). Ce problème de fusion de données hétérogènes peut alors être réinterprété comme un problème de décompositions tensorielles couplées.

Dans ce projet, nous nous intéressons à des problèmes d'analyse de données multi-dimensionnelles [RTMC11]. Les blocs de données interprétés de manière tensorielle peuvent être modélisés au moyen de décompositions multi-linéaires (Parafac, Tucker, etc.). Afin d'estimer les facteurs (variables latentes) impliqués dans ces décompositions.

Les applications que nous visons plus spécifiquement portent principalement sur :

— la spectroscopie de fluorescence (imagerie 3D ou 4D [RTMC$^+$15]) pour la surveillance environnementale et le monitoring de la qualité de l'eau

— la déconvolution multi-dimensionnelle de signaux où l'estimation de directions d'arrivées utilisant des distributions spatiales quadratiques afin d'exploiter un plus grand nombre de modalités : temps-fréquence-capteurs, fréquence Doppler-retard-capteurs voire même temps-fréquence-capteurs-polarité (tenseurs d'ordre 4) [VMCTM15] pour des applications en surveillance Radar

— l'imagerie hyperspectrale pour des applications en surveillance/observation des milieux naturels (sols, érosion et trait de côte, détection de pollution etc.) [Ben11]

La finalité des algorithmes que nous développerons est l'identification des variables latentes (ce qui peut être vu aussi comme de la réduction de données) et la séparation (ou démixage) des sources ou composés purs présents au niveau des mélanges étudiés.

Ainsi, en spectroscopie de fluorescence, les méthodes développées ont pour but de détecter, séparer voire classifier :

i) des composés de matière organique fluorescents ou

ii) ou des métaux lourds (plomb, cadmium)

iii) ou encore des radionucléides présents dans les échantillons prélevés en un même site de surveillance (projet CNRS NEEDS en partenariat avec l'IRSN).

En contrôlant l'évolution de la composition de l'eau au cours du temps, l'objectif est de parvenir à détecter, de manière précoce, des changements environnementaux tels que ceux survenant lors d'incidents de pollution et pouvoir ainsi prévenir ou minimiser leurs risques.

Alors que les traitements conjoints des données consistent à décomposer simultanément un grand nombre de blocs de même dimension, les traitements couplés des données nécessitent la manipulation et la décomposition de blocs (matrices ou tenseurs) de dimensions différentes, correspondant à des grilles d'échantillonnage spatial, spectral (ou toute autre modalité considérée) non superposable et à des rapports signal à bruit éventuellement très différents. Il est alors nécessaire de résoudre un problème de décomposition couplée où le couplage fait non seulement intervenir des opérateurs d'interpolation mais éventuellement aussi d'autres types d'opérateurs (décalage, atténuation, entre autres). La plupart des méthodes de la littérature procèdent de manière itérative en alternant d'un problème de décomposition à l'autre [YYI11] ce qui est sous-optimal au lieu de considérer directement le problème couplé

[CFC16, AKD11]. Ce type de problème peut également se poser en dimension encore plus grande en surveillance par des réseaux de capteurs.

Nous y traiterons les aspects modélisation, couplage et choix des fonctions de coût à optimiser. Nous étudierons la robustesse des algorithmes développés vis-à- vis d'erreurs [KRTMC14] (rang, modèle, bruit additif, non-linéarités, non-stationnarité, etc.) et testerons également l'intérêt dans un tel contexte d'introduire des régularisations. Nous envisageons également de proposer des versions adaptatives [BATMZ13] afin de mieux prendre en compte l'aspect dynamique des phénomènes observés ou parallélisables (en utilisant MapReduce et l'architecture Hadoop popularisée dans le domaine des mégadonnées (Big Data [BKP14]).

8.3 Dynamique des foules

Je compte soutenir l'effort de recherche d'Abdelghani Elmoussaoui (section 7.2) et d'Abdelilah Hakim sur la théorie cinétique et les règles stochastiques pour la modélisation d'une foule en mouvement : ici nous nous intéresserons au développement d'un modèle mathématique se basant sur la théorie cinétique pour les particules actives où la gestion des interactions entre les individus est probabiliste. Cela conduira à une implantation numérique de schémas numériques basés sur des règles stochastiques (simulation par Monte Carlo-DSMC). L'idée derrière tout cela est d'augmenter le nombre de piétons dans les simulations numériques (un gros frein actuel de notre étude est une montée en échelle[nombre de piétons]), idéalement de pouvoir simuler des foules denses [BG16, BG15]. Un autre point très intéressant que nous souhaitons abordé est la gestion des interactions d'un point de vue probabiliste, en se basant sur des outils de la théorie des jeux pour gérer le cas de la panique et ses nombreuse applications pratiques [bel, BCG+16, Bel08].

8.4 Références bibliographique

[AKD11] E. Acar, T. G. Kolda et D. M. Dunlavy, *All at once optimization for coupled matrix and tensor factorizations*, Proceedings of Mining and Learning with Graphs, 2011.

[BATMZ13] A. Belouchrani, M. Amin, N. Thirion-Moreau et Y. Zhang, *Source separation and localization using time-frequency distributions*, IEEE Signal Processing Magazine (2013), 97–107.

[BCG+16] N. Bellomo, D. Clarke, L. Gibelli, P. Townsend et B.J. Vreugdenhil, *Human behaviours in evacuation crowd dynamics : From modelling to « big data » toward crisis management*, Physics of Life Reviews **18** (2016), 1–21.

[Bea06] R. Beare, *A locally constrained watershed transform*, IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (2006), 1063–1074.

[bel] *On multiscale models of pedestrian crowds from mesoscopic to macroscopic*, Communications in Mathematical Sciences **13**, no. 7.

[Bel08] N. Bellomo, *Modeling complex living systems : A kinetic theory and stochastic game approach*, Modeling and Simulation in Science, Engineering and Technology, 2008.

[Ben11] Z. Ben Rabah, *Démixage spectral d'images hyperspectrales et fusion possibiliste des connaissances. application á l'érosion hydrique*, Ph.D. thesis, ENST Bretagne, 2011.

[BG15] Nicola Bellomo et Livio Gibelli, *Toward a mathematical theory of behavioral-social dynamics for pedestrian crowds*, Mathematical Models and Methods in Applied Sciences **25** (2015), no. 13, 2417–2437.

[BG16] N. Bellomo et L. Gibelli, *Behavioral crowds : Modeling and monte carlo simulations toward validation*, Comput. & Fluids. Advances in Fluid-Structure Interaction **141** (2016), 13–21.

[BKP14] A. Beutel, A. Kumar et E. E. Papalexakis, *Flexifact : Scalable flexible factorization of coupled tensors on hadoop*, Proceedings of the 2014 SIAM International Conference on Data Mining, 2014, pp. 109–117.

[CFC16] J. E. Cohen, R. Cabral Farias et P. Comon, *Joint tensor compression for coupled cpd*, Proceedings of the 24th EUSIPCO Signal Processing Conference, 2016.

[CP11] A. Chambolle et T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vis. **40** (2011), no. 1, 120–145.

[CW98] T. Chan et C. K. Wong, *Total variation blind deconvolution*, IEEE Tternational Conference on Image Processing **7** (1998), no. 3, 370–375.

[GB99] I. Groma et P. Balogh, *Investigation of dislocation pattern formation in a two-dimensional self-consistent field approximation*, Acta mater **47** (1999), 3647–3654.

[KH96] D. Kundur et D. Hatzinakos, *Blind deconvolution*, IEEE Signal Processing Magazine, Vol. 13, no. 3 (1996), 343–64.

[KRTMC14] I. Kopriva, J.-P. Royer, N. Thirion-Moreau et P. Comon, *Error analysis of low-rank three-way tensor factorization approach to blind source separation*, Proceedings of the International Conference on Acoustic Speech and Signal Processing (ICASSP'2014), 2014.

[Las09] J. B. Lasserre, *Moments, positive polynomials and their applications*, Imperial College Press (2009).

[PF16] D. Perrone et P. Favaro, *A clearer picture of total variation blind deconvolution*, IEEE Transactions on Pattern Analysis and Machine Intelligence **38** (2016), no. 6, 1041–1055.

[PGB03] P. Pérez, M. Gangnet et A. Blake, *Poisson image editing*, ACM Transactions on Graphics (SIGGRAPH'03) **22** (2003), no. 3, 313–318.

[RT92] E. Rouy et A. Tourin, *A viscosity solutions approach to shape-from-shading*, SIAM J. Numer. Anal **47** (1992), 867–884.

[RTMC11] J.-P. Royer, N. Thirion-Moreau et P. Comon, *Computing the polyadic decomposition of nonnegative third order tensors*, Eurasip Signal Processing **91** (2011), no. 9, 2159–2171.

[RTMC+15] J.-P. Royer, N. Thirion-Moreau, P. Comon, R. Redon et S. Mounier, *A regularized nonnegative canonical polyadic decomposition algorithm with preprocessing for 3d fluorescence spectroscopy*, Journal of Chemometrics **29** (2015), 253–265.

[RW04] Z. Rahman et G. A. Woodell, *Retinex processing for automatic image enhancement*, Journal of Electronic Imaging **13** (2004), 100–110.

[ST43] J. A. Shohat et J.D Tamarkin, *The problem of moments*, American Mathematical Society Mathematical surveys ; American Mathematical Society, New York **II** (1943).

[VMCTM15] X. T. Vu, S. Maire, C. Chaux et N. Thirion-Moreau, *A new stochastic algorithm to decompose large nonnegative third order tensors*, IEEE Signal Processing Letters **22** (2015), no. 10, 1713–1717.

[YYI11] N. Yokoya, T. Yairi et A. Iwasaki, *Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion*, Eurasip Signal Processing **99** (2011), 1–10.

9 Liste des publications

9.1 Revues internationales avec comité de lecture

[1] A. Jebrane, P. Argoul, M. El Rhabi et A. Hakim. «Estimating contact forces/pressure in a dense crowd from the non-smooth approach view for dynamics of solids». *Submitted to Journal of Applied Mathematical Modelling*.

[2] A. Elmoussaoui, A. Jebrane, P. Argoul, M. El Rhabi et A. Hakim. «2D meso modelling of crowd motion with the kinetic theory approach : Influence of obstacles on the evacuation time». *Submitted to Journal of Engineering Mathematics*.

[3] F.-Z. Aït Bella, M. El Rhabi, A. Hakim et A. Laghrib. «Reduction of the non-uniform illumination by using nonlocal variational models for document image analysis». *Submitted to Journal-of-the-Franklin-institute*.

[4] A. Elmoussaoui, P. Argoul, M. El Rhabi et A. Hakim. «Discrete kinetic theory for 2D modeling of a moving crowd : Application to the evacuation of a non-connected bounded domain». *Computers & Mathematics with Applications* **75**, numéro 4 (2018), 1159 – 1180.

[5] «Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution». *Computer Vision and Image Understanding* **168** (2018), 50 – 63.

[6] D. El Morabit, M. El Rhabi, A. Hakim, A. Laghrib, et E. Moreau. «A new denoising model for multi-frame super-resolution image reconstruction». *Signal Processing* **132** (2017), 51–65.

[7] M. EL Rhabi, H. Fenniri, A. Ghazdali et A. Keziou. «Blind noisy mixture separation for independent/dependent sources through a regularized criterion on copulas». *Signal Processing* **131** (2017), 502–513.

[8] I. El Morabit, M. El Rhabi et A. Hakim. «Orthogonal projection algorithm for first and second order total variation denoising». *Annals of the University of Craiova - Mathematics and Computer Science Series* **42**, numéro 1 (2015), 140–147.

[9] A. Laghrib, A. Hakim, S. Raghay et M. El Rhabi. «Robust multi-frame super-resolution with non-parametric deformations using diffusion registration». *Annals of the University of Craiova - Mathematics and Computer Science Series* **42**, numéro 1 (2015), 117–128.

[10] A. Laghrib, A. Hakim, S. Raghay et M. El Rhabi. «A robust multi-frame super resolution based on curvature registration and second order variational regularization». *International Journal of Tomography & Simulation* **28**, numéro 2 (2015), 63–71.

[11] A. Laghrib, A. Hakim, S. Raghay et M. El Rhabi. «Robust super resolution of images with non-parametric deformations using an elastic registration». *Applied Mathematical Sciences* **8**, numéro 179 (2014), 8897–8907.

[12] S. Saoud, Z. Mahani, A. Hakim et M. El Rhabi. «Document scanning in a tough environment : application to cameraphones». *International Journal of Imaging & Robotics (IJIR), Special Issue on Practical Perspective of Digital Imaging for Computational Applications* **9**, numéro 1 (2013), 1–16.

[13] M. El Rhabi, H. Fenniri, A. Kéziou et E. Moreau. «A robust algorithm for convolutive blind source separation in presence of noise». *Signal Processing* **93**, numéro 4 (2013), 818–827.

[14] M. El Rhabi, A. Hakim, Z. Mahani et S. Saoud. «A General Formulation for Time-Harmonic Maxwell Equations in 3D Cavities and Approximation by a Spectral Method». *Advances in Dynamical Systems and Applications, ISSN 0973-5321* **8**, numéro 1 (2013), 13–23.

[15] M. Azaïez, F. Ben Belgacem, C. Bernardi et M. El Rhabi. «The mortar spectral element method in domains of operators. Part II : the curl operator and the vector potential problem». *IMA Journal of Numerical Analysis* **28**, numéro 1 (2008), 106–120.

[16] M. El Rhabi, H. Fenniri, G. Gelle et G. Delaunay. «Blind Separation of rotating machines signals using PMI criterion and Minimal Distorsion Principle». *MSSP Mechanical System and Signal Processing* **19**, numéro 6 (2005), 1282–1292.

[17] T.Z. Boulmezaoud et M. El Rhabi. «A mortar spectral method for $3D$ Maxwell's equations». *IMA Journal on Numerical Analysis* **25**, numéro 3 (2005), 577–610.

[18] M. El Rhabi, G. Gelle, H. Fenniri et G. Delaunay. «A penalized mutual information criterion for blind separation of convolutive mixtures». *Signal Processing* **83** (2004), 1979–1984.

9.2 Conférences internationales avec comité de lecture

[BEFM10] T.Z. Boulmezaoud, M. El Rhabi, H. Fenniri et E. Moreau, *On convolutive blind source separation in a noisy context and a total variation regularization*, The 11th IEEE International Workshop on Signal Processing Advances in Wireless Communications, SPAWC2010, http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5671015, 2010.

[BER10] T.Z. Boulmezaoud, M. El Rhabi et G. Rochefort, *A novel model for deblurring images*, SIAM Conference on Imaging Science (IS10) 2010, 2010.

[DER10] L. Dumas, M. El Rhabi et G. Rochefort, *A robust method for joint correction of nonuniform illumination and blind deconvolution of barcode signals*, SIAM Conference on Imaging Science (IS10) 2010, 2010.

[DER11] L. Dumas, M. El Rhabi et G. Rochefort, *An evolutionary approach for blind deconvolution of barcode images with nonuniform illumination*, IEEE Congress on Evolutionary Computation, CEC 2011, http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5949917, 2011.

[EFD06] M. El Rhabi, H. Fenniri et G. Delaunay, *Blind source separation using penalized mutual information criterion and minimal distortion principle*, International Symposium on Communications, Control and Signal Processing 2006, 2006.

[EFGD04] M. El Rhabi, H. Fenniri, G. Gelle et G. Delaunay, *Convolutive mixtures separation by penalized mutual information criterion. application to rotating machine diagnosis and monitoring*, 5th International Conference "Acoustical and Vibratory Surveillance Methods and Diagnostic Techniques", 2004.

[EFHM13] M. El Rhabi, H. Fenniri, A. Hakim et E. Moreau, *A new image deblurring approach using a special convolution expansion*, Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European, Sept 2013, pp. 1–5.

[EGFD04] M. El Rhabi, G. Gelle, H. Fenniri et G. Delaunay, *Convolutive blind source separation using a penalized mutual information criterion*, EUSIPCO 2004, Proceedings (CD-ROM and paper hard copy), 2004.

[EHM+12] M. El Rhabi, A. Hakim, Z. Mahani, K. Messou et S. Saoud, *A tool for scanning document-images with a photophone or a digicam*, Computer Applications for Communication, Networking, and Digital Contents (Tai-hoon Kim, Dae-sik Ko, Thanos Vasilakos, Adrian Stoica et Jemal Abawajy, eds.), Communications in Computer and Information Science, vol. 350, Springer Berlin Heidelberg, 2012, pp. 331–341 (English).

[MZS+12] Zouhir Mahani, Jalal Zahid, Sahar Saoud, Mohammed El Rhabi et Abdelilah Hakim, *Text enhancement by pde s based methods*, Image and Signal Processing (Abderrahim Elmoataz, Driss Mammass, Olivier Lezoray, Fathallah Nouboud et Driss Aboutajdine, eds.), Lecture Notes in Computer Science, vol. 7340, Springer Berlin Heidelberg, 2012, pp. 65–76 (English).

[TE03] T.Z. Boulmezaoud et M. El Rhabi, *On time-harmonic maxwell's equation in lipschitz and multiply-connected domain in $\mathbb{R}^3$*, Monografias del Seminario Matematico Garcia de Galdeano **27** (2003), 127–134.

9.3 Conférences nationales avec comité de lecture

[AE06] A. Ghorbel, R. Monneau et M. El Rhabi, *Comportement mécanique par homogénéisation de la dynamique des dislocations*, Colloque National MECAMAT - Aussois 2006, École de Mécanique des Matériaux, « Approches multi-échelles en mécanique des matériaux », Janvier 2006.

[JAHE17] A. Jebrane, P. Argoul, A. Hakim et M. El Rhabi, *Sur l'évolution d'un système de solides rigides, à nombre de degrés de liberté fin*, Congrès SMAI 2017, 8ème Biennale Française des Mathématiques Appliquées et Industrielles, 5 juin au 9 juin 2017.

[TE01] T.Z. Boulmezaoud et M. El Rhabi, *On time-harmonic maxwell's equation in lipschitz and multiply-connected domain in $\mathbb{R}^3$. theory and approximation by a spectral method*, 1er Congrès National de Mathématiques Appliquées et Industrielles, Mai 2001.

[TE02] T.Z. Boulmezaoud et M. El Rhabi, *A mortar spectral method for 3d maxwell's equations*, 34ème Congrès National d'Analyse Numérique (CANUM 2002), Mai 2002.

9.4 Brevets

[EMHS14] M. El Rhabi, Z. Mahani, A. Hakim et S. Saoud, *Procédé pour la restauration d'une image floue acquise au moyen d'une caméra numérique en vue de permettre la lecture et l'interprétation d'un motif inclus dans ladite image. OMPIC*

(Moroccan Industrial and Commercial Property Office) : MA 34997 B1. Cl. Internationale : G06t 5/00 ; g06k 9/56 ; g06k 9/62 ; h04n 5/243. Date de publication : 03.04.2014. n° Dépôt : 35219., Tech. report, Université Ibn Zohr-Agadir-Maroc, 2014.

[ER08] M. El Rhabi et G. Rochefort, *Method of restoring a blurred image acquired by means of a camera fitted to a communication terminal : Pub. no. :wo/2009/112710, Publication Date : 17.09.2009, IPC : G06T 5/00 (2006.01), G06K 9/56 (2006.01), G06K 9/62 (2006.01), H04N 5/243 (2006.01). Priority Data : 08 00693 11.02.2008 FR.*, Tech. report, Realeyes3d, SA, 2008.

9.5 Article en cours de rédaction

[AAHE] A.Jebrane, P. Argoul, A. Hakim et M. El Rhabi, *An overview on the non smooth dynamic of rigid bodies*, in progress.

[EAEH] A. Elmoussaoui, P. Argoul, M. El Rhabi et H. Hakim, *Discrete kinetic theory for 2d modeling of crowd motion : crowd characterized by a state of equilibrium*, in progress.

[EFM] M. El Rhabi, H. Fenniri et E. Moreau, *A quasi-newton method for a convolutive blind source separation*, in progress.

9.6 Prix

[Rea07a] Realeyes 3D team , *Mobile monday global peer awards : Un jury, composé d'experts et d'acteurs de la mobilité, a élu meilleure start up innovante realeyes 3d*, Prix, Realeyes3d, SA, 2007.

[Rea07b] Realeyes 3D team, *Le prix du red herring, ce prix classe les 100 startup européennes les plus prometteuses dans le monde*, Prix, Realeyes3d, SA, 2007.

Mes six dernières publications

A Traitement d'images et traitement du signal [3, 5, 6, 7]

Nous présentons trois articles de revues (acceptés ou soumis) en traitement d'images (restauration d'images par des méthodes variationnelles) et un article en traitement du signal (séparation aveugle de source). Pour chaque article, il est d'abord proposé un résumé en anglais, puis l'article dans son intégralité.

A.1 Reduction of the non-uniform illumination by using nonlocal variational models for document image analysis [3]

In this paper,we investigate two new reflectance and illumination decomposition models based on a nonlocal partial differential equation (PDE) applied to text images. Taking into consideration the higher regularity level of the illumination compared to the reflectance, we propose a high regular nonlocal PDE which deal with repetitive structures and textures that characterized the text image much better compared to the classical local PDEs. The aim of this approach is to use the repetitive features of the reflectance to efficiently extract it from the nonuniform illumination. This idea is motivated by extending the range of application of the nonlocal operators to such problem. Numerical experiments on both grayscale and color text images show the performance and strength of the proposed nonlocal PDE.

A.2 Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution [5]

The aim of a Super resolution (SR) technique is to construct a high-resolution image from a sequence of observed low-resolution ones of the same scene. One major roadblock of an SR reconstitution is removing noise and blur without destroying edges. We propose a novel multiframe image SR algorithm based on a convex combination of Bilateral Total Variation and a non-smooth second order variational regularization, using a controlled weighting parameter. We prove the existence of a minimizer of the proposed energy in the space of functions of bounded Hessian. The minimization of the convex functional is performed with a fast primal-dual algorithm. The simulation results and real experiments show the performance of the proposed algorithm in avoiding undesirable artifacts compared to other methods in the literature.

A.3 A new denoising model for multi-frame super-resolution image reconstruction [6]

Multi-frame image super-resolution (SR) aims to combine the sub-pixel information from a sequence of low resolution (LR) images to build a high-resolution (HR) one. SR techniques usually suffers from annoying restoration artifacts such as noise, jagged edges, and staircasing effect. In this paper, we aim to increase the performance

of SR reconstitution under a variational framework using adaptive diffusion-based regularization term. We propose a new tensor based diffusion regularization that takes the benefit from the diffusion model of Perona-Malik in the flat regions and use a nonlinear tensor derived from the diffusion process of Weickert filter near boundaries. Thus, the proposed SR approach can preserve important image features (sharp edges and corners) much better while avoiding artifacts. The synthetic and real experimental results show the effectiveness of the proposed regularization compared to other methods in both quantitatively and visually.

A.4 Blind noisy mixture separation for independent/dependent sources through a regularized criterion on copulas [7]

The paper introduces a new method for Blind Source Separation (BSS) in noisy instantaneous mixtures of both independent or dependent source component signals.This approach is based on the minimization of a regularized criterion.Precisely, it consists in combining the total variation method for denoising with the Kullback-Leibler divergence between copula densities.The latter takes advantage of the copulato model the structure of the dependence between signal components.The obtained algorithm achieves separation in a noisy context where standard BSS methods fail.The efficiency and robustness of the proposed approach are illustrated by numerical simulations.

Reflectance and illumination decomposition using nonlocal models

Ait Bella Fatim Zahra[a,*], Mohammed El Rhabi [b], Abdelilah Hakim[a], Amine Laghrib[c]

[a] *Cadi Ayyad University, LAMAI FST Marrakech, Morocco;*
[b] *Applied Mathematics and Computer Science department, Ecole des Ponts ParisTech (ENPC), Paris, France.*
[c] *LMA FST Béni-Mellal, Sultan Moulay Slimane University, Morocco*
[*] *Correspondig author : aitbella.fatimzehrae@gmail.com*

Abstract

In this paper, we investigate two new reflectance and illumination decomposition models based on a nonlocal partial differential equation (PDE) applied to text images. Taking into consideration the higher regularity level of the illumination compared to the reflectance, we propose a nonlocal PDE which deals with repetitive structures and textures that characterize the text image much better compared to the classical local PDEs. The aim of this approach is to use the repetitive features of the reflectance to efficiently extract it from the nonuniform illumination. This idea is motivated by extending the range of application of the nonlocal operators to such problem. Numerical experiments on both grayscale and color text images show the performance and strength of the proposed nonlocal PDE.

Keywords: Nonlocal operators, Partial differential equation, Text image, Viscosity solutions theory

1. Introduction

In image analysis, reflectance and illumination decomposition is a basic process which aims to enhance the image by reducing or removing the effects of the nonuniform illumination. This technique, which is named the Retinex problem, represents a necessary pretreatment step in a wide range of image processing applications.

The image decomposition is considered as an ill-posed problem since it consists in extracting two unknown parts from an observed image. The Retinex theory was first proposed by Land and McCann in [27] to address and improve the quality of an image when the lighting conditions are not satisfactory [26, 33]. The main idea of this theory is to consider the human visual system as a color perception where we can define the reflectance associated to a field in which both illumination and reflectance are unknown. Indeed, for a given scene, our visual system detects and sees the same color even if the illuminations conditions vary. This explains that the color of the objects is considered invariant in spite of illumination transformations.

The main goal of Retinex theory is to decompose a given image u into two components represented by the following model:

$$u(x,y) \;=\; I(x,y)R(x,y). \tag{1}$$

The model assumes that the observed image u is the product of two main components [18], namely the illumination I and the reflectance image R. According to the single-scale Retinex algorithm, the reflectance is estimated as described by the following equation:

$$\log(R_i(x,y)) = \log(u_i(x,y)) - \log(F(x,y) * u_i(x,y)), \tag{2}$$

where $*$ is the convolution operator, R_i is the reflectance image on the ith channel and F is a Gaussian filter defined by

$$F(x,y) = C\exp^{-(\frac{x^2+y^2}{\sigma^2})}, \tag{3}$$

where C is a normalization factor.
This single-scale Retinex formula is extended to a multi-scale version in [24, 40]. This approach is pursued and introduced by Brelstaff [4] and Horn [19], who proposed a Retinex algorithm based on Poisson equation. In 2003, Kimmel and Elad proposed a variational Retinex approach [25]; to form a penalty functional, the authors call into question some hypotheses about the illumination image: the illumination is spatially smooth, it is close to the intensity image u and $I \geq u$, moreover they assume that the illumination continues smoothly as a constant beyond the image boundaries. Other approaches are presented using the same principle. First, Ma and Osher have introduced a total variation regularized formulation and an efficient algorithm based on the split Bregman method, then, they took advantage of

the nonlocal total variation to enhance the resulting images [31]. Further, Chen et al have used a TV-L1 based variational Retinex approach [31] . More recently, Liang and Zhang suggested a reflectance and illumination decomposition model for the Retinex problem using a convex variational model based on high-order total variation and L^1 decomposition [28]. This method can effectively decompose the gradient component of images into salient edges and slightly smoother illumination field.

Another famous method to correct the non-uniform illumination is the Homomorphic filtering. This approach was introduced by Oppenheim et Al [37], the main idea of this technique is to decompose the image into high and low frequencies components, as a result the enhanced image is obtained through a high-pass filter in order to remove the illumination component. The improvement is then achieved by the simultaneous compression of the intensity range and the contrast enhancement. This method gave reasonable results and have been used in many applications [12, 10, 17].

On the other hand and precisely in document-image processing, many approaches are proposed to enhance the quality of a document-image. This generates a text which is easier to read in order to extract it thereafter using an optical character recognition (OCR). In 1995, W. Jiang has investigated a thresholding and enhancement method [22] which consists of converting text images of low spatial resolution to bi-level images of higher spatial resolution. In the same principle, a simple algorithm based on empirical mode decomposition was proposed in [39] to improve the illumination problem for document-image. However, there are several other methods specified for document-image enhancement with different degrees of success, see for examples [41, 7, 23, 11].

The main objective of this paper is to provide an improved nonlocal approach to remove the shadow from a document-image. Nonlocal operators have received much attention in the past decade. Such operators were first introduced by Gilboa and Osher [16]. The purpose of their paper is the use of nonlocal operators to treat signal and image processing by adopting the so called nonlocal PDE. Convinced by the effectiveness and efficiency of the nonlocal operators in dealing better with image textures, we introduce two new nonlocal PDEs to enhance the text quality and make it more meaningful and expressive for the human eye. These two nonlocal PDEs take into consideration the fact that illumination varies more smoothly than reflectance and also consider the effect of the illumination on image features such as texture. This can help the model to be able to take into account the surroundings of

the object much better compared to some classical approaches.

The paper is organized as follows. In section 2, we will focus on introducing a local model which we aim to extend to the nonlocal case, we recall some essential and basic definitions and properties of viscosity solutions theory and we prove the existence of a unique solution of this problem, section 3 deals with the formulation of the new nonlocal models, in section 4, two algorithms are presented to solve the proposed nonlocal models, and in order to assess the validity of the proposed approach we will present some numerical tests in section 5, finally we shall sum up by a conclusion where we will outline the big points covered in the paper.

2. Problem Formulation

We consider a document-image u acquired by a camera phone. In general, many degradation factors are involved by this acquisition, such as : noise, non-uniform illumination and motion blur. Note that in this paper, the blur is not considered. In fact, the studied model is based on the separation of the two components of the image: the illumination I and the reflectance R

$$u(x,y) = I(x,y)R(x,y) + n(x,y), \tag{4}$$

where n is the noise. As discussed above, we firstly neglect the noise and we apply the logarithm function, we obtain then the additive model

$$\log(u(x,y)) = \log(I(x,y)) + \log(R(x,y)). \tag{5}$$

Our aim is to consider a decomposition approach to recover the illumination I by estimating the term " $\log I$". In particular, we focus on extracting illumination I using a nonlocal PDE in a well posed functional framework. Our proposed model is considered as a nonlocal version of the PDE proposed in [32]. The idea behind this choice is related to recently developed nonlocal operators [16]. Therefore, in the following, we recall the studied PDE and also we present some theoretical results. The local PDE is defined by

$$\begin{cases} \dfrac{\partial u}{\partial t} &= \max(0, \Delta u) \quad \text{in } \Omega \times (0,T) \\ \dfrac{\partial u}{\partial n} &= 0 \qquad\quad \text{on } \partial\Omega \times (0,T) \\ u(.,0) &= u_0 \qquad\quad \text{in } \Omega \end{cases} \tag{6}$$

where Ω is a bounded open subset of $\mathbb{R}^2$ and Δ is the Laplacian operator defined by

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}. \tag{7}$$

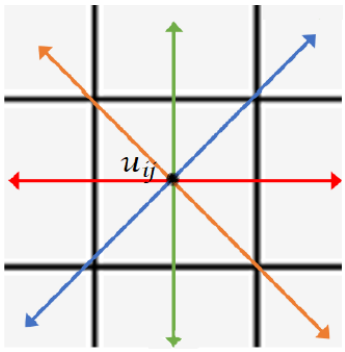
The method proposed in [32] which aims to resolve the equation (6) supposes that the document-image u has been acquired by a cameraphone. The authors consider two problem types; namely: image distortions and noise. Especially, they address the so-called non uniform illumination or variations of brightness. In image processing applications, a shadow is considered as a region with low lightness and high gradient contours. There are in fact two kinds of shadows, the own shadow and the shadows arising from the acquisition process. Own shadow occurs when the light hits a surface with a slope change. Otherwise, the brightness of pixels decreases when the angle of incidence deviates from the normal of the surface. While the brightness reaches its minimum value in the case when the incident light and the surface normal are orthogonal. A shadow always occurs when the light source is obscured by an object before the light reflection on the surface. Taking into account that the image contains only text, the background is preponderant compared to the text in the image (or there is much more background than text in the image). Also knowing that the intensity of the background is superior to the intensity of the text (light background on dark writing), it is easy to approximate the background. Indeed, intuitively, in order to estimate the background, we replace each pixel by considering the maximum mean of its neighboring pixels and we iterate that is to say if u_{ij}^n represents the intensity of the image in the spatial location (i, j) at the n^{th} iteration, we get:

$$u_{ij}^{n+1} = \max\left(u_{ij}^n, \widehat{u}_{ij}^{1,n}, \widehat{u}_{ij}^{2,n}, \widehat{u}_{ij}^{3,n}, \widehat{u}_{ij}^{4,n}\right), \tag{8}$$

where $\widehat{u}^k$ is the mean calculated in the k^{th} direction using the neighboring pixels (see Fig.1). By subtracting u_{ij}^n from both sides of the equation (8), the discrete derivatives in all directions appear and we get the discretized version of the equation:

$$\frac{\partial u}{\partial t} = \frac{1}{2}\max\left(0, \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial y^2}, \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + 2\frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - 2\frac{\partial^2 u}{\partial x \partial y}\right) \tag{9}$$

in a similar way, considering only the horizontal and vertical directions, one can easily get the equation (6).



$$\widehat{u}_{ij}^{1,n} = \frac{u_{i+1j}^n + u_{i-1j}^n}{2}$$

$$\widehat{u}_{ij}^{2,n} = \frac{u_{ij+1}^n + u_{ij-1}^n}{2}$$

$$\widehat{u}_{ij}^{3,n} = \frac{u_{i+1j+1}^n + u_{i-1j-1}^n}{2}$$

$$\widehat{u}_{ij}^{4,n} = \frac{u_{i-1j+1}^n + u_{i+1j-1}^n}{2}$$

Figure 1: The mean at the neighboring pixels of the image u.

Realeyes 3D company and M. El Rhabi have successfully implemented this equation as an embedded application for document scanning in more than 150 million smartphones in the world from 2006 to 2010. In this paper, as a first step, we choose a simplified form of this equation which concides with equation (6). Then, we rewrite (9) in its non local formulation which we analyse.

The equation (6) has demonstrated its robustness in the extraction of the variable luminance and it has given quite satisfactory results (see [32]). However, the theoretical aspect of this equation has not been investigated, which left a lack of information on the behavior and the existence of the solution. As a first contribution in this work, we are interested in the theoretical side of this model. The complexity to show the existence of a solution to the equation (6) comes from the fact that we cannot say anything about the maximum of two operators. We intend here to find a continuous viscosity solution of (6). We first give a brief overview of some essential notions of the viscosity solutions theory [8, 15, 2].
The theory deals with some partial differential equations of the form :

$$F(x, u, \nabla u, \nabla^2 u) = 0, \tag{10}$$

where x is defined in an open subset Ω of $\mathbb{R}^n$. F is a real-valued function on $D = \Omega \times \mathbb{R} \times \mathbb{R}^n \times S^n$, where S^n stands for the set of symmetric $n \times n$ matrices, u denotes the real-valued unknown function on Ω, ∇u and $\nabla^2 u$ designate respectively, the gradient and the Hessian matrix of u.
The function F is called degenerate elliptic if it verifies the following inequality:

$$F(x, r, p, X) \leq F(x, r, p, Y) \quad \text{for } Y \leq X, \tag{11}$$

where $Y \leq X$ signifies:

$$< X\xi, \xi > \leq < Y\xi, \xi > \quad \forall \xi \in \mathbb{R}^n. \tag{12}$$

If F is degenerate elliptic, we can say that it is proper if:

$$F(x, r, p, X) \leq F(x, s, p, Y) \quad \text{for } Y \leq X, \;\; r \leq s. \tag{13}$$

Let $u : E \to \mathbb{R} \cup \{\pm\infty\}$, where E is a set of a metric space. The upper semi-continuous envelope u^*, and the lower semi-continuous envelope u_* of u are given respectively by

$$u^*(y) = \lim_{r \to 0} \sup\{u(z)/\ z \in B(y, r) \cap E\},$$

$$u_*(y) = \lim_{r \to 0} \inf\{u(z)/\ z \in B(y, r) \cap E\},$$

where $B(y, r)$ is a closed ball of radius r centred at $y \in \overline{E}$.

Definition 1. *[15] Let Ω be an open set in $\mathbb{R}^n$ and $T > 0$ a fixed time parameter.*
Let $F : \overline{\Omega} \times [0, T] \times \mathbb{R} \times \mathbb{R}^n \times S^n \to \mathbb{R}$ be a continuous function.
Let U be an open set in $\Omega \times]0, T[$.
A function $u : U \to \mathbb{R} \cup \{-\infty\}$ is a viscosity sub-solution of

$$\frac{\partial u}{\partial t} \; + \; F(x, t, u, \nabla u, \nabla^2 u) = 0, \tag{14}$$

in U if:
- *$u^*(x, t) < \infty$ for $(x, t) \in U$.*
- *If $(\varphi, \hat{x}, \hat{t}) \in \mathcal{C}^2(U) \times U$ satisfies*

$$\max_U (u^* - \varphi) \; = \; (u^* - \varphi)(\hat{x}, \hat{t}). \tag{15}$$

then

$$\frac{\partial \varphi}{\partial t}(\hat{x}, \hat{t}) \; + \; F(\hat{x}, \hat{t}, u^*(\hat{x}, \hat{t}), \nabla\varphi(\hat{x}, \hat{t}), \nabla^2\varphi(\hat{x}, \hat{t})) \; \leq \; 0. \tag{16}$$

Definition 2. *[15] A function $u : U \to \mathbb{R} \cup \{+\infty\}$ is a viscosity supersolution of (14) in U, if*
- *$u_*(x, t) > -\infty$ for $(x, t) \in U$.*
- *If $(\varphi, \hat{x}, \hat{t}) \in \mathcal{C}^2(U) \times U$ satisfies*

$$\min_U (u_* - \varphi) \; = \; (u_* - \varphi)(\hat{x}, \hat{t}), \tag{17}$$

then

$$\frac{\partial \varphi}{\partial t}(\hat{x}, \hat{t}) \; + \; F(\hat{x}, \hat{t}, u_*(\hat{x}, \hat{t}), \nabla\varphi(\hat{x}, \hat{t}), \nabla^2\varphi(\hat{x}, \hat{t})) \; \geq \; 0 \tag{18}$$

Definition 3. *[15] A continuous function $u : U \to \mathbb{R} \cup \{\pm\infty\}$ is a viscosity solution of (14) in U, if it is both a viscosity subsolution and a viscosity supersolution of (14) in U.*

We apply theorem 3.1 in [21] to get the following existence result:

Theorem 1. *Let Ω be a regular bounded open subset of $\mathbb{R}^2$.*
Let $u \in USC([0,T[\times\overline{\Omega})$ (upper semicontinuous), and $v \in LSC([0,T[\times\overline{\Omega})$ (lower semicontinuous), be, respectively, a viscosity sub and a supersolutions of

$$\frac{\partial u}{\partial t} - \max(0, \Delta u) \;=\; 0 \;\; in \;\; \Omega\times]0,T[\tag{19}$$

and

$$\frac{\partial u}{\partial n} \;=\; 0 \qquad\qquad on \;\; \partial\Omega\times]0,T[\tag{20}$$

If

$$u(x,0) \leq v(x,0) \quad for \;\; x \in \overline{\Omega} \tag{21}$$

then

$$u \leq v \quad in \;\;]0,T[\times\overline{\Omega}$$

Moreover, for each $u_0 \in C(\overline{\Omega})$, there exists a unique continuous viscosity solution $u \in C([0,T[\times\overline{\Omega})$ of (6).

Proof
First of all, we use the following decomposition of the max operator

$$\max(0, \Delta u) \;=\; H(\Delta u)\Delta u \tag{22}$$

where H is the Heaviside step function, defined by

$$\forall\, x \;\in \mathbb{R}, \quad H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases} \tag{23}$$

Afterwards, we consider the function $F : S^n \to]-\infty, 0]$, given by:

$$F(X) \;=\; -H(trace(X))trace(X). \tag{24}$$

According to the theorem 2.1 in [21], we have to prove that F is continuous and there exists a continuous function $w : [0, \infty[\rightarrow [0, \infty[$ satisfying $w(0) = 0$, such that if $X, Y \in S^n$ and $\mu_1, \mu_2 \in [0, \infty[$ satisfy

$$\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \leq \mu_1 \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \mu_2 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \tag{25}$$

then

$$F(X) - F(-Y) \geq -w(\mu_2). \tag{26}$$

We start then by the continuity which is evident, since the functions

$$\begin{array}{rccc} f & : & S^n & \rightarrow & \mathbb{R} \\ & & X & \mapsto & trace(X) \end{array}$$

and

$$\begin{array}{rccc} g & : & \mathbb{R} & \rightarrow &]-\infty, 0] \\ & & x & \mapsto & -H(x)x \end{array}$$

are continuous. Now, we suppose that there exist $\mu_1, \mu_2 \in [0, \infty[$ such that the inequality (25) is satisfied.

We multiply both sides of the inequality by the matrix $\begin{pmatrix} I & I \\ I & I \end{pmatrix}$,

We then apply the trace operator, we obtain

$$trace(X) + trace(Y) \leq 2n\mu_2. \tag{27}$$

Afterwards, taking the decreasing function $g(x) = -H(x)x$ yields

$$\begin{array}{rcl} -H(2n\mu_2)2n\mu_2 & = & -2n\mu_2 \\ & \leq & -H(trace(X) + trace(Y))(trace(X) + trace(Y)) \\ & \leq & \frac{1}{2}(-trace(X) - trace(Y) - |trace(X) + trace(Y)|) \\ & \leq & \frac{1}{2}(-trace(X) - trace(Y) - |trace(X)| + |trace(Y)|) \\ & = & F(X) - F(-Y). \end{array} \tag{28}$$

Finally, we take $w(r) = 2nr$ which satisfies $w(0) = 0$.
According to the theorem 2.1 and theorem 3.1 in [21], the results of comparison and existence of a unique viscosity solution is reached. □

As demonstrated in [32], the used PDE (6) has given good results, but does not preserve well the fine details of the text. This can be interpreted

by the use of a local image reconstruction algorithm using observations in a neighborhood of a pixel of the interest. However, in some cases the selection of the neigborhood plays a major role in the quality of the image reconstitution. This led us to think to the nonlocal version of the equation (6). In the nonlocal techniques, algorithms analyse data in a larger neigborhood and collect the observations from the whole image, searching similar features, which is much better and accurate. Indeed, the nonlocal methods, compared to local methods, are more effective, in particular to restore fine and repetitive structures of the image, which is the case of document-images. These approaches which are called also patch methods, allow to characterize a pixel with a vector of attributes wider than the intensity or color value of the pixel only. A patch representing a pixel is considered as a square sub-image of the original image, centered at this pixel, of a selected width. The patch-based methods were first used in the texture synthesis [9]. Also, they were then used for filtering and restoring degraded image [5]. More recently, it have also been extended to inpainting problems [16]. In the following section, we introduce the proposed nonlocal model to extract the illumination I and we give a result of the existence and uniqueness of the solution.

3. The proposed nonlocal approaches

In this section, we introduce the two nonlocal PDEs and we give some theoretical results. We start by the definition of the first model.

3.1. The first model

As discussed above, we propose a new model for document-image enhancing. Therefore, we intend to obtain a new nonlocal generalization of the local model presented in [32]. As a preliminary, we recall some definitions of nonlocal operators introduced by Gilboa and Osher [16].
Let $\Omega \subset \mathbb{R}^2$, and $u : \Omega \to \mathbb{R}$ denote a real function.
The nonlocal gradient $\nabla_w u(x)$ at $x \in \Omega$ is defined as the vector of all partial derivatives $\nabla_w u(x,.)$:

$$(\nabla_w u)(x,y) := (u(y) - u(x))\sqrt{w(x,y)}, \quad y \in \Omega, \tag{29}$$

where w stands for a symmetric and positive weight function.
For any pair of $v_1, v_2 : \Omega \times \Omega \to \mathbb{R}$, one defines

$$(v_1.v_2)(x) := \int_\Omega v_1(x,y)v_2(x,y)dy, \tag{30}$$

and the inner product

$$< v_1, v_2 > := \int_{\Omega\times\Omega} v_1(x,y)v_2(x,y)dxdy. \tag{31}$$

With the above definition of inner product and nonlocal gradient, the nonlocal divergence of $v : \Omega \times \Omega \to \mathbb{R}$ at x is defined as the negative adjoint of the nonlocal gradient:

$$(\mathrm{div}_w\, v)(x) \ := \ \int_\Omega (v(x,y) - v(y,x))\sqrt{w(x,y)}dy. \tag{32}$$

The nonlocal Laplacian of $u : \Omega \to \mathbb{R}$ is then defined by:

$$(\Delta_w u)(x) \ := \ \frac{1}{2}\,\mathrm{div}_w((\nabla_w u)(x)) = \ \int_\Omega (u(y) - u(x))w(x,y)dy. \tag{33}$$

At this point, we can use these operators to introduce the proposed nonlocal PDE. Indeed, to estimate the variable illumination we introduce the following model:

$$\begin{cases} \dfrac{\partial u}{\partial t} \ = \ \max(0, \Delta_w u) \ \text{ in } \ \Omega \times [0, \infty[\\ u(.,0) = u_0 \ \text{ in } \ \Omega \end{cases} \tag{34}$$

where $u_0 \in L^2(\Omega)$ is a given degraded image. This problem is seen as an evolution equation with homogeneous Neumann boundary condition, which mean that when integrating in Ω, we suppose that there is no flux across the boundary. Before resolving this PDE, we have to check the existence and uniqueness of the solution. We are now ready to state our main result

Theorem 2. *Let Ω be a bounded open subset of $\mathbb{R}^2$. Then, the problem (34) admits a unique generalized global solution $u \in \mathcal{C}([0, +\infty[; L^2(\Omega))$.*

Proof

Basically, we use the following decomposition of the max operator

$$\max(0, \Delta_w u) = \frac{1}{2}(\Delta_w u + \ | \ \Delta_w u \ |) \tag{35}$$

According to the definition of $\Delta_w u$ and (35), we consider the bounded linear operator A given by

$$Au = \frac{1}{2}u(.) \int_\Omega w(.,y)dy. \tag{36}$$

The operator A is the infinitesimal generator of $S(t) = e^{-tA},\ t \geq 0$. On the other hand, the function:

$$f(u) = \frac{1}{2}(\int_\Omega u(y)w(.,y)dy + |\int_\Omega (u(y) - u)w(.,y)dy|) \tag{37}$$

is an M-Lipschitz continuous function, indeed $\forall t > 0,\ \forall u, v \in L^2(\Omega)$:

$$\begin{aligned} | f(u(t)) - f(v(t)) |_{L^2(\Omega)} &\leq | \Omega | | w |_{L^\infty(\Omega\times\Omega)} (2 | u - v |_{L^1(\Omega)} + | u - v |_{L^2(\Omega)}) \\ &\leq | \Omega | | w |_{L^\infty(\Omega\times\Omega)} (2c + 1) | u - v |_{L^2(\Omega)} \end{aligned}$$

We introduce now the operator ψ given by:

$$\psi(u)(t) = S(t)u_0 + \int_0^t S(t - s)f(u(s))ds. \tag{38}$$

We shall show that ψ maps

$$E_\alpha = \{u \in \mathcal{C}([0, \infty[; L^2(\Omega)); | u |_{E_\alpha} < +\infty\}$$

equipped with the norm $| u |_{E_\alpha} = \sup_t(e^{-\alpha t} | u(t) |_{L^2(\Omega)})$ into $E_\alpha\ \forall \alpha > 0$.
E_α is a complete space since it is a closed subset of $\mathcal{C}([0, \infty[; L^2(\Omega))$.
we have

$$\begin{aligned} \forall t > 0,\ | \psi(u)(t) |_{L^2(\Omega)} &\leq | S(t)u_0 |_{L^2(\Omega)} + \int_0^t \| S(t - s) \|_{\mathcal{L}(L^2(\Omega))} | f(u(s)) |_{L^2(\Omega)}\, ds \\ &\leq | u_0 |_{L^2(\Omega)} + \int_0^t | f(u(s)) |_{L^2(\Omega)}\, ds \\ &\leq | u_0 |_{L^2(\Omega)} + \int_0^t (| f(u(s)) - f(0) |_{L^2(\Omega)} + | f(0) |_{L^2(\Omega)})ds \\ &\leq | u_0 |_{L^2(\Omega)} + \int_0^t (M | u(s) |_{L^2(\Omega)} + C)ds. \end{aligned} \tag{39}$$

We multiply this last inequality by $e^{-\alpha t}$ for $t > 0$. We obtain then :

$$\begin{aligned} \forall t > 0;\ e^{-\alpha t} | \psi(u)(t) |_{L^2(\Omega)} &\leq e^{-\alpha t} | u_0 |_{L^2(\Omega)} + \int_0^t e^{-\alpha(t-s)}e^{-\alpha s}(M | u(s) |_{L^2(\Omega)} + C)ds \\ &\leq e^{-\alpha t} | u_0 |_{L^2(\Omega)} \\ &+ M \sup_s (e^{-\alpha s} | u(s) |_{L^2(\Omega)}) \int_0^t e^{-\alpha(t-s)}ds + tCe^{-\alpha t} \\ &\leq | u_0 |_{L^2(\Omega)} + \frac{1}{\alpha}M | u |_{E_\alpha} + C \sup_t (te^{-\alpha t}). \end{aligned} \tag{40}$$

Thus

$$\sup_{t>0} e^{-\alpha t} \mid \psi(u)(t) \mid_{L^2(\Omega)} < +\infty. \tag{41}$$

Therefore, we can define the function ψ as follows, $\forall \alpha > 0$:

$$\psi : E_\alpha \to E_\alpha$$

Now, we shall show that ψ is a contraction on E_α: let $u, v \in E_\alpha$ and $t \geq 0$, we have

$$\begin{aligned} \mid \psi(u)(t) - \psi(v)(t) \mid_{L^2(\Omega)} &= \mid \int_0^t S(t-s)(f(u(s)) - f(v(s)))ds \mid_{L^2(\Omega)} \\ &\leq M \int_0^t \mid u(s) - v(s) \mid_{L^2(\Omega)} ds. \end{aligned} \tag{42}$$

Hence

$$\forall t > 0, \quad e^{-\alpha t} \mid \psi(u)(t) - \psi(v)(t) \mid_{L^2(\Omega)} \leq \frac{M}{\alpha} \mid u - v \mid_{E_\alpha} . \tag{43}$$

This shows that the operator ψ is a contraction on E_α if $\alpha > M$. Consequently, using the Banach fixed-point theorem [38] we deduce the existence of a unique fixed point of ψ in E_α. This fixed point is the desired solution of the integral equation

$$u(t) = S(t)u_0 + \int_0^t S(t-s)f(u(s))ds \tag{44}$$

Which concludes the proof. □

3.2. The second model

We conclude this section by presenting another new nonlocal model using the nonlocal second derivatives. The main idea is to take into consideration the second order opposite derivatives instead of considering only the nonlocal Laplacian operator. In fact, we consider all the directions presented in Fig.1. This technique is inspired by the benefits of considering the different directions in a document-image to handle the complexity of text texture in the degraded image. In [16], Gilboa and Osher derived also the nonlocal second

order derivatives. Let $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ be a function, and $w : \Omega \times \Omega \to \mathbb{R}$ be a nonnegative and symmetric weight function, then for $i = 1, 2$:

$$\begin{aligned}\left(\frac{\partial^2 u}{\partial x_i^2}\right)_w &:= \nabla_w(\nabla_w(u).\nabla_w(x_i)).\nabla_w(x_i) \\ &= \nabla_w(\int_\Omega (u(y) - u(x))w(x,y)(y_i - x_i)dy).\nabla_w(x_i) \end{aligned} \quad (45)$$

Following the same idea, we establish the following text enhancement model to extract the variable illumination from an image:

$$\begin{cases} \frac{\partial u}{\partial t} &= \max(0, (\frac{\partial^2 u}{\partial x_1^2})_w, (\frac{\partial^2 u}{\partial x_2^2})_w, (\frac{\partial^2 u}{\partial x_1^2})_w + (\frac{\partial^2 u}{\partial x_2^2})_w + 2(\frac{\partial^2 u}{\partial x_1 x_2})_w, (\frac{\partial^2 u}{\partial x_1^2})_w \\ & \quad +(\frac{\partial^2 u}{\partial x_2^2})_w - 2(\frac{\partial^2 u}{\partial x_1 x_2})_w) \ \text{ in } \ \Omega \times (0,T) \\ u(,0) &= u_0 \qquad\qquad \text{in } \ \Omega \end{cases} \quad (46)$$

In the next section, we describe briefly the proposed algorithms for solving the problems (34) and (46).

4. Numerical implementation

We begin by defining the weight function to measure the similarities between the pixels which is generally based on a gray level intensity vector comparison between two pixels. A pixel is characterized by its neighboring pixels belonging to a neighborhood window centered at its position, and the weight is calculated using the difference between these patches. Indeed, several forms of weight functions can be used [43, 6]. Note that the choice of the weight function is crucial for the success of non-local methods. In this paper, we propose a non-negative real-valued and symmetric weight function depending on the spatial distance only:

$$w(x,y) = exp\left(\frac{-\mid x - y \mid^2}{h}\right), \quad (47)$$

where $u : \Omega \to \mathbb{R}$ is a given image, $\Omega \in \mathbb{R}^2$ is a bounded open domain and h is a positive parameter.
Before introducing the two proposed algorithms, we consider the following discretization of the nonlocal Laplacian operator:

$$\triangle_w(u_i) = \sum_{j \in N_i} (u_j - u_i)w_{i,j} \quad j \in N_i, \quad (48)$$

where u_i is the value of a pixel i $(1 \leq i \leq N)$, $N_i = \{j : |i-j| \leq r\}$ denotes the neighbors set of the pixel i, and $w_{i,j}$ is the discretized form of the weight function :

$$w_{i,j} = \exp\left(\frac{-\mid i-j \mid^2}{h}\right). \tag{49}$$

To discretize the non-local second derivative operators, we give the discretized version of the non-local gradient:

$$\nabla_w(u_i) = (u_j - u_i)\sqrt{w_{i,j}} \quad j \in N_i, \tag{50}$$

and the discrete inner product for vectors is given by:

$$(v_1.v_2)_i \ := \sum_{j \in N_i} (v_{1_{i,j}} v_{2_{i,j}}) \quad j \in N_i. \tag{51}$$

The second order discretized nonlocal derivatives are given by:

$$(u_i)''_{k,l} = \nabla_w(\nabla_w(u_i).\nabla_w(i_k)).\nabla_w(i_l) \quad k,l \in \{1,2\}, \tag{52}$$

where i_k is the k^{th} component of the pixel i and

$$\nabla_w(i_k)(j) = (j_k - i_k)\sqrt{w_{i,j}} \quad j \in N_i. \tag{53}$$

Now we are ready to implement the two proposed algorithms based on the discretization form of non-local operators given above. We give firstly the algorithm 1 for solving the problem (34) as follows:

Algorithm 1 for solving (34)

Input : The acquired image u
initialization : We set $I^0 = \log(u+1)$ and choose $dt > 0$, $h > 0$
Compute :
$I_i^{n+1} = I_i^n + dt \max(0, \Delta_w(I_i^n)) \quad i = 1, ..., N$
Output : The illumination I.
The reflectance $R = \exp(\log(u+1) - I)$

In the same way, the algorithm 2 is given for resolving the second model in (46) using the nonlocal second order derivatives instead of the nonlocal Laplacian operator:

Algorithm 2 for solving (46)

Input : The acquired image u
initialization : We set $I^0 = \log(u+1)$ and choose $dt > 0$, $h > 0$
Compute :
$I_i^{n+1} = I_i^n + dt \max(0, (I_i^n)''_{1,1}, (I_i^n)''_{2,2}, (I_i^n)''_{1,1} + (I_i^n)''_{2,2} + 2(I_i^n)''_{1,2},$
$(I_i^n)''_{1,1} + (I_i^n)''_{2,2} - 2(I_i^n)''_{1,2}) \quad i = 1, ..., N$
Output : The reflectance image $R = \exp(\log(u+1) - I)$

5. Numerical results

This section is devoted to the experimental part to test the two proposed nonlocal algorithms. Our aim is to confirm the performance and the robustness of the nonlocal models compared to some local methods in extracting the reflectance of a document-image.
The considered images in the reflectance and illumination decomposition problem are assumed to be degraded images acquired under variable illumination. For that, we consider six document-images taken in low light conditions. The first four images are given in grayscale level while the two last ones are color document-images. We compare the two nonlocal algorithms with some available methods in the literature, namely the variational method proposed in [25] (called Variational Retinex), the multiscale Retinex approach [40] (MSRetinex) and two Retinex methods based on partial differential equations proposed in [34, 29] (Screened Poisson Equation) and (PDE Retinex) respectively.
Figs. 2-7 show the obtained results using the two proposed algorithms compared to the other methods. Throughout this section, for the proposed methods, we set $h = 80$, while the numerical results by the variational Retinex method are reported with $\alpha = 0.0001$ and $\beta = 0.1$, we set the threshold parameter $t = 4$ for the PDE-Retinex model, and $\lambda = 0.0001$ the tradeoff parameter for the Screened Poisson Equation, while the parameters in the implementation of Multiscale Retinex are chosen as follows:

parameter	N	σ_1	σ_2	σ_3	α	β	w_n	G	b
value	3	15	80	250	125	46	1/3	192	-30

Note that for the first four grayscale images, we consider also a binarization step to show better the estimated text. In the following, we investigate a quantitative comparison in terms of OCR using the restored images in the previous tests apart the handwriting images. Optical Character Recognition tools are efficiently used to convert an image file or PDF file into editable and searchable text format. However, there are many limitations of OCR that can lead to inaccurate or missing text, the OCR engines fail usually to take into account blurry documents, colored paper, mathematical formulas or handwritten text.
The table 1 compares the character recognition rates for the grayscale images

using three different types of OCR tools, the "Abbyy FineReader 12" [1], the software "Tesseract" [2] and the free web-based optical character recognition software "OnlineOCR.net" [3] that return a text with the best confidence rating. The rate "%" is obtained by a careful calculation of the correct characters from the OCR output. The table proves, as might be expected, and identically to the visual comparison in the six different examples, that the two non-local approaches can better recover the degraded text compared to the other approaches, providing then promising numerical results. In comparison, considering the greyscale images, we established that the two proposed methods can reduce the effects of the nonuniform illumination to a greater degree than the other methods. The model based on the nonlocal second derivatives seemed to enhance better color document-images than the first nonlocal model. Our two nonlocal methods, while different, performed comparably at estimating the reflectance part.

[1] https://en.wikipedia.org/wiki/ABBYY_FineReader
[2] https://en.wikipedia.org/wiki/Tesseract_(software)
[3] https://www.onlineocr.net/

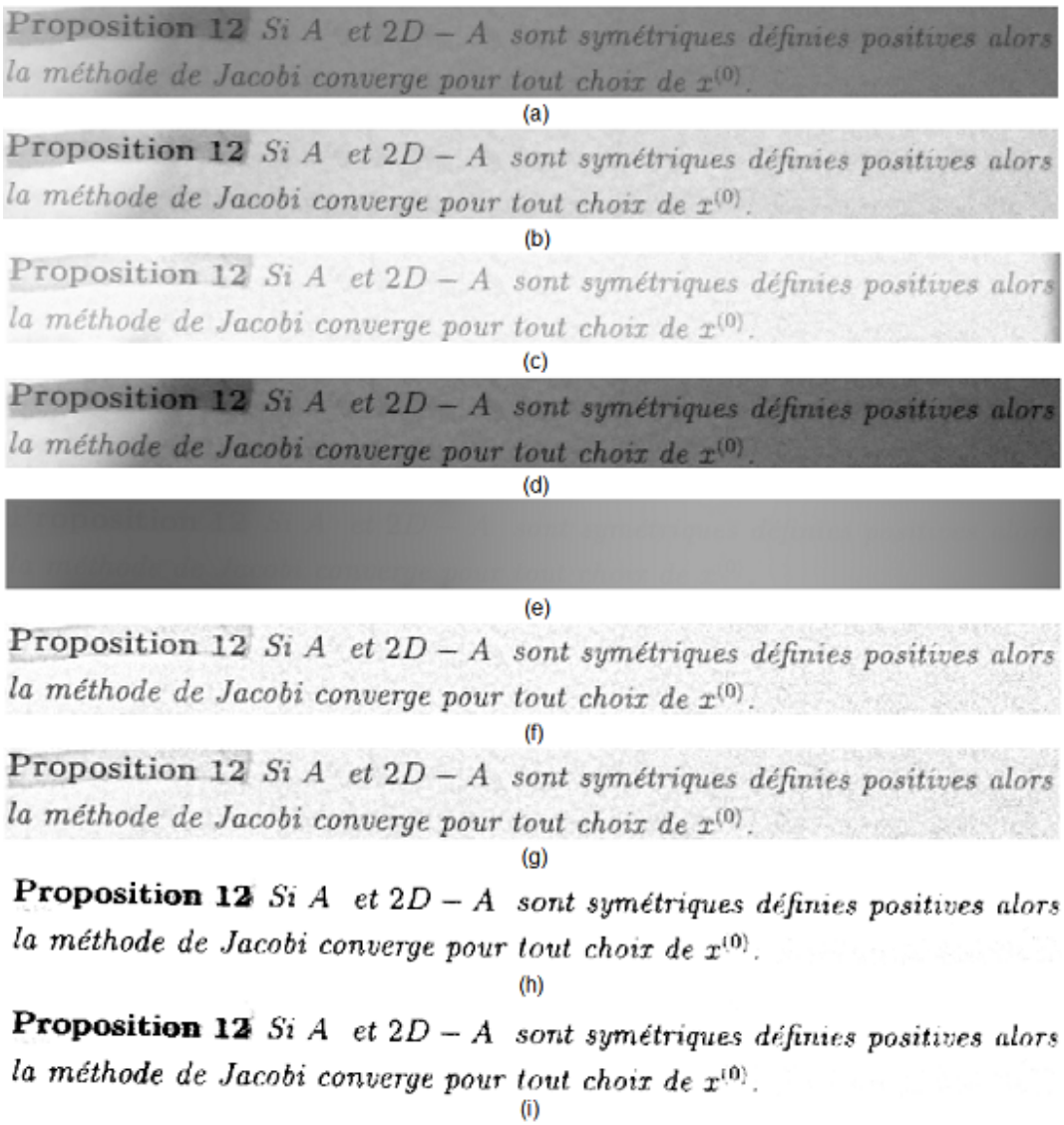


Figure 2: From top to bottom, (a) the original image 'Proposition', the recovered reflectance by (b) MSRetinex [40], (c) Variational Retinex [25], (d) Screened Poisson Equation [34], (e) PDE Retinex [29], (f) local laplacian model [32], (g) the local derivatives model [32], (h) the first proposed nonlocal laplacian model, (i) the second proposed nonlocal derivatives model

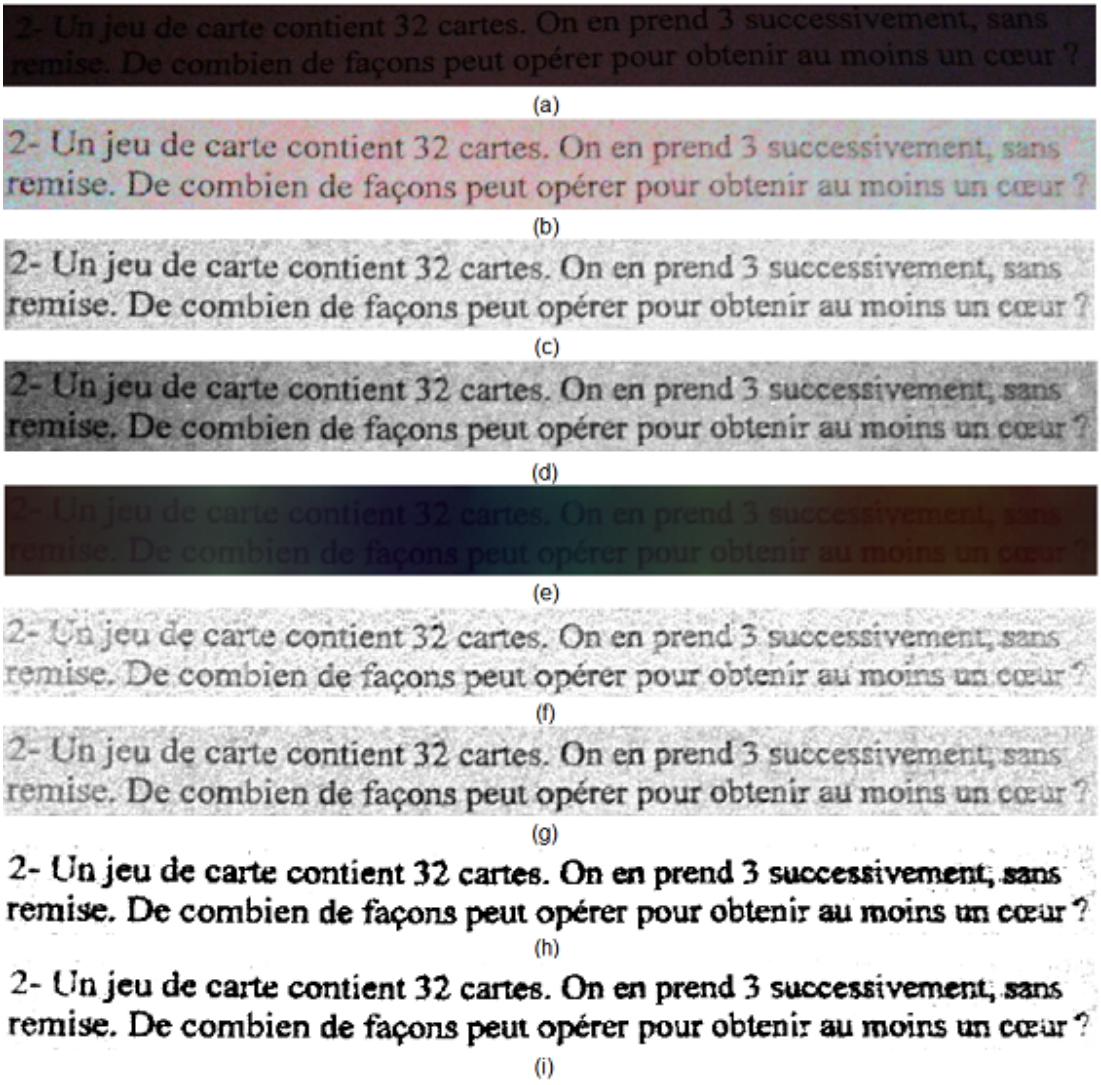


Figure 3: From top to bottom, (a) the original image 'Exercise', the recovered reflectance by (b) MSRetinex [40], (c) Variational Retinex [25], (d) Screened Poisson Equation [34], (e) PDE Retinex [29], (f) local laplacian model [32], (g) the local derivatives model [32], (h) the first proposed nonlocal laplacian model, (i) the second proposed nonlocal derivatives model

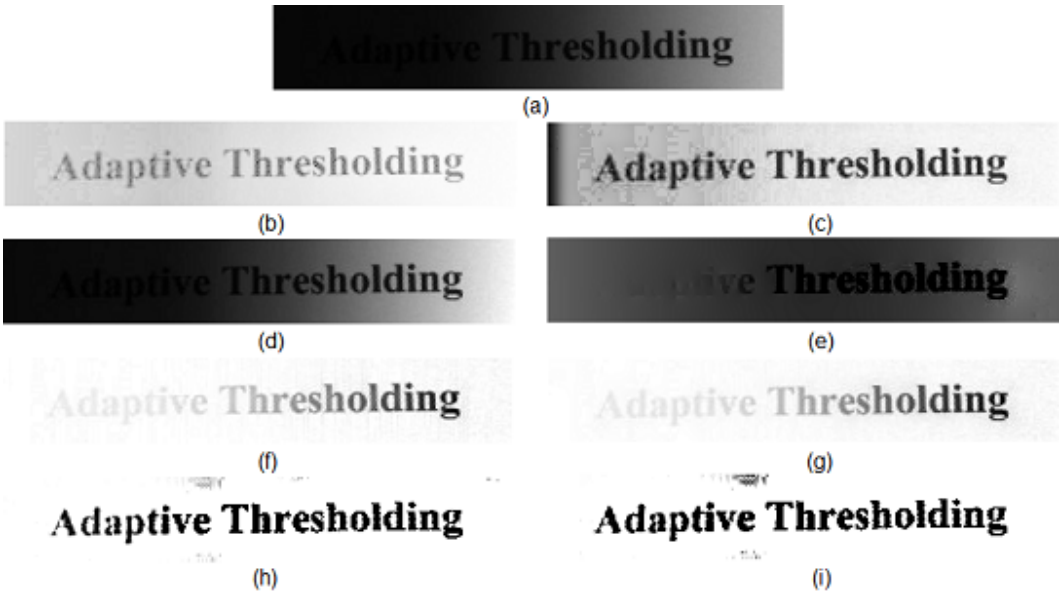


Figure 4: From top to bottom and from left to right, (a) the original image 'Phrase', the recovered reflectance by (b) MSRetinex [40], (c) Variational Retinex [25], (d) Screened Poisson Equation [34], (e) PDE Retinex [29], (f) local laplacian model [32], (g) the local derivatives model [32], (h) the first proposed nonlocal laplacian model, (i) the second proposed nonlocal derivatives model

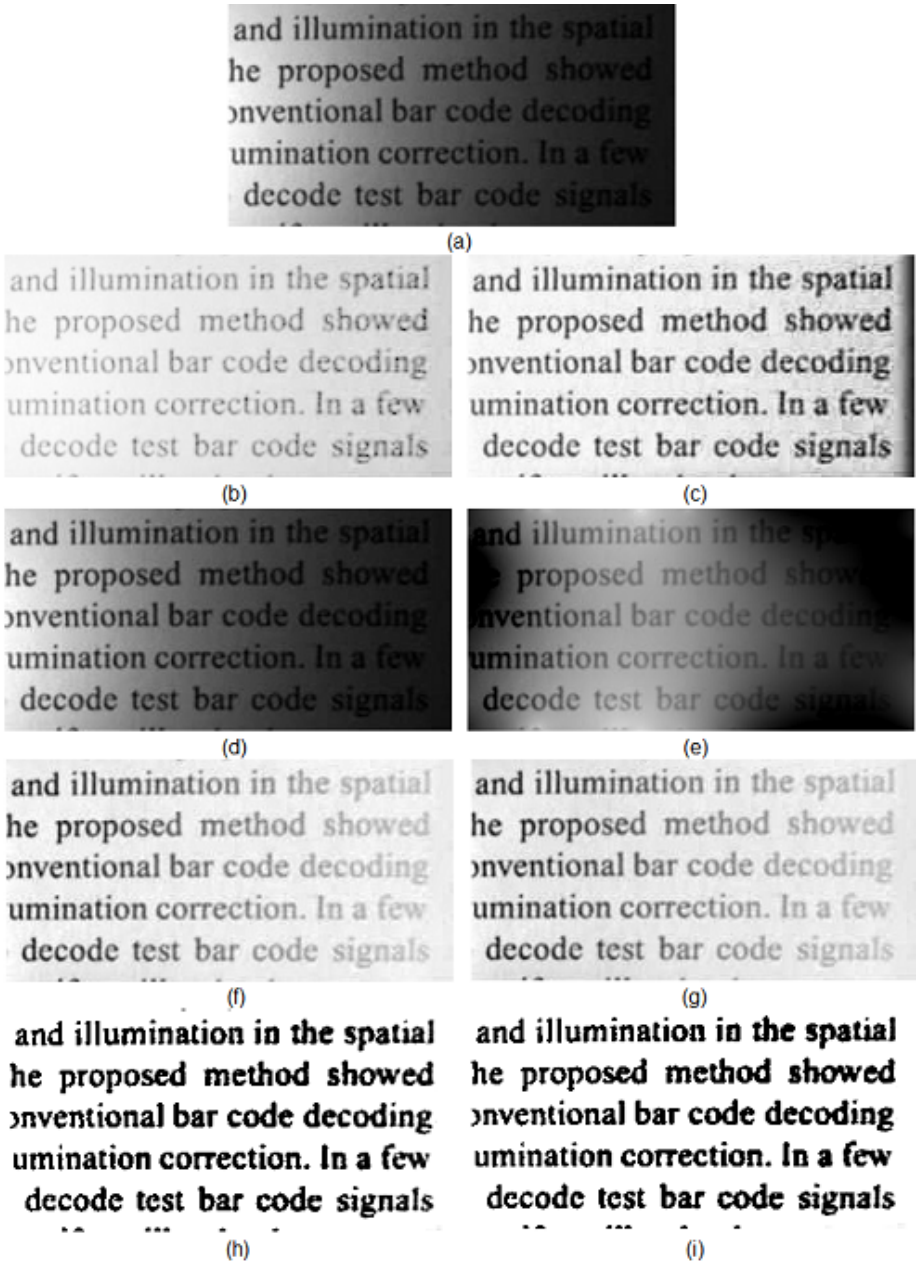


Figure 5: From top to bottom and from left to right, (a) the original image 'Text', the recovered reflectance by (b) MSRetinex [40], (c) Variational Retinex [25], (d) Screened Poisson Equation [34], (e) PDE Retinex [29], (f) local laplacian model [32], (g) the local derivatives model [32], (h) the first proposed nonlocal laplacian model, (i) the second proposed nonlocal derivatives model

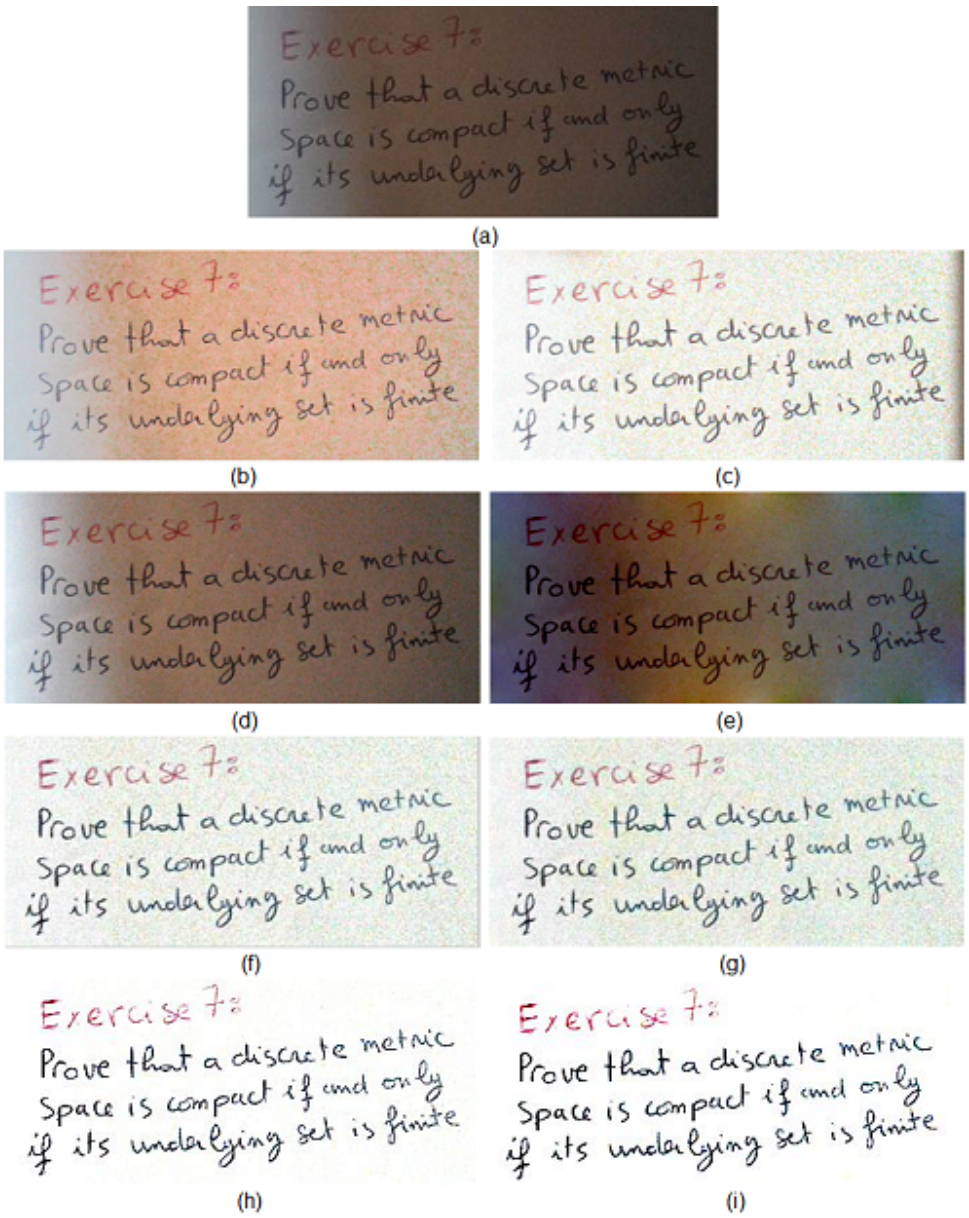


Figure 6: From top to bottom and from left to right, (a) the original image 'Exercise 7', the recovered reflectance by (b) MSRetinex [40], (c) Variational Retinex [25], (d) Screened Poisson Equation [34], (e) PDE Retinex [29], (f) local laplacian model [32], (g) the local derivatives model [32], (h) the first proposed nonlocal laplacian model, (i) the second proposed nonlocal derivatives model

Figure 7: From top to bottom and from left to right, (a) the original image 'Board', the recovered reflectance by (b) MSRetinex [40], (c) Variational Retinex [25], (d) Screened Poisson Equation [34], (e) PDE Retinex [29], (f) local laplacian model [32], (g) the local derivatives model [32], (h) the first proposed nonlocal laplacian model, (i) the second proposed nonlocal derivatives model

Image	Method	FineReader %	OnlineOCR %	Tesseract %
Proposition	the original	72.11	77.88	0
	MSRetinex	82.69	95.19	0
	Var. Retinex	95.19	95.19	**82.69**
	Screned Poisson Equ	91.35	82.96	0
	PDE Retinex	0	0	0
	L.Laplacian Model	96.15	**98.08**	75
	L.Derivatives Model	96.15	**98.08**	**80.77**
	The proposed Model 1	**97.11**	**98.08**	75
	The proposed Model 2	**98.08**	**96.15**	76.92
Exercise	the original	0	0	0
	MSRetinex	76.67	97.5	0
	Var. Retinex	98.33	95.83	10
	Screned Poisson Equ	98.33	**100**	0
	PDE Retinex	0	0	0
	L.Laplacian Model	95.83	93.33	37.5
	L.Derivatives Model	93.33	95	40.83
	The proposed Model 1	**100**	**100**	**82.5**
	The proposed Model 2	**99.17**	**97.5**	**82.5**
Phrase	the original	0	0	0
	MSRetinex	95	**100**	0
	Var. Retinex	**100**	95	60
	Screned Poisson Equ	35	0	0
	PDE Retinex	0	60	0
	L.Laplacian Model	45	60	50
	L.Derivatives Model	50	60	55
	The proposed Model 1	**100**	**100**	**100**
	The proposed Model 2	**100**	**100**	**100**
Text	the original	39.2	62.4	0
	MSRetinex	92	95.2	43.2
	Var. Retinex	**99.19**	**100**	89.6
	Screned Poisson Equ	56.8	64.8	0
	PDE Retinex	18.4	31.2	0
	L.Laplacian Model	68.8	99.19	85.6
	L.Derivatives Model	68.8	99.19	86.4
	The proposed Model 1	**99.19**	**100**	**98.39**
	The proposed Model 2	**100**	**100**	**98.39**

Table 1: **OCR accuracy results**

6. Conclusion

In the present paper, an existence result for the local PDE's text enhancement model proposed in (6) was given, then, based on the nonlocal operators properties, two nonlocal models were proposed to estimate the reflectance image. Moreover, the two nonlocal algorithms are simple, easy to implement and give quite satisfactory results compared to some available methods, which gives confidence in the efficiency of our models. Beyond the two proposed nonlocal PDEs, we could also extend these two models to other PDEs with more regularity properties.

References

[1] Gilles Aubert and Pierre Kornprobst. *Mathematical problems in image processing: partial differential equations and the calculus of variations*, volume 147. Springer Science & Business Media, 2006.

[2] M Bardi, MG Crandall, LC Evans, HM Soner, and PE Souganidis. Viscosity solutions and applications. lectures given at the 2nd cime session held in montecatini terme, june 12–20, 1995. edited by i. capuzzo dolcetta and pl lions. *Lecture Notes in Mathematics*, 1660.

[3] Andrew Blake. On lightness computation in mondrian world. In *Central and peripheral mechanisms of colour vision*, pages 45–59. Springer, 1985.

[4] Gavin Brelstaff and Andrew Blake. Computing lightness. *Pattern Recognition Letters*, 5(2):129–138, 1987.

[5] Antoni Buades, Bartomeu Coll, and J-M Morel. A non-local algorithm for image denoising. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 60–65. IEEE, 2005.

[6] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. Nonlocal image and movie denoising. *International journal of computer vision*, 76(2):123–139, 2008.

[7] Gulcin Caner and Ismail Haritaoglu. Shape-dna: Effective character restoration and enhancement for arabic text documents. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 2053–2056. IEEE, 2010.

[8] Michael G Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27(1):1–67, 1992.

[9] Alexei A Efros and Thomas K Leung. Texture synthesis by non-parametric sampling. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1033–1038. IEEE, 1999.

[10] Chun-Nian Fan and Fu-Yan Zhang. Homomorphic filtering based illumination normalization method for face recognition. *Pattern Recognition Letters*, 32(10):1468–1479, 2011.

[11] Ya-Ru Fan, Ting-Zhu Huang, Tian-Hui Ma, and Xi-Le Zhao. Cartoon–texture image decomposition via non-convex low-rank texture regularization. *Journal of the Franklin Institute*, 354(7):3170–3187, 2017.

[12] R Fries and J Modestino. Image enhancement by stochastic homomorphic filtering. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 27(6):625–637, 1979.

[13] R Fries and J Modestino. Image enhancement by stochastic homomorphic filtering. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 27(6):625–637, 1979.

[14] Marius Ghergu and Vicentiu Radulescu. *Nonlinear PDEs: Mathematical models in biology, chemistry and population genetics*. Springer Science & Business Media, 2011.

[15] Yoshikazu Giga. *Surface evolution equations: A level set approach*, volume 99. Springer Science & Business Media, 2006.

[16] Guy Gilboa and Stanley Osher. Nonlocal operators with applications to image processing. *Multiscale Modeling & Simulation*, 7(3):1005–1028, 2008.

[17] Pelin Gorgel, Ahmet Sertbas, and Osman N Ucan. A wavelet-based mammographic image denoising and enhancement with homomorphic filtering. *Journal of medical systems*, 34(6):993–1002, 2010.

[18] Berthold Horn. *Robot vision*. MIT press, 1986.

[19] Berthold KP Horn. Determining lightness from an image. *Computer graphics and image processing*, 3(4):277–299, 1974.

[20] Yali Huang, Yuehua Gao, Hong Wang, Dongmei Hao, Jinhui Zhao, and Zhen Zhao. Enhancement of ultrasonic image based on the multi-scale retinex theory. In *Recent Advances in Computer Science and Information Engineering*, pages 115–120. Springer, 2012.

[21] Hitoshi Ishii and Moto-Hiko Sato. Nonlinear oblique derivative problems for singular degenerate parabolic equations on a general domain. *Nonlinear Analysis: Theory, Methods & Applications*, 57(7):1077–1098, 2004.

[22] Wei W Cindy Jiang. Thresholding and enhancement of text images for character recognition. In *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, volume 4, pages 2395–2398. IEEE, 1995.

[23] Wei W Cindy Jiang. Thresholding and enhancement of text images for character recognition. In *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, volume 4, pages 2395–2398. IEEE, 1995.

[24] Daniel J Jobson, Zia-ur Rahman, and Glenn A Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image processing*, 6(7):965–976, 1997.

[25] Ron Kimmel, Michael Elad, Doron Shaked, Renato Keshet, and Irwin Sobel. A variational framework for retinex. *International Journal of computer vision*, 52(1):7–23, 2003.

[26] Edwin H Land. The retinex theory of color vision. *Scientific American*, 237(6):108–129, 1977.

[27] Edwin H Land and John J McCann. Lightness and retinex theory. *Josa*, 61(1):1–11, 1971.

[28] Jingwei Liang and Xiaoqun Zhang. Retinex by higher order total variation l^{1} decomposition. *Journal of Mathematical Imaging and Vision*, 52(3):345–355, 2015.

[29] Nicolas Limare, Ana Belén Petro, Catalina Sbert, and Jean-Michel Morel. Retinex poisson equation: a model for color perception. *Image Processing On Line*, 1:39–50, 2011.

[30] Guowen Ma and Jinfeng Yang. Shadow removal using retinex theory. In *Intelligent Visual Surveillance (IVS), 2011 Third Chinese Conference on*, pages 25–28. IEEE, 2011.

[31] Wenye Ma and Stanley Osher. A tv bregman iterative model of retinex theory. *Ucla Cam Report*, pages 10–13, 2010.

[32] Zouhir Mahani, Jalal Zahid, Sahar Saoud, Mohammed El Rhabi, and Abdelilah Hakim. Text enhancement by pde's based methods. *Image and Signal Processing*, pages 65–76, 2012.

[33] John McCann. Lessons learned from mondrians applied to real images and color gamuts. In *Color and imaging conference*, volume 1999, pages 1–8. Society for Imaging Science and Technology, 1999.

[34] Jean-Michel Morel, Ana-Belen Petro, and Catalina Sbert. Screened poisson equation for image contrast enhancement. *Image Processing On Line*, 4:16–29, 2014.

[35] Michael K Ng and Wei Wang. A total variation model for retinex. *SIAM Journal on Imaging Sciences*, 4(1):345–365, 2011.

[36] Uche Nnolim and Peter Lee. Homomorphic filtering of colour images using a spatial filter kernel in the hsi colour space. In *Instrumentation and Measurement Technology Conference Proceedings, 2008. IMTC 2008. IEEE*, pages 1738–1743. IEEE, 2008.

[37] A van Oppenheim, Ronald Schafer, and Thomas Stockham. Nonlinear filtering of multiplied and convolved signals. *IEEE transactions on audio and electroacoustics*, 16(3):437–466, 1968.

[38] A. Pazy. Semigroups of linear operators and applications to partial differential equations. 1983.

[39] Soo-Chang Pei, Mary Tzeng, and Yu-Zhe Hsiao. Enhancement of uneven lighting text image using line-based empirical mode decomposition. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 1249–1252. IEEE, 2011.

[40] Ana Belén Petro, Catalina Sbert, and Jean-Michel Morel. Multiscale retinex. *Image Processing On Line*, pages 71–88, 2014.

[41] Aishwarya Visvanathan, T Chattopadhyay, and Ujjwal Bhattacharya. Enhancement of camera captured text images with specular reflection. In *Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2013 Fourth National Conference on*, pages 1–4. IEEE, 2013.

[42] Guoxin Xue, Pei Xue, and Qiang Liu. A method to improve the retinex image enhancement algorithm based on wavelet theory. In *Computational Intelligence and Design (ISCID), 2010 International Symposium on*, volume 1, pages 182–185. IEEE, 2010.

[43] Leonid Yaroslavsky. *Digital picture processing: an introduction*, volume 9. Springer Science & Business Media, 2012.

Conflict of interest

The authors declare that they have no conflict of interest.

Computer Vision and Image Understanding 000 (2017) 1–14

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution

Amine Laghrib[a,*], Mahmoud Ezzaki[b], Mohammed El Rhabi[b], Abdelilah Hakim[a], Pascal Monasse[c], Said Raghay[a]

[a] University Cadi Ayyad, Faculty of Science and Technology of Marrakech, Department of Mathematics and Informatics, Marrakech, Morocco
[b] Applied Mathematics and Computer Science department, Ecole des Ponts ParisTech (ENPC), Paris, France
[c] Université Paris-Est, LIGM (UMR CNRS 8049), ENPC, F-77455 Marne-la-Vallée, France

ARTICLE INFO

Article history:
Received 12 December 2016
Revised 18 June 2017
Accepted 15 August 2017
Available online xxx

Keywords:
Multiframe super resolution
Bilateral TV filter
Bounded hessian space
Second order regularization
Relaxed function

ABSTRACT

The aim of a Super resolution (SR) technique is to construct a high-resolution image from a sequence of observed low-resolution ones of the same scene. One major roadblock of an SR reconstitution is removing noise and blur without destroying edges. We propose a novel multiframe image SR algorithm based on a convex combination of Bilateral Total Variation and a non-smooth second order variational regularization, using a controlled weighting parameter. We prove the existence of a minimizer of the proposed energy in the space of functions of bounded Hessian. The minimization of the convex functional is performed with a fast primal-dual algorithm. The simulation results and real experiments show the performance of the proposed algorithm in avoiding undesirable artifacts compared to other methods in the literature.

1. Introduction

Image super resolution (SR) reconstruction is a challenging problem and an active research area in image processing. The principle of this technique is to reconstruct a high-resolution (HR) image by fusing degraded low-resolution (LR) ones. The SR is used in many applications, such as medical diagnostics (Hill et al., 2001), video surveillance (Luong, 2009), satellite imaging, etc.

After the first seminal work proposed by Tsai et al. (Tsai and Huang, 1984), several techniques of the muliframe SR (Laghrib et al., 2016; Lee and Kang, 2003; Park et al., 2003; Schultz et al., 1998; Su et al., 2012; Xiao et al., 2016) have been proposed to enhance the quality of the obtained HR image. Since the SR approach is an ill-posed problem, many SR methods are based on a regularization framework (Caselles et al., 2000; El Mourabit et al., 2017; Jin et al., 2016; Laghrib et al., 2015; Lee and Kang, 2003; Luong, 2009; Villena et al., 2013). Farsiu et al. proposed a fast and robust SR (Farsiu et al., 2004); this technique reduces the complexity of the SR method by decomposing it into two steps: first, finding a blurred HR image from the LR measurements, then, deblurring and denoising it. Both steps are critical to the quality of the HR image. We concentrate here on the second one; we impose some prior on the HR image in a Bayesian framework. Some of the widely-used prior functions are Tikhonov-type regularizers (Lee and Kang, 2003; Patanavijit and Jitapunkul, 2006) and Total Variation (TV)-type regularizers (Ng et al., 2007). Another successful regularization is the bilateral total variation (BTV) with the L^1 norm proposed by Farsiu et al. (2004) and an adaptive norm with BTV regularizer (Zeng and Yang, 2013). These approaches are successful in recovering images with sharp edges but fail on images with smooth surface, suffering from the staircasing effect. To overcome this defect, a second and fourth order variational approach (Bergounioux and Piffet, 2010; Lysaker et al., 2003; Min et al., 2014; Prasath and Vorotnikov, 2014; You and Kaveh, 2000) can be used. Despite its success in numerous denoising problems, it suffers from blur and sometimes does not preserve edges correctly. To perform the process of simultaneous deblurring and denoising, a combined first and second order regularization (Lysaker and Tai, 2006; Papafitsoros and Schönlieb, 2014) has been used and proved its robustness in image restoration, but no such attempts have been proposed to increase the performance of SR methods with the BTV regularizer. An alternative to BTV is the total generalized variation (TGV) (Bredies et al., 2010), TGV^2_α, known for its effectiveness in avoiding the staircasing effect near discontinuity regions. However, Papafitsoros and Schönlieb (Papafitsoros and Schönlieb, 2014) demonstrate that the combined first and second order regularization yields sometimes better results than TGV^2_α at a much lower computation cost. The best performing regularization terms are learned from training data in

* Corresponding author.
E-mail address: laghrib.amine@gmail.com (A. Laghrib).

http://dx.doi.org/10.1016/j.cviu.2017.08.007
1077-3142/

the work of Chen et al. (2013); 2014). Nevertheless, the mathematical well-posedness of the problems using these priors is not ascertained, and we did not investigate those. Motivated by the success of the combination models in restoration problem, we propose in this paper a new combination of the BTV regularizer and a second order differential regularization in the choice of the prior function.

The main idea behind this combination is to use a large weight for the BTV regularization to enhance sharp edges and also to correct the misregistration errors, and use the second order regularizer to eliminate the staircasing caused by the BTV regularization without blurring the image. In the following, we will prove existence of the deblurring and denoising minimization problem (second step of SR), which is a combination of the BTV and second order variational regularization. We use as functional space the bounded Hessian space $HB(\Omega)$ (Demengel, 1985) and we use the classical steps based on relaxed functions (Aubert and Kornprobst, 2006; Demengel and Demengel, 2012).

In Section 2, we present the general SR problem and describe the proposed SR algorithm steps. In Section 3, we introduce the variational problem and the space $HB(\Omega)$ and its essential properties. From there, we will prove the existence of a minimum of the relaxed functional associated to the deblurring and denoising problem, using standard techniques from calculus of variations. In Section 4, we present some experimental results using both synthetic and real images, and we compare our model with some existing algorithms in the SR literature.

2. Problem formulation

The observed images of a real scene are usually in low resolution due to inconvenient atmospheric conditions and low sensor quality. In practice, the acquired image is corrupted by noise, blur and sampling (Milanfar, 2010; Park et al., 2003; Tsai and Huang, 1984). We assume that the LR images are taken under the same environmental conditions using the same sensor. The relationship between an ideal HR image $X \in \mathcal{H}(\Omega)$, $\Omega \subset \mathbb{R}^2$ where $\mathcal{H}$ is an appropriate Hilbert space of real functions defined on Ω and the LR images Y_k where $k = 1 \dots n$ and n is the number of the LR images, is given by the relation

$$Y_k = DF_kHX + e_k, \qquad \forall k = 1, 2, \dots, n, \tag{1}$$

where H represents the linear blur operator in $\mathcal{H}$, i.e., $H \in \mathcal{L}(\mathcal{H})$, D represents the sampling operator that reduces the number of pixels vertically and horizontally by a factor of r such as $D \in \mathcal{L}(\mathcal{H})$, F_k is a geometric warp matrix ($F_k \in \mathcal{L}(\mathcal{H})$) representing a non-parametric transformation that differs in all frames, and $e_k = e_k(Y_k)$ denotes the additive noise for each image, assumed to follow a zero mean Gaussian distribution.

In the presence of different operators of degradation (sampling, blur and noise), the problem becomes very unstable. To deal with it, we use the same approach as in Farsiu et al. (2004) that suggests to separate it into three steps:

1. Computing the warp matrix F_k for each image.
2. Fusing the low-resolution images Y_k into a blurred HR version $\widehat{B}$.
3. Finding the estimation of the HR image X from B.

To align the input images, we have to compute the exact warp matrix F_k, which always considered as a crucial factor for super-resolution success. There are many approaches in the literature (for more details see Milanfar, 2010). In our implementation, we use a non parametric registration (Laghrib et al., 2014; Modersitzki, 2003) to allow for arbitrary motion. The main idea of this approach is to take a bicubic upsamplings $\overline{Y}_i$ of the input images Y_i. We select then the reference image that is used for super-resolution as $\hat{Y}_k$ with $1 \le k \le n$. From the image $\hat{Y}_k$ we calculate transformations to all input images. As a result we get n transformation vector fields which are stored in the linear operator F_i.

For the fusion step (2), we use the technique of Farsiu et al. (2004), recalled in the following section.

2.1. The fusion step

The first part of our algorithm is to compute the blurred HR version $B = HX$. We assume that the additive noise is Gaussian distributed and follows the same distribution for all low resolution images. The blurred image $\widehat{B}$ can be found via the principle of the maximum likelihood estimator (ML) (Milanfar, 2010). The ML suggests the choice of $\widehat{B}$ that maximizes the likelihood function, which also minimizes the negative log-likelihood function

$$\begin{aligned}\widehat{B} &= \arg\max_B \{p(Y_k|B)\} \\ &= \arg\min_B \{-\log(p(Y_k|B))\} \\ &= \arg\min_B \sum_{k=1}^{n} \|Y_k - DF_kB\|^2_{L^2}. \end{aligned} \tag{2}$$

The steepest descent algorithm (Nocedal and Wright, 2006) can be used to resolve this minimization problem, or, even better, the primal-dual algorithm (Chambolle and Pock, 2011).

2.2. Deconvolution and denoising step

In this step we try to find the HR image X by deblurring and denoising the image $\widehat{B}$. Unfortunately we are facing an unstable inverse problem due to the presence of noise and blur at the same time. To overcome this difficulty, we impose some prior knowledge about the image in a Bayesian framework. Knowing that the measured vectors Y_k undergo a Gaussian noise $e_k(Y_k)$ and blur, $\widehat{B}$ has also white Gaussian noise. Via the Bayes rule, finding the HR image X leads us to look for the solution of the following minimization problem (3) using the maximum a posteriori (MAP) estimator:

$$\begin{aligned}\widehat{X}_{MAP} &= \arg\max_X \{p(X|\widehat{B})\} \\ &= \arg\max_X \left\{ \frac{p(\widehat{B}|X)p(X)}{p(\widehat{B})} \right\} \\ &= \arg\min_X \left\{ -\log(p(\widehat{B}|X)) - \log(p(X)) \right\}, \end{aligned} \tag{3}$$

where $p(\widehat{B}|X)$ represents the likelihood term (data attachment term) and $p(X)$ denotes the prior knowledge on the HR image. To formulate precisely this problem, we need to assume a prior image distribution.

2.3. The prior Gibbs function (PGF)

A well-known manner to represent the image prior distribution p is given by

$$p(X) \propto c_G \cdot \exp\{-G(X)\}, \tag{4}$$

where c_G is a normalization constant and G is a non negative energy function. Since we know the robustness of the BTV regularizer to remove noise and preserve edges, and the success of the variational second order regularization to remove the blocky effect, we propose to use a combination of these two regularizers in the choice of the Gibbs function, controlled by a weight parameter that depends on the gradient of the image:

$$p(X) = \exp\left\{ -\delta \sum_{i=-p}^{p} \sum_{j=-p}^{p} \alpha^{|i|+|j|} \|X - S_x^i S_y^j X\|_1 - (1-\delta) \|f(\nabla^2 X)\|_1 \right\}$$

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007

where the operators S_x^i and S_y^j shift X by i and j pixels in horizontal and vertical directions respectively:

$$S_x^i X(x,y) = X(x+i,y) \text{ and } S_y^j X(x,y) = X(x,y+j).$$

and ∇^2 is the second order operator defined as $\nabla^2 X = (\nabla_{xx}X, \nabla_{xy}X, \nabla_{yx}X, \nabla_{yy}X)$. This formulation depends on the following parameters:

- α: a scalar weight, applied to give a spatially decaying effect to the summation of the regularization terms, $0 < \alpha < 1$.
- p: the spatial window size, $p \geq 1$.
- f: a function defined from $\mathbb{R}^4$ to $\mathbb{R}^+$ verifying some assumptions that we define below. For more details and for some examples of this function, see Aubert and Kornprobst (2006).
- δ: the weight that controls the regularization combination. We define in the following how it is calculated from the gradient of the image.

The first term is a measure of bilateral variation, containing a spatial decaying influence and the intensity difference in the neighborhood.

3. Resolution of the final SR problem

In this section, we rewrite the problem (3) with the defined prior function and the likelihood terms, which will constitute the final problem. To do so, we use the operators F_k defined for each image and the computed blurred HR image $\widehat{B}$ obtained in (2).

$$\widehat{X} = \arg\min_X F(X), \tag{5}$$

with

$$F(X) = \|HX - \widehat{B}\|_1 + \delta \sum_{i=-p}^{p} \sum_{j=-p}^{p} \alpha^{|i|+|j|} \|X - S_x^i S_y^j X\|_1 + (1-\delta)\|f(\nabla^2 X)\|_1. \tag{6}$$

The norm of the Lebesgue space $L^1(\Omega)$ in the expression $\|HX - \widehat{B}\|_1$, where Ω contains all the pixels on the HR grid X, is used because it is robust against outliers (Farsiu et al., 2004). The choice of L^1 instead of the L^2 norm is essentially due to the better effectiveness of the former in removing impulse noise. Also, it is contrast invariant in a weak sense: if $C \in \mathbb{R}^+$, CX is a solution for $C\widehat{B}$ in (6). Before solving the minimization problem (5), we have to check the existence of a solution. In the following, we construct the functional framework and we ensure the existence of a solution.

3.1. Functional framework

We consider a compact set $\Omega \subset \mathbb{R}^2$ smooth enough. We construct the functional framework where the function F is well defined. In other words, we look for a Banach space of functions X defined from Ω to $\mathbb{R}$, which ensures the existence of the different norms in (6). We suppose that $X \in L^1(\Omega)$; this ensures that $\|HX(x) - \widehat{B}(x)\|_1$ is well defined. Moreover, the operator $(I - S_x S_y)$ is defined such that $I - S_x S_y : L^1(\Omega) \longrightarrow L^1(\Omega)$ which implies the existence of $\|X - S_x S_y X\|_1$. Moreover, we assume that f is convex and satisfies the following condition:

$$\exists k, k' > 0,\ \forall x \in \mathbb{R}^4,\ k|x|_2 - k' \leq f(x) \leq k|x|_2 + k', \tag{7}$$

where $|.|_2$ is the Euclidean norm in $\mathbb{R}^4$. The condition (7) ensures that if $\nabla^2 X \in [L^1(\Omega)]^4$ then $\|f(\nabla^2 X)\|_1$ is well defined. There is in fact various choices of the function f, for more details see (Aubert and Kornprobst, 2006); we will define this function in the proposed algorithm part. From the conditions $X \in L^1(\Omega)$, $(X - S_x S_y X) \in L^1(\Omega)$ and $\nabla^2 X \in [L^1(\Omega)]^4$, we recognize as an adequate functional framework, for $q = 1$, the Sobolev space (Brezis, 2010)

$$W^{2,q} = \{X \in L^q(\Omega) \mid D^\alpha X \in L^q(\Omega), [\alpha] \leq 2\}, \tag{8}$$

where

$$D^\alpha = \frac{\partial^{\alpha_1}\partial^{\alpha_2}}{\partial x^{\alpha_1}\partial y^{\alpha_2}} \text{ and } [\alpha] = \alpha_1 + \alpha_2.$$

Unfortunately, this space is not reflexive (Brezis, 2010) and we cannot apply directly the following theorem of optimization:

Theorem 1. *We consider a reflexive space V, and a function $F : V \mapsto \mathbb{R}$ such that F is coercive, lower semi-continuous (l.s.c) and convex. Then the minimum of F exists. Moreover if F is strictly convex, then this minimum is unique.*

We cannot say anything about a bounded minimizing sequence in $W^{2,1}(\Omega)$. To overcome the ill-posedness of this problem, we use the procedure of relaxation (Demengel and Temam, 1984). A typical choice of the space that guarantees the convenient compactness results is the space of Bounded Hessian $HB(\Omega)$ noted also $BV^2(\Omega)$, introduced for the first time by Demengel (1985). We present next the space $HB(\Omega)$ and its useful properties.

3.2. The space HB(Ω)

We define the space $HB(\Omega)$ of bounded Hessian functions. First, we define the space of bounded variation functions $BV(\Omega)$ as

$$BV(\Omega) = \left\{X \in L^1(\Omega) \text{ such that } \nabla X \in \mathcal{M}_b(\Omega, \mathbb{R}^2)\right\}, \tag{9}$$

where $\mathcal{M}_b(\Omega, \mathbb{R}^2)$ is the space of bounded measures from Ω to $\mathbb{R}^2$. We define the space $HB(\Omega)$ using the space $BV(\Omega)$ by

$$HB(\Omega) = \left\{X \in W^{1,1}(\Omega), \nabla X \in (BV(\Omega))^2\right\} \tag{10}$$

It is clear that $W^{2,1}(\Omega) \subset HB(\Omega)$. We define now a norm of the space $HB(\Omega)$.

Theorem 2. *The space $HB(\Omega)$ is a Banach space, when endowed with the norm $\|.\|_{HB}$ defined as*

$$\|X\|_{HB} = \|X\|_{W^{1,1}} + |DX|,$$

where

$$\|X\|_{W^{1,1}} = \|X\|_1 + \|\nabla X\|_1,$$

and

$$|DX| = \sup\{\int_\Omega X \operatorname{div}^2(\phi) : \phi \in \mathcal{C}_0^2(\Omega),\ \|\phi\|_{L^\infty(\Omega)} \leq 1\};$$

div^2 *is the second order divergence operator. In the discrete setting, this operator is presented with the adjointness property*

$$\operatorname{div}^2 X \cdot Y = X \cdot \nabla^2 Y \quad \forall Y \in \mathbb{R}^M,\ X \in \left(\mathbb{R}^M\right)^4,$$

and the dot "$\cdot$" denotes the Euclidean inner product. $\mathcal{C}_0^2(\Omega)$ represents the space of twice continuously differentiable functions with compact support in Ω.

We have also the embedding results proved by Demengel (1985) as well:

Theorem 3. *Assuming that the dimension $n > 1$, we have*

$$HB(\Omega) \hookrightarrow W^{1,q}(\Omega), \text{ when } q \leq \frac{n}{n-1},$$

with continuous embedding. Moreover the embedding is compact if $q < \frac{n}{n-1}$. In particular, we have

$$HB(\Omega) \hookrightarrow L^q(\Omega), \text{ for } q \leq \frac{n}{n-2}, \text{ if } n > 2,$$

$$HB(\Omega) \hookrightarrow L^q(\Omega), \text{ for } q \geq 1, \text{ if } n = 2.$$

Since in our problem we treat the case $n = 2$ and Ω is a subset of $\mathbb{R}^2$, we have in fact $HB(\Omega) \hookrightarrow W^{1,2}(\Omega)$ with continuous embedding and $HB(\Omega) \hookrightarrow W^{1,1}(\Omega)$ with compact embedding.

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007

Definition 1. (Weak * convergence in $HB(\Omega)$)

Let $(X_k)_{k\in\mathbb{N}}$ and $X\in HB(\Omega)$. We say that (X_k) converges to X weakly* if

$$X_k \overset{L^1(\Omega)}{\longrightarrow} X \quad \text{and} \quad \nabla X_k \underset{BV-\omega*}{\rightharpoonup} \nabla X,$$

where $BV-\omega*$ is the weak* topology in $BV(\Omega)$ defined as:

$$X_k \underset{BV-\omega*}{\rightharpoonup} X \Longleftrightarrow \begin{cases} X_k \underset{L^1}{\longrightarrow} X \\ DX_k \overset{*}{\underset{M}{\rightharpoonup}} DX, \end{cases}$$

and $DX_k \overset{*}{\underset{M}{\rightharpoonup}} DX$ means

$$\int_\Omega \varphi\, DX_k \longrightarrow \int_\Omega \varphi\, DX \qquad \forall \varphi \in \mathcal{C}_0^1(\Omega).$$

We also need this compactness result in $BV(\Omega)$:

Theorem 4. *(Compactness in $BV(\Omega)$)*

Every uniformly bounded sequence $(X_k)_{k\in\mathbb{N}}$ in $BV(\Omega)$ is relatively compact in $L^q(\Omega)$, for $1\le q<\frac{n}{n-1}$, $n\ge 1$. Moreover, there exist a subsequence $(X_{kj})_{j\in\mathbb{N}}$ and $X\in BV(\Omega)$ such as $X_{kj} \underset{BV-\omega*}{\rightharpoonup} X$.

Another very useful result to demonstrate the l.s.c is the Buttazo and Freddi theorem:

Theorem 5. *Let Ω be an open subset of $\mathbb{R}^n$, ν, $(\nu_k)_{k\in\mathbb{N}}$ be $\mathbb{R}^m$-valued finite Radon measures and μ, $(\mu_k)_{k\in\mathbb{N}}$ be positive Radon measure in Ω. Let $g:\mathbb{R}^m\to\mathbb{R}$ be a convex function and suppose that $\mu_k\to\mu$ and $\nu_k\to\nu$ weakly* in Ω. Consider the Lebesgue decompositions:*

$$\nu=\left(\frac{\nu}{\mu}\right)\mu+\nu_s, \quad \nu_k=\left(\frac{\nu_k}{\mu_k}\right)\mu_k+\nu_{s_k}, \quad k\in\mathbb{N}.$$

If $\mu=\mu_k=dx^n$, then

$$g(\nu)(\Omega)\le\liminf_{k\to\infty} g(\nu_k)(\Omega).$$

3.3. The variational problem

In this subsection, we prove the existence of a solution to the problem (5) by using the direct method in calculus of variations.

Theorem 6. *Under the hypothesis (7), and assuming that the operators $(I-S_xS_y)$ and H defined: $L^1(\Omega)\longrightarrow L^1(\Omega)$ are continuous; assuming also that H does not annihilate the constants (in particular $H.1\neq 0$); then, the optimization problem*

$$\min_X \overline{F}_{ext}(X), \tag{11}$$

has a solution in $HB(\Omega)$, where $\overline{F}_{ext}$ is the lower semicontinuous envelope of F defined in the proof.

Proof. First of all, we extend the function F in $HB(\Omega)$ by :

$$F_{ext}(X)=\begin{cases} F(X) & \text{if } X\in W^{2,1}(\Omega) \\ +\infty & \text{if } X\in HB(\Omega)\setminus W^{2,1}(\Omega) \end{cases} \tag{12}$$

□

We can check easily that F_{ext} is not lower semi continuous for the weak* topology defined in $HB(\Omega)$ following the same steps as in Papafitsoros and Schönlieb (2014), so we should determine its l.s.c envelope defined in the whole space $HB(\Omega)$ by

$$\begin{aligned}\overline{F}_{ext}(X) &= \|HX-\widehat{B}\|_1 \\ &\quad +\delta\sum_{i=-p}^{p}\sum_{j=-p}^{p}\alpha^{|i|+|j|}\|X-S_x^iS_y^jX\|_1 \\ &\quad +(1-\delta)\|f(D^2X)\|_1,\end{aligned} \tag{13}$$

where $D^2X=D\nabla X$ and D is the distributional gradient of X. With these notations $\overline{F}_{ext}$ coincides with F on $W^{2,1}(\Omega)$. Let us prove now the coerciveness of $\overline{F}_{ext}$.

- Let $(X_n)_{n\in\mathbb{N}}$ be a minimizing sequence of the function F_{ext} such that

$$\lim_{n\to+\infty}\overline{F}_{ext}(X_n)=\inf_{X\in HB(\Omega)}\overline{F}_{ext}(X).$$

Using the hypothesis (7) on F_{ext}, we can deduce that

$$\begin{cases} |D^2X_n|(\Omega)\le M_1, \\ \|HX_n-\widehat{B}\|_1\le M_2, \\ \sum_{i=-p}^{p}\sum_{j=-p}^{p}\alpha^{|i|+|j|}\|X_n-S_x^iS_y^jX_n\|_1\le M_3. \end{cases} \tag{14}$$

We prove now that $\|X_n\|_1$ is also bounded. We use a classical approach (Vese, 1996). We construct two sequences $Y_n=\frac{1}{|\Omega|}\int_\Omega X_n\,dx$ and $Z_n=X_n-Y_n$; then

$$\int_\Omega Z_n\,dx=0, \quad \text{and} \quad \nabla Z_n=\nabla X_n. \tag{15}$$

Using the generalized Poincaré–Wirtinger inequality (Brezis, 2010) for the norm $\|.\|_2$ of the Lebesgue space $L^2(\Omega)$, there is a universal constant C that we will use in the following, such that

$$\|Z_n\|_2\le C\|\nabla Z_n\|_2. \tag{16}$$

Based on the approximation given to the TV prior suggested in Li and Santosa (1996), we can deduce that

$$\|\nabla X_n\|_1\simeq\|Q_xX_n\|_1+\|Q_yX_n\|_1,$$

This equivalence is obtained computationally, where the operators Q_x and Q_y can be defined in the regularization BTV such that $i=1$, $j=0$ and $i=0$, $j=1$ respectively, which is equivalent to the choice $Q_x=(I-S_x^1)$ and $Q_y=(I-S_y^1)$. Since we choose $p\ge 1$ in the BTV regulizer, we can deduce that

$$\|\nabla X_n\|_1\le\sum_{i=-p}^{p}\sum_{j=-p}^{p}\alpha^{|i|+|j|}\|X_n-S_x^iS_y^jX_n\|_1. \tag{17}$$

Using the third inequality in (14) and relations (15) and (16), we have

$$\|Z_n\|_2\le C, \tag{18}$$

then

$$\begin{aligned}\|X_n\|_2 &= \|X_n-Y_n+Y_n\|_2 \\ &= \|Z_n+Y_n\|_2 \\ &\le \|Z_n\|_2+\|Y_n\|_2 \\ &\le C+\|Y_n\|_2\end{aligned} \tag{19}$$

with $\|Y_n\|_{L^2(\Omega)}=\|\int_\Omega X_n\,dx\|$. We also have

$$\begin{aligned}\left\|H\left(\frac{1}{|\Omega|}\int_\Omega X_n\,dx\right)\right\|_1 &\le \|HY_n-HX_n\|_1 \\ &\quad +\|HX_n-\widehat{B}\|_1+\|\widehat{B}\|_1 \\ &\le \|H\|_{L^\infty(\Omega)}\|Z_n\|_1+C+\|\widehat{B}\|_1 \\ &\le \|H\|_{L^\infty(\Omega)}\|Z_n\|_2+C \\ &\le C,\end{aligned} \tag{20}$$

and finally

$$\begin{aligned}\left\|H\left(\frac{1}{|\Omega|}\int_\Omega X_n\,dx\right)\right\|_1 &= \left|\frac{1}{|\Omega|}\int_\Omega X_n\,dx\right|\|H.1\|_1 \\ &\le C.\end{aligned} \tag{21}$$

We can deduce that the sequence $(X_n)_{n\in\mathbb{N}}$ is bounded in $L^2(\Omega)$, and Ω is bounded so it is also bounded in $L^1(\Omega)$. In addition we have that, $\forall X_n\in HB(\Omega)$:

$$\int_\Omega|\nabla X_n|\,dx\le C_1|D^2X_n|(\Omega)+C_2\|X_n\|_1. \tag{22}$$

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007

Since $(X_n)_{n\in\mathbb{N}}$ is bounded in $L^1(\Omega)$, $(\nabla X_n)_{n\in\mathbb{N}}$ is bounded in $L^1(\Omega)$ and $(D^2X_n)_{n\in\mathbb{N}}$ is bounded in $L^1(\Omega)$, we deduce finally that (X_n) is bounded in $HB(\Omega)$. Using the embedding results in Theorem 3, the sequence $(X_n)_{n\in\mathbb{N}}$ is also bounded in $W^{1,1}(\Omega)$ and $HB(\Omega)\hookrightarrow W^{1,1}(\Omega)$ compactly. Thus, we can extract a subsequence $(X_{n_k})_{k\in\mathbb{N}}$ such that

$$X_{n_k} \xrightarrow{W^{1,1}(\Omega)} X.$$

In addition $(\nabla X_n)_{n\in\mathbb{N}}$ is also bounded in $[BV(\Omega)]^2$, so that, based on the result of compactness in Theorem 4, we can extract a subsequence $(\nabla X_{n_k})_{k\in\mathbb{N}}$ satisfying

$$\nabla X_{n_k} \overset{W^*}{\rightharpoonup} Y.$$

Using the weak* convergence in $BV(\Omega)$ we have

$$\begin{cases} \nabla X_{n_k} \underset{L^1}{\longrightarrow} Y, \\ D\nabla X_{n_k} \underset{M}{\overset{*}{\rightharpoonup}} DY, \end{cases}$$

where $D\nabla X_{n_k} \underset{M}{\overset{*}{\rightharpoonup}} DY$ is equivalent to

$$\int_\Omega \varphi\, D\nabla X_{n_k} \longrightarrow \int_\Omega \varphi\, DY \qquad \forall \varphi \in \mathcal{C}_0^1(\Omega).$$

Therefore $\nabla X = Y$, and we finally get

$$\begin{cases} X_{n_k} \longrightarrow X & \text{in } W^{1,1}(\Omega) \\ \nabla X_{n_k} \overset{W^*}{\rightharpoonup} \nabla X & \text{in } [BV(\Omega)]^2 \end{cases}$$

$$\Longleftrightarrow X_{n_k} \overset{*}{\rightharpoonup} X \text{ in } HB(\Omega).$$

To finish this proof it remains to show that F_{ext} is weak sequentially l.s.c in $HB(\Omega)$.

- To prove that $\overline{F}_{ext}$ is weak sequentially l.s.c, it suffices to establish the following inequality

$$\liminf_{X_{n_k}\rightharpoonup X} \overline{F}_{ext}(X_n) \geq \overline{F}_{ext}(X).$$

Let X and $(X_n)_{n\in\mathbb{N}}$ be functions taken from $HB(\Omega)$ such that

$$X_n \underset{HB(\Omega)}{\overset{*}{\rightharpoonup}} X.$$

This implies

$$\|X_n - X\|_1 \longrightarrow 0.$$

Since H and $(I - S_x^i S_y^j)$ are continuous operators from L^1 to itself we can check that

$$\|HX_n - B\|_1 \longrightarrow \|HX - B\|_1,$$

and

$$\sum_{i=-p}^{p}\sum_{j=-p}^{p} \alpha^{|i|+|j|} \|X_n - S_x^i S_y^j X_n\|_1 \longrightarrow \sum_{i=-p}^{p}\sum_{j=-p}^{p} \alpha^{|i|+|j|} \|X - S_x^i S_y^j X\|_1. \tag{23}$$

Finally we use Theorem 5 of Buttazo and Freddi

$$\nu_n = dx^2,\ \mu = D^2X \quad \text{and} \quad \mu_n = D^2X_n$$

to deduce that

$$\liminf_{X_n\rightharpoonup X} f(D^2X_n)(\Omega) \geq f(D^2X)(\Omega),$$

which gives

$$\liminf_{X_n\rightharpoonup X} \overline{F}_{ext}(X_n) \geq \overline{F}_{ext}(X),$$

i.e., X is the minimum of $\overline{F}_{ext}$ in $HB(\Omega)$.

We cannot say anything about the uniqueness of the solution since $\|HX_n - B\|_1$ is not strictly convex, but if we replace this term by $\|HX_n - B\|_2$, the solution is unique since the blur operator H is injective. In the next section, we define the proposed iterative SR algorithm.

3.4. The proposed SR algorithm

Let us describe the minimization algorithm used in our SR approach. We can simply use the alternative approach proposed in Rudin et al. (1992), based on the discretization of the gradient descent PDE associated to the minimizing function defined in (6). Unfortunately this approach is very slow and may sometimes generate some errors, hence a faster and more consistent algorithm is needed. We use an efficient primal-dual algorithm (Chambolle and Pock, 2011), well adapted to our non-smooth convex optimization problem. Let us detail the proposed algorithm.

First, we define the function f in (6) using the second order differential operators noted ∇^2 as

$$f(\nabla^2 X) = \nabla_{xx}X + \nabla_{xy}X + \nabla_{xy}X + \nabla_{yy}X,$$

where

$$\nabla^2 X = (\nabla_{xx}X, \nabla_{xy}X, \nabla_{xy}X, \nabla_{yy}X).$$

We start with the following notation:

$$K_1 = \delta \begin{pmatrix} \alpha^{2p}(I - S_x^{-p}S_y^{-p}) \\ \alpha^{2p-1}(I - S_x^{-p}S_y^{-p+1}) \\ \vdots \\ \alpha^{2p}(I - S_x^{p}S_y^{p}) \end{pmatrix} \Bigg\} (2p+1)^2, \tag{24}$$

$$K_2 = (1-\delta) \sum_{a,b\in\{x,y\}} \nabla_{ab}, \tag{25}$$

and

$$K = \begin{pmatrix} K_1 \\ K_2 \end{pmatrix} \Bigg\} (2p+1)^2 + 1, \tag{26}$$

and we define also the function F_1 as follows

$$\begin{aligned} \mathrm{F}_1(KX) &= \|K_1X\|_1 + \|K_2X\|_1 \\ &= \delta \sum_{i=-p}^{p}\sum_{j=-p}^{p} \alpha^{|i|+|j|} \|(I - S_x^i S_y^j)X\|_1 \\ &\quad + (1-\delta)\, \| \sum_{a,b\in\{x,y\}} \nabla_{ab}X \|_1. \end{aligned} \tag{27}$$

Also, the function F_2 is defined as

$$\mathrm{F}_2(\tilde{K}X) = \|HX - \widehat{B}\|_1, \tag{28}$$

where $\mathrm{F}_2(X) = \|X\|_1$. Using the notations above, the minimization problem (5) becomes

$$\widehat{X} = \arg\min_X \{\mathrm{F}_1(KX) + \mathrm{F}_2(\tilde{K}X)\}. \tag{29}$$

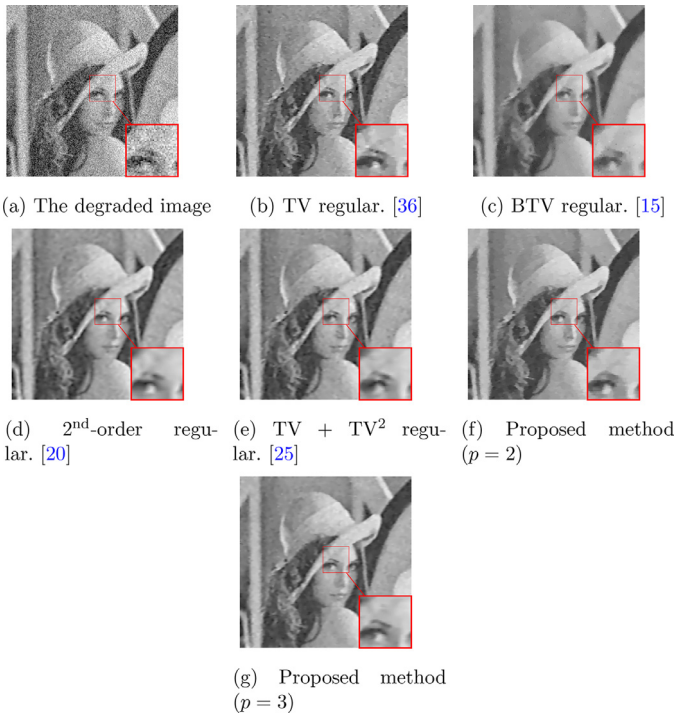
Now, we can apply the primal-dual algorithm (Chambolle and Pock, 2011) to minimize the general problem (29), where F_1 and F_2 are convex functions. $\tilde{K}$ is an affine operator defined from $L^1(\Omega) \to L^1(\Omega)$, and K is a linear operator defined from $L^1(\Omega) \to [L^1(\Omega)]^{(2p+1)^2+1}$. Thus, using the saddle point problem (Chambolle and Pock, 2011), we get the equivalent primal-dual problem

$$\inf_X \sup_{Y,Z} \{< KX, Y > - \mathrm{F}_1^*(Y) + < HX - \widehat{B}, Z > - \mathrm{F}_2^*(Z)\}, \tag{30}$$

where $Y = \begin{pmatrix} Y^1 \\ Y^2 \\ \vdots \\ Y^{(2p+1)^2+1} \end{pmatrix} \in [L^\infty(\Omega)]^{(2p+1)^2+1}$ and $Z \in L^\infty(\Omega)$ denoting the dual variables, F_1^* and F_2^* are the dual functions of respectively F_1 and F_2 defined in a similar way as

$$\mathrm{F}_i^*(Y) = \iota_P(Y) = \begin{cases} 0 & Y \in P \\ +\infty & Y \notin P \end{cases} \tag{31}$$

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007

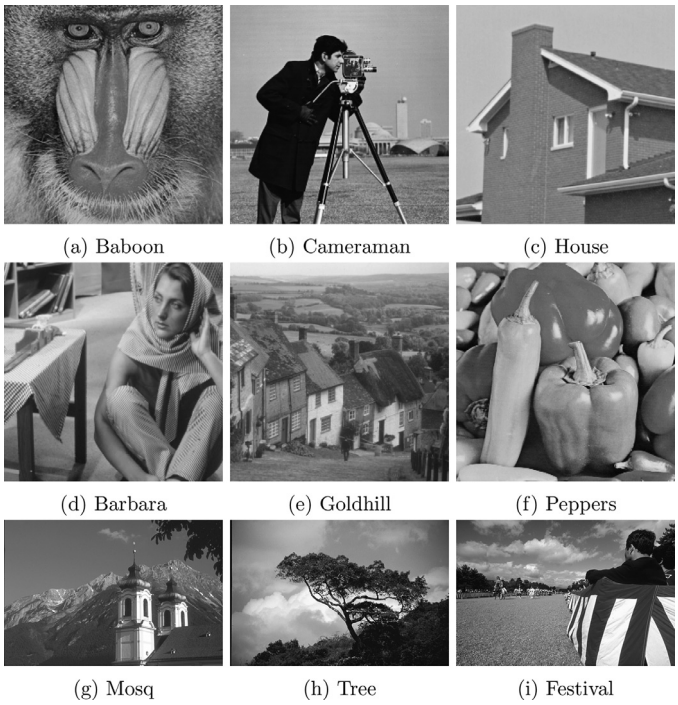


(a) The degraded image (b) TV regular. [36] (c) BTV regular. [15]

(d) 2nd-order regular. [20] (e) TV + TV^2 regular. [25] (f) Proposed method ($p = 2$)

(g) Proposed method ($p = 3$)

Fig. 1. Simulation results of deconvolution and denoising of *Lena* using different regularization methods.

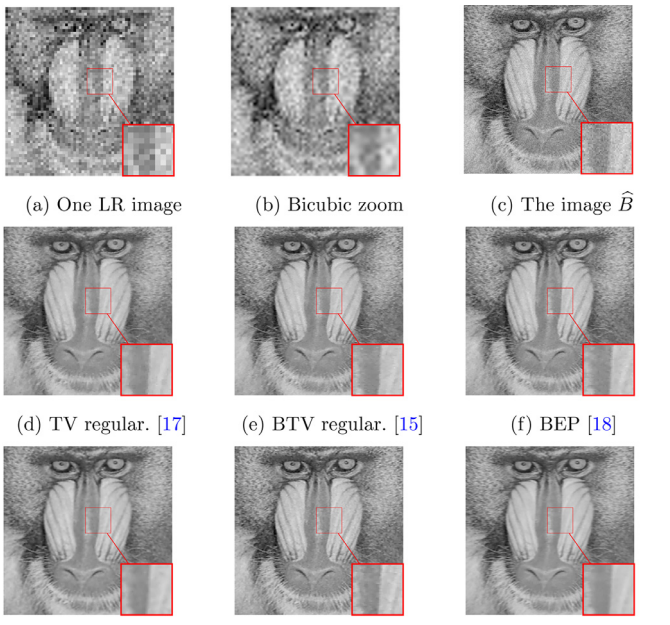


(a) Baboon (b) Cameraman (c) House

(d) Barbara (e) Goldhill (f) Peppers

(g) Mosq (h) Tree (i) Festival

Fig. 2. Set of benchmark images used in tests.



(a) One LR image (b) Bicubic zoom (c) The image $\widehat{B}$

(d) TV regular. [17] (e) BTV regular. [15] (f) BEP [18]

(g) 2nd-order regular. [20] (h) TV + TV^2 regular. [25] (i) Proposed method (with $p = 2$)

Fig. 3. Super resolution of *Baboon* by different methods from 120 LR images.

where $P = \{Y : \|Y\|_\infty \leq 1\}$, and $\|\cdot\|_\infty$ denotes the discrete maximum norm defined as

$$\|Y\|_\infty = \max_i |Y_i|.$$

For the primal-dual algorithm, we have to define the proximal operator functions $(I + \sigma\partial \mathrm{F}_1^*)^{-1}$ and $(I + \sigma\partial \mathrm{F}_2^*)^{-1}$ using the projection on a convex set P (noted Π_P) as follows

$$Y = (I + \sigma\partial \mathrm{F}_1^*)^{-1}(\widehat{Y}) = \Pi_P(\widehat{Y}), \tag{32}$$

and,

$$Z = (I + \sigma\partial \mathrm{F}_2^*)^{-1}(\widehat{Z}) = \Pi_P(\widehat{Z}), \tag{33}$$

where

$$\Pi_P(\widehat{W}_i) = \frac{\widehat{W}_i}{\max(\|\widehat{W}_i\|_\infty, 1)}.$$

To complete the primal-dual algorithm, we have to compute the resolvent operator $(I + \tau\partial \mathrm{G})^{-1}$ for the function $G(X) = 0$ (in our case) as follows

$$X = (I + \tau\partial G)^{-1}(\widehat{X}) = \widehat{X}. \tag{34}$$

Now, we can apply the iterations of the primal-dual algorithm (Unger et al., 2010) to our saddle-point problem (30), we obtain the following iterative algorithm:

$$\begin{cases} \widehat{Y}^{n+1} = (I + \sigma\partial \mathrm{F}_1^*)^{-1}(\widehat{Y}^n + \sigma K X^n), \\ \widehat{Z}^{n+1} = (I + \sigma\partial \mathrm{F}_2^*)^{-1}(\widehat{Z}^n + \sigma \tilde{K} X^n), \\ \widehat{X}^{n+1} = \widehat{X}^n - \tau(K^*\widehat{Y}^{n+1} + \tilde{K}^*\widehat{Z}^{n+1}), \\ \widehat{X}^{n+1} = 2\widehat{X}^{n+1} - \widehat{X}^n. \end{cases} \tag{35}$$

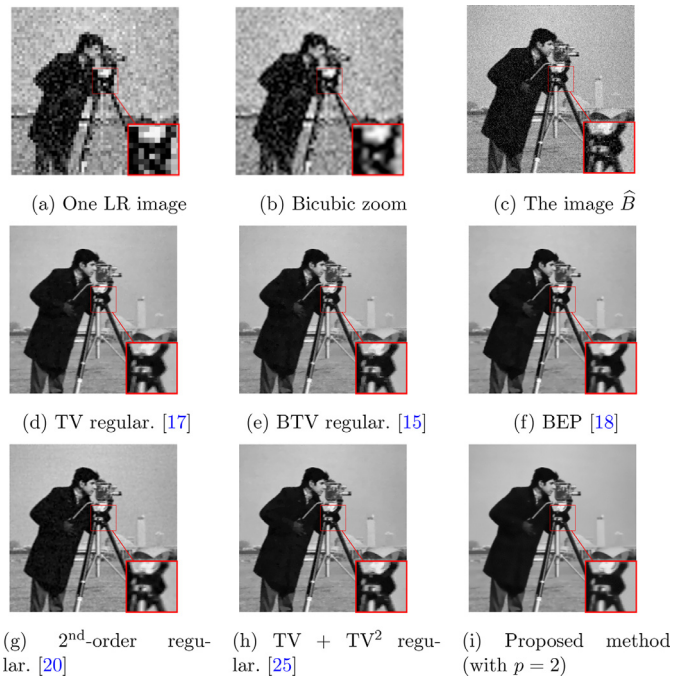
Using the expressions of F_1^* and F_2^*, the solution of the iterative problem (35) is given as

$$\begin{cases} \widehat{Y}^{n+1} = \Pi_P(\widehat{Y}^n + \sigma K X^n), \\ \widehat{Z}^{n+1} = \Pi_P(\widehat{Z}^n + \sigma \tilde{K} X^n), \\ \widehat{X}^{n+1} = \widehat{X}^n - \tau(K^*\widehat{Y}^{n+1} + \tilde{K}^*\widehat{Z}^{n+1}), \\ \widehat{X}^{n+1} = 2\widehat{X}^{n+1} - \widehat{X}^n. \end{cases} \tag{36}$$

where Π_P is the projection on the convex set P defined above. We need also to define the operators K^* and $\tilde{K}^*$, which are respectively the adjoint of the operators K and $\tilde{K}$.

$$K^*Y = \sum_{k=1}^{(2p+1)^2} \left(\delta \sum_{i=-p}^{p} \sum_{j=-p}^{p} \alpha^{|i|+|j|} (I - S_y^{-j} S_x^{-i}) Y_k \right) + (1-\delta) \sum_{a,b\in\{x,y\}} \overline{\nabla_{ab}} Y_{(2p+1)^2+1}, \tag{37}$$

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007

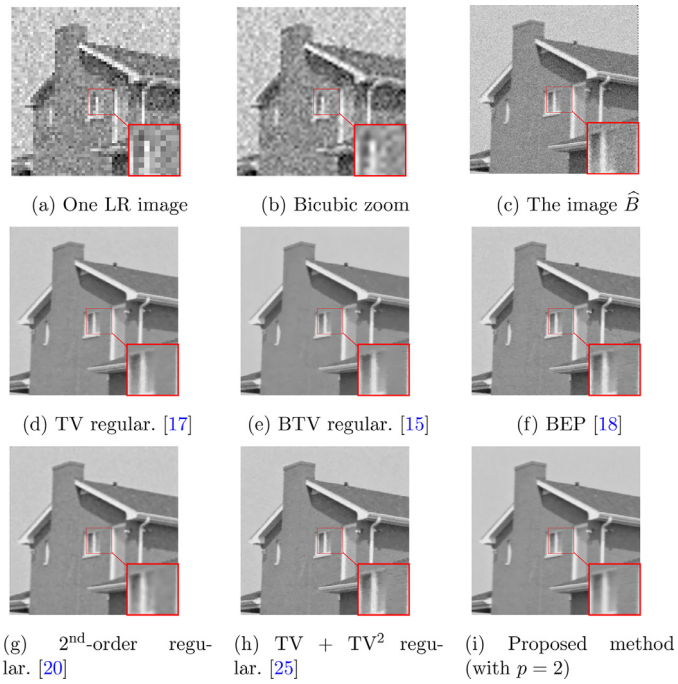


(a) One LR image (b) Bicubic zoom (c) The image $\widehat{B}$

(d) TV regular. [17] (e) BTV regular. [15] (f) BEP [18]

(g) 2^{nd}-order regular. [20] (h) TV + TV^2 regular. [25] (i) Proposed method (with $p = 2$)

Fig. 4. Super resolution of *Cameraman* by different methods from 120 LR images.



(a) One LR image (b) Bicubic zoom (c) The image $\widehat{B}$

(d) TV regular. [17] (e) BTV regular. [15] (f) BEP [18]

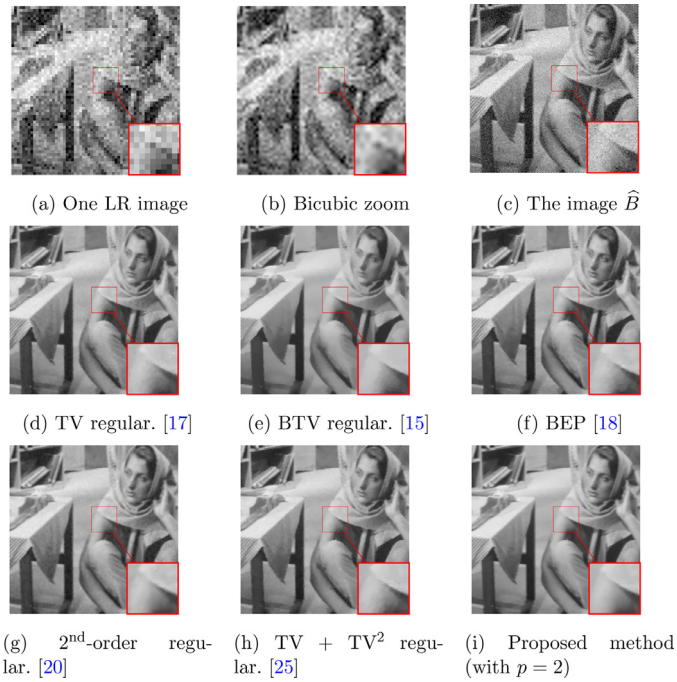
(g) 2^{nd}-order regular. [20] (h) TV + TV^2 regular. [25] (i) Proposed method (with $p = 2$)

Fig. 5. Super resolution of *House* by different methods from 120 LR images.

where the operators S_x^{-i} and S_y^{-j} define the transpose of matrices S_x^i and S_y^j respectively and have a shifting effect in the opposite directions. $\overline{\nabla_{ab}}$ is the adjoint operator of ∇_{ab} that we will define later in a discrete form. We also have

$$\tilde{K}^* = H^\intercal.$$

This provides all the ingredients needed to implement the primal-dual algorithm associated to the problem (29). Let us introduce the discrete setting.

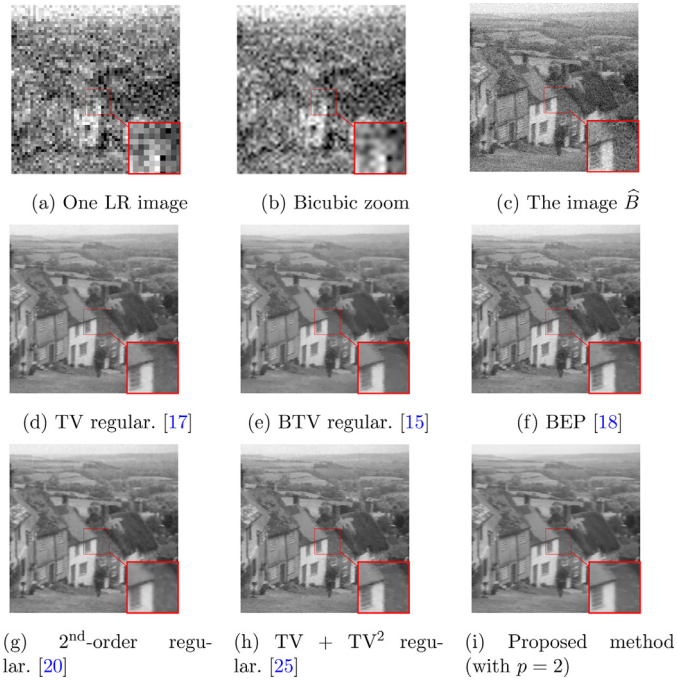


(a) One LR image (b) Bicubic zoom (c) The image $\widehat{B}$

(d) TV regular. [17] (e) BTV regular. [15] (f) BEP [18]

(g) 2^{nd}-order regular. [20] (h) TV + TV^2 regular. [25] (i) Proposed method (with $p = 2$)

Fig. 6. Super resolution of *Barbara* by different methods from 120 LR images.



(a) One LR image (b) Bicubic zoom (c) The image $\widehat{B}$

(d) TV regular. [17] (e) BTV regular. [15] (f) BEP [18]

(g) 2^{nd}-order regular. [20] (h) TV + TV^2 regular. [25] (i) Proposed method (with $p = 2$)

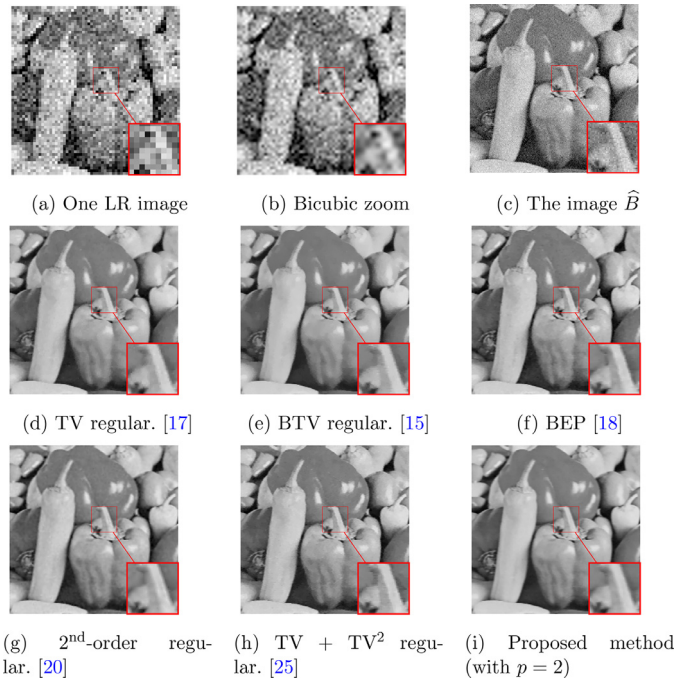
Fig. 7. Super resolution of *Goldhill* by different methods from 120 LR images.

We define the second order differential operators ∇_{xx}, ∇_{yy}, and ∇_{xy} using convolutions with the kernels

$$k_{xx} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad k_{yy} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$k_{xy} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007



(a) One LR image (b) Bicubic zoom (c) The image $\widehat{B}$

(d) TV regular. [17] (e) BTV regular. [15] (f) BEP [18]

(g) 2^{nd}-order regular. [20] (h) TV + TV^2 regular. [25] (i) Proposed method (with $p = 2$)

Fig. 8. Super resolution of *Peppers* by different methods from 120 LR images.

As a consequence, the matrix K_2 in (27) can be interpreted as the kernel k_2 defined as:

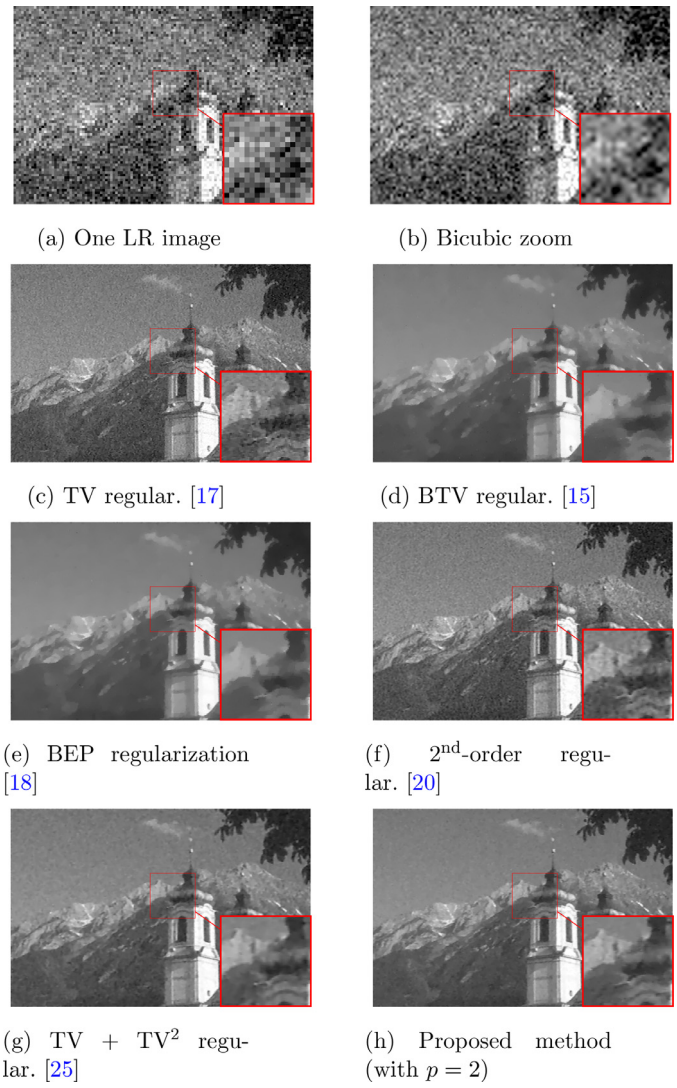
$$k_2 = k_{xx} + k_{yy} + 2k_{xy} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

Therefore, the result of K_2X is merely the convolution of X with the above linear kernel k_2. Moreover, the transpose of K_2^* is given by the convolution with the rotated kernel $\tilde{k}_2$ in the case of periodic boundary condition exploited in this paper. Algorithm 1 is used to perform the deconvolution and denoising of the proposed SR method.

Data: Choose τ, $\sigma > 0$, $(\widehat{Y}^0, \widehat{Z}^0, \widehat{X}^0) \in \mathbb{R}^M \times \mathbb{R}^M$ and set $\widehat{Y}^0 = \widehat{Z}^0 = \widehat{X}^0$.
Result: the restored HR image $\widehat{X}$
Iterate until negligible variation of $\widehat{X}^n$:
begin
$\widehat{Y}^{n+1} = \Pi_P(\widehat{Y}^n + \sigma K X^n)$; // Use (32)
$\widehat{Z}^{n+1} = \Pi_P(\widehat{Z}^n + \sigma \tilde{K} X^n)$; // Use (33)
$\widehat{X}^{n+1} = \widehat{X}^n - \tau(K^*\widehat{Y}^{n+1} + \tilde{K}^*\widehat{Z}^{n+1})$; // Use (37)
$\widehat{X}^{n+1} = 2\widehat{X}^{n+1} - \widehat{X}^n$

Algorithm 1: Primal-Dual algorithm.

In this algorithm, we choose carefully the weighting parameter δ. Indeed, we propose a new manner to control this parameter, using a function depending on the gradient of the image. Since we know the success of the BTV regularization in preserving sharp edges and of the second variational approach in recovering smoother surfaces, the ideal choice of the weighting parameter is in the surrounding of 1 along edges; and to avoid the artifacts due to the BTV regularization, the choice of the weighting parameter δ is $0 \leq \delta < 1$ in smooth regions and small jumps. Thus, we use the following formula, which yields the parameter δ giving usually the



(a) One LR image (b) Bicubic zoom

(c) TV regular. [17] (d) BTV regular. [15]

(e) BEP regularization [18] (f) 2^{nd}-order regular. [20]

(g) TV + TV^2 regular. [25] (h) Proposed method (with $p = 2$)

Fig. 9. Super resolution of *Mosq* by different methods from 20 LR images.

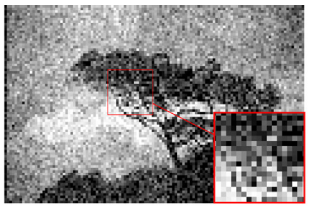
best quality of the reconstitution HR image:

$$\delta = \begin{cases} \exp(-\frac{\|\nabla X\|_1^2}{2M^2}) & \text{if } \|\nabla X\|_1/M \leq T \\ \exp(-\frac{T^2}{2}) & \text{if } \|\nabla X\|_1/M > T \end{cases} \quad (38)$$

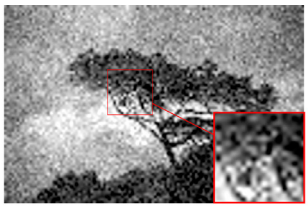
where T is the Otsu binarization threshold (Otsu, 1975). To approve this choice, we used several test images; while sometimes we find a better result for a hand tuned parameter, this parameter is very close to the controlled weighting parameters calculated by the function.

To compare the performance of the proposed algorithm to other regularization methods, we set up the following deconvolution and denoising experiment. First, we added a normalized 5×3 Gaussian kernel of zero mean and unity variance blur, then, we added Gaussian white noise of zero mean and standard deviation $\sigma = 30$ to the original *Lena* image. Fig. 1 shows the obtained image using different regularizations. We compare our method with TV (Ng et al., 2007), BTV (Farsiu et al., 2004), 2nd-order regularization (Bergounioux and Piffet, 2010), and also the combined first order and second order (TV + TV^2) regularizations (Papafitsoros and Schönlieb, 2014). We select the regularization parameters according to the best visual result. This example demonstrates that the proposed regularization functional not only produces sharp edges but also avoids the staircasing effect.

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007

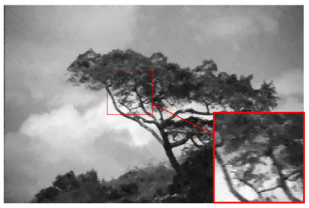


(a) One LR image

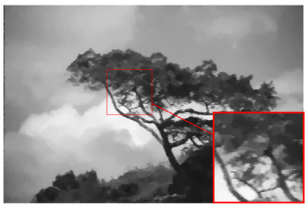


(b) Bicubic zoom

(c) TV regular. [17]

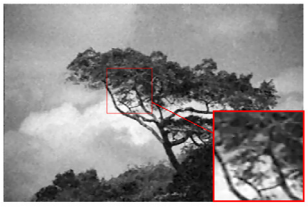


(d) BTV regular. [15]

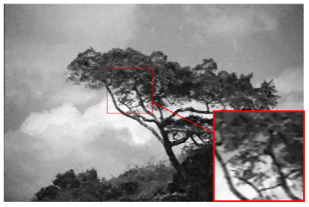


(e) BEP regularization [18]

(f) 2^{nd}-order regular. [20]

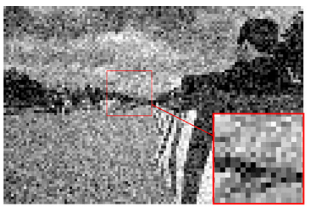


(g) TV + TV^2 regular. [25]



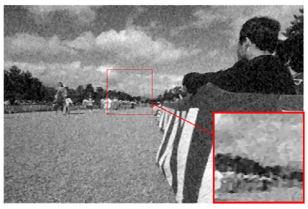
(h) Proposed method (with $p = 2$)

Fig. 10. Super resolution of *Tree* by different methods from 20 LR images.

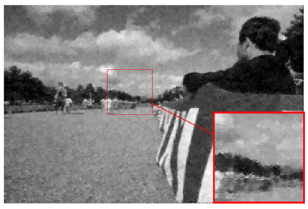


(a) One LR image

(b) Bicubic zoom



(c) TV regular. [17]



(d) BTV regular. [15]

(e) BEP regularization [18]

(f) 2^{nd}-order regular. [20]

(g) TV + TV^2 regular. [25]

(h) Proposed method (with $p = 2$)

Fig. 11. Super resolution of *Festival* by different methods from 20 LR images.

4. Experimental results

In this section, both synthetic and real images are tested to judge the performance of our proposed method, and compare it to some popular approaches. Some of the used benchmark images are listed in Fig. 2. The chosen images are often used in image processing applications, since they differ in texture, smoothness and gray level histogram. To generate the LR images from these HR images, we follow the classical steps: First, we define the sub-pixel motion using a non parametric registration (Laghrib et al., 2014), after, we align the images using the registration process (Laghrib et al., 2014), in the third and fourth step we blur and decimate the image respectively and finally we add an additive zero-mean white Gaussian noise with a chosen Signal-to-Noise Ratio level (SNR). We can now justify the contribution of this proposed combination in noise, blur and misregistration errors removing using different generated LR images. Hence, to evaluate the performance of the SR process, we use two measures such as peak signal to noise ratio (PSNR) and the structural similarity (SSIM). The PSNR is a popular metric used to measure the quality of the estimated HR image, while the SSIM is a complementary measure, which gives an indication of image quality based on known characteristics of the human visual system. In the next section, we describe the parameters chosen in the simulation results.

4.1. Simulation results

4.1.1. Effectiveness of the proposed regularization

We use the first six image of Fig. 2, all these images have the same size 248×248. Then, we generate 120 LR images for each of the six images: we use a Gaussian kernel with standard deviation $\sigma = 1.5$ truncated in a 3×3 window to blur it, and subsample it by a factor $r = 4$. Notice that this introduces severe aliasing (a blur of $\sigma = 4 \times 0.8 = 3.2$ would be necessary according to Morel and Yu, 2011) and ringing (small window size). In addition to measure the robustness of our proposed method against noise, we add an additive white Gaussian noise with standard deviation $\sigma = 10$ to all the generated LR obtained from the six benchmark images. The initial HR image X^0 in the primal-dual algorithm is obtained by a bicubic interpolation of the LR image Y_1, the scalar weight $\alpha = 0.7$, the best choice of the spatial window size coincides always with $p = 2$. We stop at the first iteration n where $\|\widehat{X}^{n+1} - \widehat{X}^n\|_1 < 10^{-4}\|\widehat{X}^n\|_1$. We justify the effectiveness of our method by comparing it with popular methods used in SR, such as bicubic, TV (Ng et al., 2007), BTV (Farsiu et al., 2004), BEP regularizer (Zeng and Yang, 2013), the second order variational regularization used in denoising problem (Bergounioux and Piffet, 2010), and also the combined first order and second order (TV + TV^2) regularization (Papafitsoros and Schönlieb, 2014). We show the reconstructed HR images $\widehat{X}$ and also the obtained im-

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007

Table 1
PSNR and SSIM results obtained by applying different methods to the benchmark images. In ***bold*** the best (highest) score of each row is shown.

Image			Method						
	σ noise	Metric	Bicubic	TV	second order	BEP	BTV	TV+TV^2	proposed
Baboon 	20	PSNR	17.95	25.69	25.79	26.50	26.45	26.50	**26.74**
		SSIM	0.229	0.738	0.744	0.788	0.781	0.789	**0.801**
	30	PSNR	16.98	24.22	24.57	25.11	24.96	25.09	**25.77**
		SSIM	0.201	0.685	0.688	0.713	0.699	0.704	**0.746**
Cameraman 	20	PSNR	18.22	26.48	26.52	27.52	26.63	26.93	**27.58**
		SSIM	0.288	0.852	0.837	0.848	0.855	0.864	**0.892**
	30	PSNR	17.88	25.43	25.49	26.48	26.04	26.18	**26.89**
		SSIM	0.200	0.783	0.768	0.796	0.795	0.802	**0.827**
House 	20	PSNR	18.62	33.27	32.45	33.53	33.32	33.90	**34.40**
		SSIM	0.326	0.870	0.858	0.875	0.872	0.878	**0.900**
	30	PSNR	17.55	31.82	31.49	32.88	32.05	32.92	**33.21**
		SSIM	0.256	0.838	0.833	0.814	0.810	0.838	**0.876**
Barbara 	20	PSNR	18.83	26.97	27.31	27.68	27.02	27.34	**27.87**
		SSIM	0.412	0.772	0.787	0.796	0.770	0.787	**0.813**
	30	PSNR	17.41	26.15	26.93	26.94	26.55	26.85	**27.03**
		SSIM	0.328	0.709	0.750	0.752	0.750	0.754	**0.763**
Goldhill 	20	PSNR	20.19	29.16	29.15	**29.99**	29.02	29.24	29.98
		SSIM	0.487	0.832	0.829	0.847	0.820	0.836	**0.868**
	30	PSNR	19.45	28.00	28.11	**28.68**	28.22	28.29	28.66
		SSIM	0.338	0.771	0.763	0.809	0.779	0.790	**0.820**
Peppers 	20	PSNR	18.39	30.74	30.58	31.40	30.81	31.37	**31.65**
		SSIM	0.438	0.889	0.885	0.893	0.889	0.883	**0.920**
	30	PSNR	17.88	29.73	29.80	30.15	29.24	30.08	**30.72**
		SSIM	0.318	0.855	0.856	0.853	0.839	0.835	**0.896**

ages $\widehat{B}$ from the fusion step in Figs. 3–8. By a visual evaluation, we can see the effectiveness of the proposed method in removing jagged artifacts. Note that we select the regularization parameters according to the best visually pleasant result and the highest PSNR in all the experiments for the other methods. In Table 1, the SSIM and PSNR values are computed, with two standard deviations σ for noise. The best results are represented by a bold number. Usually, the proposed method outperforms the others in terms of both PSNR and SSIM. Even if the result obtained by the BEP method exceeds ours by a narrow margin in the *Goldhill* figure in terms of PSNR, we can find a better result if we change the regularization parameter's value δ.

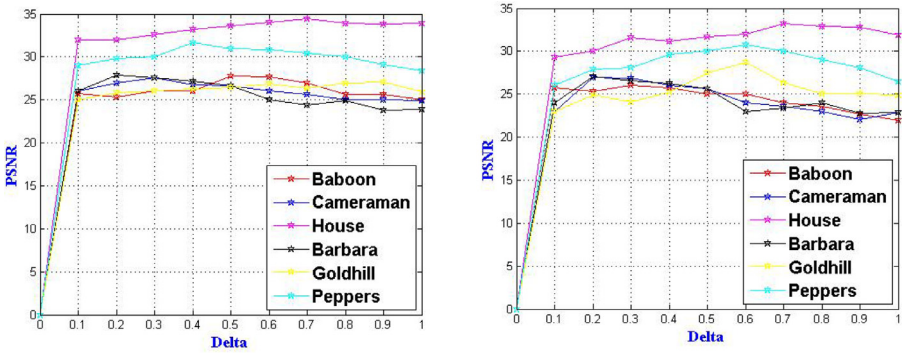
We reduce now the number of the LR frames and we present two simulated tests to show that even if the number of the chosen LR images is small, the proposed approach still present the best performance. We construct then 20 LR images for two given images: we blur firstly the LR sequence using a Gaussian kernel with $\sigma = 1.5$ truncated in a 3×3 window, and then subsample it by a factor $r = 4$. Finally we introduce an additive white Gaussian noise with high standard deviation (we take $\sigma = 30$) to all the generated LR sequence for the two tests. In Figs. 9 and 10, we show the reconstructed HR images compared with the other methods. To evaluate quantitatively the performance of the proposed method, the information fidelity criterion (IFC) index (Sheikh et al., 2005) is used which has highest correlation with perceptual scores for super resolution evaluation (Yang et al., 2014). The main principle of the IFC metric is to evaluate the image information loss by extracting wavelet features with focus on highfrequency details rather than low-frequency ones. This metric is known for its adequation to human perception, which the visual perception is more sensitive to high-frequency details of SR images rather than low-frequency components. We present in Table 2 the IFC value corresponding to the obtained result for the two sequences *Mosq* and *Tree* compared with the other methods. Knowing that the best result is in bold number, we can see that the proposed method is always better than the others.

Table 2
IFC index obtained by applying different methods. In ***bold*** the best (highest) score of each dataset is shown.

The algorithm used	IFC value (Sheikh et al., 2005)
Example 1 (*Mosq*)	
bicubic interpolation	0.1586
Super resolution with TV	1.7156
Super resolution with TV2	1.8000
Super resolution with BTV	1.6450
Super resolution with BEP	1.6803
Super resolution with TV+TV2	1.7622
The proposed method	**2.1412**
Example 2 (*Tree*)	
bicubic interpolation	0.3835
Super resolution with TV	2.0376
Super resolution with TV2	2.1100
Super resolution with BTV	2.1868
Super resolution with BEP	2.0051
Super resolution with TV+TV2	2.1057
The proposed method	**2.4652**

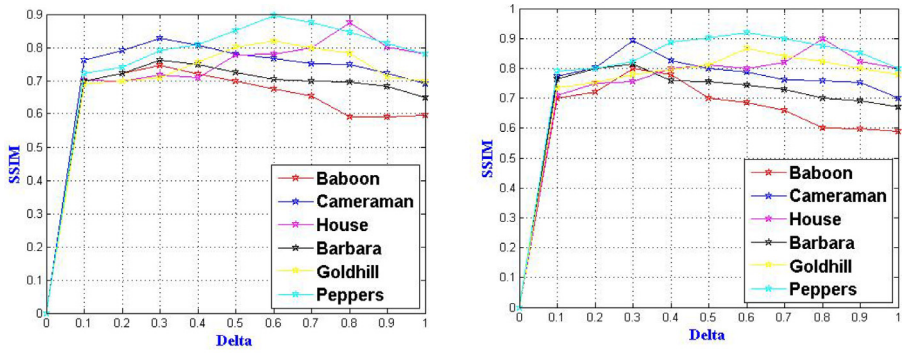
Finally, we test the performance of the proposed regularization term in handling registration and also PSF errors. We simulate then the errors due to misregistration by introducing a registration parameter error in a selected one LR *Festival* image corresponding to a translation error of 6 pixels on the HR image grid, while the PSF is assumed to be a normalized 6×4 Gaussian kernel instead of the real estimation which is a 2×2 Gaussian kernel blur. In addition, to generate more outliers, we add an additive white Gaussian noise with standard deviation $\sigma = 25$. The obtained reconstructed images are shown in Fig. 11 using different approaches. We can see that the proposed regularization is more effective in handling the registration and PSF estimation errors than the other regulariza-

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007



(a) For $\sigma = 20$ (b) For $\sigma = 30$

Fig. 12. The variation of the PSNR value with respect to δ from 20 LR images.



(a) For $\sigma = 20$ (b) For $\sigma = 30$

Fig. 13. The variation of the SSIM value with respect to δ from 20 LR images.

tions. However, we can observe the apparition of some blur near boundaries which is considered as a limitation of the proposed method.

4.1.2. Effectiveness of the parameter selection

To justify the effectiveness of the proposed adaptive parameter selection, we calculate the parameter δ for each test using the formula (38) and we compare it with the best one obtained manually with respect to the highest PSNR and SSIM values as presented in the Figs. 12 and 13. For example we find $\delta = 0.68$ when applying (38), while the obtained manual value with respect to the best PSNR and SSIM measures is about 0.69. For the *Cameraman* image, the best choice is $\delta = 0.3$ manually and also by using the Eq. (38), which confirms the effectiveness of the proposed model selection of the regularization parameter δ. Note that sometimes we do not obtain the best choice of δ using (38) but it is always very close to the best one, which is the case of the *Barbara* and *Goldhill* images. In fact, we find $\delta = 0.27$ for the Barbara image while the best value is $\delta = 0.2$, otherwise, for the *Goldhill* image, we obtain $\delta = 0.8$ while the best is $\delta = 0.9$.

4.1.3. Effectiveness of the proposed algorithm

The proposed primal-dual algorithm converges in fewer iterations compared with some classical iteratively least squares implementation. In Table 3, the execution time of the proposed method compared with the others is shown. Not that the implementation of the proposed algorithm is done in on an Intel i7 3.4 GHz Quad-Core computer using MATLAB R2014. We can deduce that the proposed algorithm has always the best execution time. We recall that the other approach re also implemented in MATLAB and used the gradient descent algorithm.

4.2. Real experiment

In this section, we discuss the performance of our proposed algorithm on natural data. We use in the first experiment the "Alpaca" video sequence of 55 compressed grayscale frames of size 96×128, downloaded from Milanfar's web site[1]. Without clue of the camera's PSF, we assume that it is a 11×11 Gaussian kernel with standard deviation equal to 1.5. In Fig. 14, we show the reconstituted HR image from 4 LR frames taken from the video. The resolution is augmented by a factor $r = 2$ and we compare our algorithm with the other SR method used in the previous tests. If we focus on what happens around the edges of the obtained HR image, we can see that our method can restore the characters in the images at least as well as the best other SR algorithms.

In the second experiment, we use 16 low-resolution images obtained from the *EIA* video sequence[2] of size 90×90. We assume that the PSF is a 3×3 Gaussian kernel with standard deviation equal to 1.5. In Fig. 15, we show the result obtained by different SR algorithms from 16 LR frames taken from the *EIA* video. The resolution is augmented by a factor $r = 3$ and we compare also our algorithm with the other SR method used in the previous tests. Since the video of *EIA* has a lot more noise than *Alpaca*, we can see the robustness of the proposed algorithm against noise compared with the other SR methods.

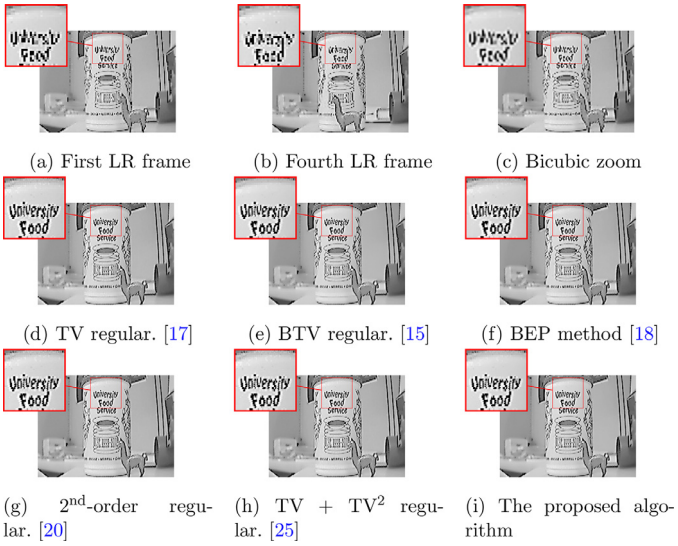
In the third and fourth experiments, we use respectively 31 frames of two sequences: *Whittard* and *Sign*, used in Ma et al. (2015), which can be downloaded from the internet.[3] These experiments differ from the above ones since the LR frames

[1] https://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html.
[2] https://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html.
[3] http://www.cse.cuhk.edu.hk/~leojia/projects/mfsr/index.html.

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007

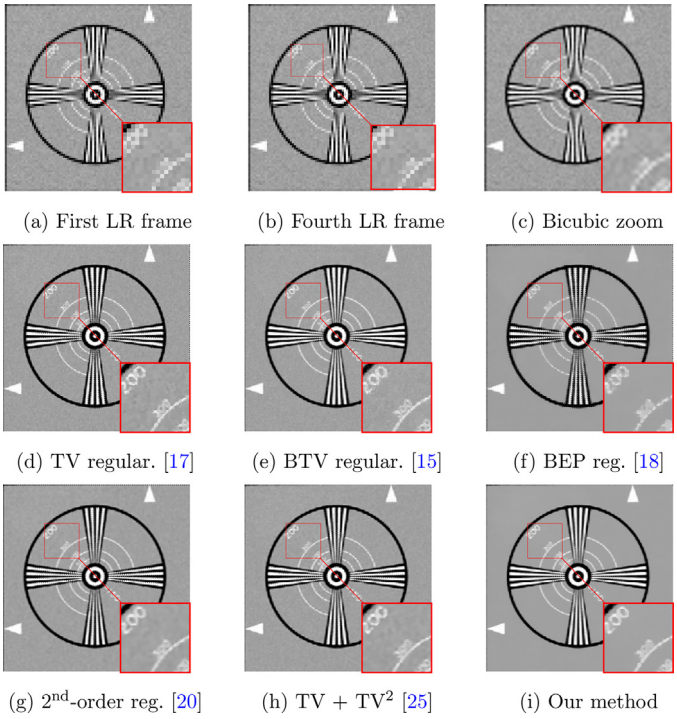
Table 3
CPU times (in seconds) when the magnification factor is 4.

Image	CPU Time			
	BTV reg. (Farsiu et al., 2004)	TV regular. (Ng et al., 2007)	BEP method (Zeng and Yang, 2013)	Our Method
Cameraman	10.24	10.26	18.0	7.00
House	11.03	11.60	17.12	5.94
Barbara	9.66	9.75	16.14	6.92
Goldhill	9.88	9.18	16.06	6.90
Baboon	10.07	12.00	18.00	6.80
Peppers	8.92	9.99	16.40	4.06
Mosq	9.6	10.6	15.55	5.02
Tree	10.20	11.2	14.44	4.6
Festival	11.90	10.88	16.00	6.3
Average time	10.16	10.60	16.41	5.94



(a) First LR frame (b) Fourth LR frame (c) Bicubic zoom
(d) TV regular. [17] (e) BTV regular. [15] (f) BEP method [18]
(g) 2nd-order regular. [20] (h) TV + TV^2 regular. [25] (i) The proposed algorithm

Fig. 14. Results on the first four frames of the *Alpaca* sequence.



(a) First LR frame (b) Fourth LR frame (c) Bicubic zoom
(d) TV regular. [17] (e) BTV regular. [15] (f) BEP reg. [18]
(g) 2nd-order reg. [20] (h) TV + TV^2 [25] (i) Our method

Fig. 15. Results on the *EIA* sequence from 16 LR images.

(a) First LR frame (b) Fourth LR frame (c) Bicubic zoom
(d) TV reg. [17] (e) BTV reg. [15] (f) BEP reg. [18]
(g) 2nd-order reg. [20] (h) TV + TV^2 [25] (i) Our method

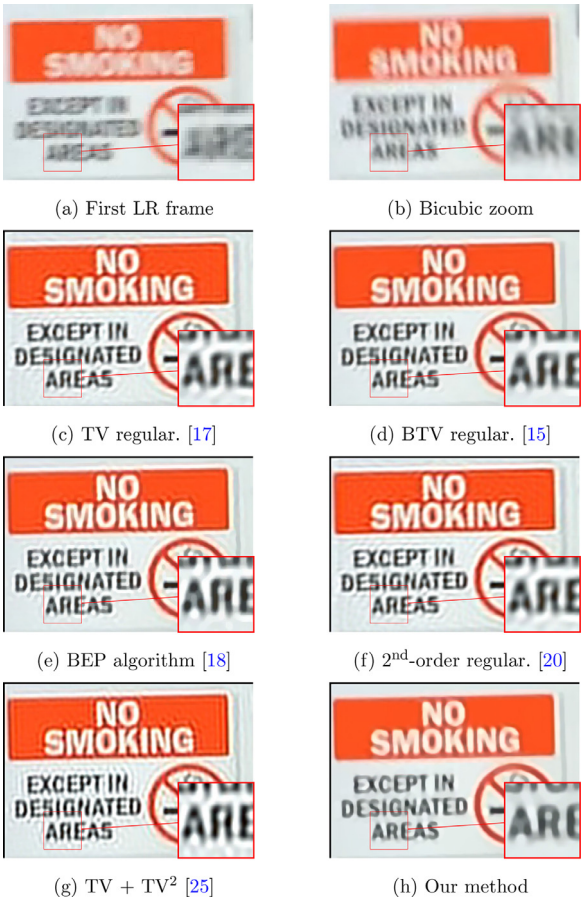
Fig. 16. Results on the *Whittard* sequence from 31 images.

are much blurred and the PSF is unknown. We use the approach in Bergen et al. (1992) to estimate the motion vectors while the unknown camera PSF is assumed to be a 5×5 and 7×6 Gaussian kernel with standard deviation equal to 1.5 for the *Whittard* and Sign sequences respectively. The desired resolution enhancement factor was set to be 4 for the two tests. In Figs. 16 and 17, we show the obtained HR image from 31 LR frames taken from the two sequences of *Whittard* and *Sign* respectively. Once again we can see that our proposed method can restore the characters in the two tests better than the compared methods, which confirms the efficiency of our method in handling high blur level.

5. Conclusion

In this paper, a novel energy has been proposed for multiframe super-resolution image reconstruction problem. We presented a new combination of BTV regulizer and a second order variational problem using a controlled weighting parameter. Since the BTV term is effective in preserving sharp edges in images when removing blur and noise, the second order term has been introduced to avoid the undesirable staircasing effect. To confirm the choice of this combination, the existence of a minimizer has been proved using a relaxation technique. Also, the results have been performed

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007



(a) First LR frame (b) Bicubic zoom

(c) TV regular. [17] (d) BTV regular. [15]

(e) BEP algorithm [18] (f) 2^{nd}-order regular. [20]

(g) TV + TV^2 [25] (h) Our method

Fig. 17. Results on the *Sign* sequence from 31 images.

on both the simulated and real images, and the proposed method has confirmed its effectiveness visually and quantitatively using the PSNR, SSIM and IFC criteria. One remaining open question is the uniqueness of the minimizer. We could neither prove nor discard this hypothesis. Another interesting path of investigation is the use of a non-local term instead of the bilateral term.

Compliance with ethical standards

- Funding: This research was entirely funded by the respective institutions of the authors.
- Conflict of interest: The authors declare that they have no conflict of interest.
- Neither human participants nor animals are involved in this research.

Acknowledgments

The authors are grateful to the anonymous reviewers for their insightful remarks and corrections. Their feedback had a great influence on the improved quality compared to the first draft version of this paper.

References

Aubert, G., Kornprobst, P., 2006. Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations, 147. Springer Science & Business Media.
Bergen, J., Anandan, P., Hanna, K., Hingorani, R., 1992. Hierarchical model-based motion estimation. In: Computer Vision–ECCV'92. Springer, pp. 237–252.
Bergounioux, M., Piffet, L., 2010. A second-order model for image denoising. Set-Valued Var. Anal. 18 (3–4), 277–306.
Bredies, K., Kunisch, K., Pock, T., 2010. Total generalized variation. SIAM J. Imaging Sci. 3 (3), 492–526.
Brezis, H., 2010. Functional Analysis, Sobolev Spaces and Partial Differential Equations. Springer Science, New York.
Caselles, V., Sapiro, G., Chung, D.H., 2000. Vector median filters, inf-sup operations, and coupled PDE's: theoretical connections. J. Math. Imaging Vis. 12 (2), 109–119.
Chambolle, A., Pock, T., 2011. A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. 40 (1), 120–145.
Chen, Y., Pock, T., Ranftl, R., Bischof, H., 2013. Revisiting loss-specific training of filter-based MRFs for image restoration. In: Pattern Recognition. Springer, pp. 271–281.
Chen, Y., Ranftl, R., Pock, T., 2014. Insights into analysis operator learning: from patch-based sparse models to higher order MRFs. IEEE Trans. Image Process. 23 (3), 1060–1072.
Demengel, F., 1985. Fonctions à hessien borné. Annales de l'Institut Fourier 34 (2), 155–190.
Demengel, F., Temam, R., 1984. Convex functions of a measure and applications. Indiana Univ. Math. J. 33 (5), 673–709.
Demengel, G., Demengel, F., 2012. Espaces fonctionnels. utilisation dans la résolution des équations aux dérivées partielles. EDP Sciences.
El Mourabit, I., El Rhabi, M., Hakim, A., Laghrib, A., Moreau, E., 2017. A new denoising model for multi-frame super-resolution image reconstruction. Signal Process. 132, 51–65.
Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P., 2004. Fast and robust multiframe super resolution. IEEE Trans. Image Process. 13 (10), 1327–1344.
Hill, D.L.G., Batchelor, P.G., Holden, M., Hawkes, D.J., 2001. Medical image registration. Phys. Med. Biol. 46 (3), R1.
Jin, R., Zhao, S., Xu, X., Song, E., 2016. Multiframe super-resolution based on half-quadratic prior with artifacts suppress. J. Vis. Commun. Image Represent. 40, 656–670.
Laghrib, A., Ghazdali, A., Hakim, A., Raghay, S., 2016. A multi-frame super-resolution using diffusion registration and a nonlocal variational image restoration. Comput. Math. Appl. 72 (9), 2535–2548.
Laghrib, A., Hakim, A., Raghay, S., 2015. A combined total variation and bilateral filter approach for image robust super resolution. EURASIP J. Image Video Process. 2015 (1), 1.
Laghrib, A., Hakim, A., Raghay, S., El Rhabi, M., 2014. Robust super resolution of images with non-parametric deformations using an elastic registration. Appl. Math. Sci. 8 (179), 8897–8907.
Lee, E.S., Kang, M.G., 2003. Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration. IEEE Trans. Image Process. 12 (7), 806–813.
Li, Y., Santosa, F., 1996. A computational algorithm for minimizing total variation in image restoration. IEEE Trans. Image Process. 5 (6), 987–995.
Luong, Q.H., 2009. Advanced Image and Video Resolution Enhancement Techniques Ph.D. thesis. Faculty of Engineering Ghent University.
Lysaker, M., Lundervold, A., Tai, X.-C., 2003. Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time. IEEE Trans. Image Process. 12 (12), 1579–1590.
Lysaker, M., Tai, X.C., 2006. Iterative image restoration combining total variation minimization and a second-order functional. Int. J. Comput. Vis. 66 (1), 5–18.
Ma, Z., Liao, R., Tao, X., Xu, L., Jia, J., Wu, E., 2015. Handling motion blur in multi-frame super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5224–5232.
Milanfar, P., 2010. Super-Resolution Imaging. CRC Press.
Min, L., Yang, X., Ye, D., 2014. Well-posedness for a fourth order nonlinear equation related to image processing. Nonlinear Anal. Real World Appl. 17, 192–202.
Modersitzki, J., 2003. Numerical Methods for Image Registration. Oxford University Press, USA.
Morel, J.-M., Yu, G., 2011. Is SIFT scale invariant? Inverse Prob. Imaging 5 (1), 115–136.
Ng, M.K., Shen, H., Lam, E.Y., Zhang, L., 2007. A total variation regularization based superresolution reconstruction algorithm for digital video. EURASIP J. Adv. Signal Process. 1–16.
Nocedal, J., Wright, S., 2006. Numerical Optimization. Springer Science.
Otsu, N., 1975. A threshold selection method from gray-level histograms. Automatica 11 (285–296), 23–27.
Papafitsoros, K., Schönlieb, C.B., 2014. A combined first and second order variational approach for image reconstruction. J. Math. Imaging Vis. 48 (2), 308–338.
Park, S.C., Park, M.K., Kang, M.G., 2003. Super-resolution image reconstruction: a technical overview. IEEE Signal Process. Mag. 20 (3), 21–36.
Patanavijit, V., Jitapunkul, S., 2006. A robust iterative multiframe superresolution reconstruction using a Huber Bayesian approach with Huber Tikhonov regularization. In: International Symposium on Intelligent Signal Processing and Communications, Yonago, Japan, pp. 13–16.
Prasath, V.S., Vorotnikov, D., 2014. Weighted and well-balanced anisotropic diffusion scheme for image denoising and restoration. Nonlinear Anal. Real World Appl. 17, 33–46.
Rudin, L., Osher, S., Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms. Physica D 60, 259–268.
Schultz, R.R., Meng, L., Stevenson, R.L., 1998. Subpixel motion estimation for super-resolution image sequence enhancement. J. Vis. Commun. Image Represent. 9 (1), 38–50.
Sheikh, H.R., Bovik, A.C., De Veciana, G., 2005. An information fidelity criterion for image quality assessment using natural scene statistics. IEEE Trans. Image Process. 14 (12), 2117–2128.

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007

Su, H., Tang, L., Wu, Y., Tretter, D., Zhou, J., 2012. Spatially adaptive block-based super-resolution. IEEE Trans. Image Process. 21 (3), 1031–1045.
Tsai, R.Y., Huang, T.S., 1984. Multiframe image restoration and registration chapter 7. In: Advances in Computer Vision and Image Processing, 1. JAI Press, Greenwich, Conn, USA, pp. 317–339.
Unger, M., Pock, T., Werlberger, M., Bischof, H., 2010. A convex approach for variational super-resolution. In: Pattern Recognition. Springer, pp. 313–322.
Vese, L., 1996. Problémes variationnels et EDP pour l'analyse d'images et l'éevolution de courbes Ph.D. thesis. Université de Nice Sophia-Antipolis.
Villena, S., Vega, M., Babacan, S., Molina, R., Katsaggelos, A.K., 2013. Bayesian combination of sparse and non-sparse priors in image super resolution. Digit. Signal Process. 23 (2), 530–541.
Xiao, J., Pang, G., Zhang, Y., Kuang, Y., Yan, Y., Wang, Y., 2016. Adaptive shock filter for image super-resolution and enhancement. J. Vis. Commun. Image Represent. 40, 168–177.
Yang, C.-Y., Ma, C., Yang, M.-H., 2014. Single-image super-resolution: A benchmark. In: European Conference on Computer Vision. Springer, pp. 372–386.
You, Y.-L., Kaveh, M., 2000. Fourth-order partial differential equations for noise removal. IEEE Trans. Image Process. 9 (10), 1723–1730.
Zeng, X., Yang, L., 2013. A robust multiframe super-resolution algorithm based on half-quadratic estimation with modified BTV regularization. Digit. Signal Process. 23 (1), 98–109.

Please cite this article as: A. Laghrib et al., Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution, Computer Vision and Image Understanding (2017), http://dx.doi.org/10.1016/j.cviu.2017.08.007

Signal Processing 132 (2017) 51–65

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

A new denoising model for multi-frame super-resolution image reconstruction

Idriss El Mourabit[a], Mohammed El Rhabi[b], Abdelilah Hakim[a], Amine Laghrib[c,*], Eric Moreau[d]

[a] University Cadi Ayyad, Faculty of Science and Technology of Marrakech, Department of Mathematics and Informatics, Marrakech, Morocco
[b] Applied Mathematics and Computer Science Department, Ecole des Ponts ParisTech (ENPC), Paris, France
[c] LMA FST Béni-Mellal, Université Sultan Moulay Slimane, Morocco
[d] CNRS, LSIS, UMR 7296, Université de Toulon, 83957 La Garde, France

ARTICLE INFO

Keywords:
Super-resolution
Multi-frame
Image restoration
Variational regularization
Tensor diffusion

ABSTRACT

Multi-frame image super-resolution (SR) aims to combine the sub-pixel information from a sequence of low-resolution (LR) images to build a high-resolution (HR) one. SR techniques usually suffers from annoying restoration artifacts such as noise, jagged edges, and staircasing effect. In this paper, we aim to increase the performance of SR reconstitution under a variational framework using adaptive diffusion-based regularization term. We propose a new tensor based diffusion regularization that takes the benefit from the diffusion model of Perona–Malik in the flat regions and use a nonlinear tensor derived from the diffusion process of Weickert filter near boundaries. Thus, the proposed SR approach can preserve important image features (sharp edges and corners) much better while avoiding artifacts. The synthetic and real experimental results show the effectiveness of the proposed regularisation compared to other methods in both quantitatively and visually.

1. Introduction

Currently, image multi-frame super-resolution reconstruction [1–5] is one of the relevant inverse problems research in image processing. The aim of this technique is to reconstruct a high-resolution (HR) image from a set of low-resolution (LR) ones that are noisy, blurred, deformed and down-sampled [6,7]. Multi-frame SR techniques is used in many applications and various fields. For instance, to achieve high recognition rates of the quality assurance in industries [8,9], agricultural surveillance [10], environmental issues [11], security and video surveillance [12,13] and also cultural heritage [14,15], resolution enhancement is therefore necessary. Moreover, to address the hardware limitations and the very expensive price that sophistication of hardware components requires, the SR technique is preferable, namely in areas like medical diagnostics [16] and satellite imaging [17] such as remote sensing and military surveillance, etc.

Multi-frame super-resolution is achieved in three basic steps [18]: first, the motions between each two frames of the captured image sequence are computed which constitutes the registration step; then, an alignment of this frames onto an HR grid is performed using the motion vectors; and, eventually, a denoising and deblurring step is required to reduce the noise, blur and misregistration errors resulting from the registration step. After the first work proposed in [19], where the authors considered a frequency domain approach, several approaches have been proposed and studied to improve the multi-frame SR problem [20–24]. Earlier works on SR algorithms are based on regularization method due to its ill-posed nature which mainly contains the likelihood and prior function [25,26]. The likelihood function measures the difference between the LR images and the obtained HR one, while the image prior function, impose some prior knowledge on the desired HR image. Some of the widely used prior functions was the Total Variation (TV) regularization as in [27] which avoid the edge-blurring effect caused by the Tikhonov regularization [22]. Besides, some minima estimators (M-estimators) based on statistics knowledge have been explored in multiframe super-resolution algorithm [28]. Namely, the M-estimation method in the context of regularization framework, such as the use of Huber function in the fidelity term [29]. More robust estimators was introduced without regularisation using the Lorentzian error norm [30] and Gaussian error norm [31]. Other robust M-estimation combine with a robust estimators and the Lorentzian error norm in both the fidelity and regularization terms [32,33]. The main drawbacks of these methods are the non-convexity of the minimizing function even if the robustness against outliers is significant.

One of the successful and simplest choice of the prior function in the SR problem was the Bilateral Total Variation (BTV) [34], which the

* Corresponding author.
E-mail address: laghrib.amine@gmail.com (A. Laghrib).

http://dx.doi.org/10.1016/j.sigpro.2016.09.014
Received 21 March 2016; Received in revised form 2 September 2016; Accepted 22 September 2016
Available online 23 September 2016

HR image is obtained by replacing every pixel with a weighted average of its neighborhood. Although the BTV regularization preserves edges typically in non-smooth regions, the use of the Lebesgue (L^1) norm in the regularization term produce artificial edges in the flat surfaces. To overcome the weakness of the BTV term Zeng and Yang proposed the adaptive Bilateral Edge-Preserving (BEP) [35] norm, which stop the diffusion process of the BTV term in smooth areas of the image. In the same principle, another adaptive combined of the TV and BTV norms was introduced to avoid the staircasing effect, which increase the performance of the restoration step of the SR algorithm [36]. Recently, an adaptive diffusion-based regularizer was treated to preserve important image features [37]. Even if this approach avoids undesirable artifacts while suppressing noise, it suffers from the blurring effect.

The main goal of this paper consists of increasing the robustness of the super-resolution techniques over the methods discussed above with respect to the misregistration errors, blurring effect and noise. To avoid the errors arising from the restoration step, we use a nonlinear diffusion-based regularizer derived from the Weickert filter [38–40]. Since the Weickert process destroys edges and generates curved structure under intense noise conditions, improvements of this filter are desired. In the context of super-resolution, we propose a more robust filter against noise and blur, which takes into account the coherence-enhancing property. This approach takes into consideration the best of Perona–Malik processes in flat regions [41], and the benefit of the Weickert filter effect near sharp edges. The proposed regularizer has proved its efficiency in avoiding the staircasing effect and blur while reducing noise, which can preserve image features much better compared with the existing diffusion-based regularizers.

The outline of the paper is the following one. In Section 2, we present the general multi- frame super-resolution problem. Then, we set the derived regularized criterion. After, we introduce the variational problem and we prove the existence of a solution of the proposed equation. In Section 3, we present some experimental results, while we compare our approach with some available methods. We finally end the paper by a conclusion.

2. Muti-frame SR problem formulation

The observed images of a real scene usually are in low resolution. This is due to some degradation operators. In practice, the acquired images are decimated, corrupted by noise and suffered from blurring [18]. We assume that all low resolution images are taken under the same environmental conditions using the same sensor. The relationship between an ideal HR image X (represented by a vector of size $[r^2N^2 \times 1]$, where r is the resolution enhancement factor) and the corresponding LR ones Y_k of size $N \times N$ (represented by a vector of size $[N^2 \times 1]$), is described by the following model:

$$Y_k = WF_kHX + V_k \quad \forall\, k = 1, 2, \ldots, n, \tag{1}$$

where n is the number of LR frames, H is the blurring operator of size $[r^2N^2 \times r^2N^2]$, W represents the decimation matrix of size $[N^2 \times r^2N^2]$, F_k is a geometric warp matrix of size $[r^2N^2 \times r^2N^2]$, representing a non-parametric transformation that differs in all frames, and V_k is a vector of size $[N^2 \times 1]$ which represents the additive noise for each image.

Given a LR sequence Y_k, $k = 1, \ldots, n$, the aim of SR consists of reconstructing the original image X. Since the super-resolution problem is ill-posed (the solution is not unique), the reconstitution is very unstable. This can be fixed by requiring some prior knowledge about the image X in a Bayesian framework. Via the maximum a posteriori (MAP) estimator, the estimation of HR image is given through the following minimization problem:

$$\begin{aligned}\widehat{X} &= \underset{X}{\operatorname{argmax}}\{p(X/Y_k)\}, = \underset{X}{\operatorname{argmax}}\{p(Y_k/X).\ p(X)\},\\ &= \underset{X}{\operatorname{argmin}}\{-\log(p(Y_k/X)) - \log(p(X))\},\end{aligned} \tag{2}$$

where $p(Y_k/X)$ represents the fidelity term (i.e. the relationship between the observed images and the HR one), while $p(X)$ describes the prior knowledge imposed on the high-resolution image X. Substituting these two terms by their expressions, we define the general super-resolution model as follows:

$$\widehat{X} = \underset{\mathbf{x}}{\operatorname{argmin}}\left\{\sum_{1}^{n} \| WF_kHX - Y_k \|_2 + \delta R(X)\right\}, \tag{3}$$

where $R(X)$ is a regularisation term representing a prior knowledge on X. There is a great number of proposed regularisations in the multi-frame super-resolution context [34–36,26]. The main principle of these approaches is to provide a global minimum to the encountered minimization problem. With the same aim, recently, Maiseli et al. [37] proposed a new regularizing potential defined by

$$R(X) = \int_{\Omega} \frac{\frac{|\nabla X|}{\beta}\left(2 + \frac{|\nabla X|}{\beta}\right)}{1 + \left(\frac{|\nabla X|}{\beta}\right)^2}\, d\Omega, \tag{4}$$

where β is a shape-defining tuning constant. The purpose of this choice is to combine the edge preserving effect of the total variation [42] with the Perona–Malik diffusion behavior [41] and the effect smoothing of backward diffusion anisotropic [43]. Which can preserve sharpen edges and details. The corresponding evolution equation of the Euler–Lagrange form of (3) using the regularisation term in (4) is

$$\frac{\partial X}{\partial t} = \frac{1}{n}\sum_{k=1}^{n} W^{\top}F_k^{\top}H^{\top}(WF_kHX - Y_k) + \operatorname{div}\left(\frac{2 + \frac{|\nabla X|}{\beta}}{1 + \left(\frac{|\nabla X|}{\beta}\right)^2}\nabla X\right). \tag{5}$$

Even if this equation preserves feature of the image, it suffers from the blurring effect in the homogeneous regions. To address this problem, we propose a new equation that performs the coherence-enhancing property and avoids blur. This approach takes in consideration the best of Perona–Malik diffusion process in flat regions [41], and the benefit of Weickert filter effect [38] near sharp edges and corners. The proposed equation with Neumann boundary conditions is given as follows:

$$\begin{cases} \dfrac{\partial X}{\partial t}(t,\ x) - \operatorname{div}(D(J_\rho(\nabla X_\sigma))\nabla X) - \frac{1}{m}\sum_{i=1}^{m}(WF_kH)^{\top}(WF_kHX - Y_k) \\ \quad = 0 \quad \text{on} \quad]0,\ T[\times\Omega, \\ \langle D(J_\rho(\nabla X_\sigma))\nabla X,\ n\rangle = 0 \quad \text{on} \quad]0,\ T[\times\partial\Omega, \\ X(0,\ x) = X_0(x), \end{cases} \tag{6}$$

where D is an anisotropic diffusion tensor and J_ρ is the structure tensor defined by

$$J_\rho(\nabla X_\sigma) = K_\rho * (\nabla X_\sigma \otimes \nabla X_\sigma) = K_\rho * (\nabla K_\sigma * X \nabla K_\sigma * X^{\top}). \tag{7}$$

With K_ρ and K_σ represents a two Gaussian convolution kernels such as $K_\tau(x) = \frac{1}{2\pi\tau^2}\exp(-\frac{|x|^2}{2\tau^2})$. The function D is chosen in order to preserve edges and corners as much as possible. This function is computed using the eigenvalues and the eigenvectors of the structure tensor J_ρ as follows:

$$D := f_+(\lambda_+,\ \lambda_-)\theta_+\theta_+^{\top} + f_-(\lambda_+,\ \lambda_-)\theta_-\theta_-^{\top}, \tag{8}$$

where $\lambda_{+/-}$ and $\theta_{+/-}$ are respectively the eigenvalues and the eigenvectors of the tensor structure J_ρ, the eigenvalues $\lambda_{+/-}$ are calculated as

$$\lambda_{+/-} = \frac{1}{2}(\operatorname{trace}(J_\rho) \pm \sqrt{\operatorname{trace}^2(J_\rho) - 4\det(J_\rho)}). \tag{9}$$

This choice is justified using the geometric characteristics of the restored image, for more details see [44]. Indeed, if $\lambda_{+/-} \approx 0$, the smoothing within homogeneous areas is isotropic, while if $\lambda_+ \gg \lambda_- \approx 0$

(or $\lambda_- \gg \lambda_+ \approx 0$), the smoothing process is anisotropic and directed along the straight edges. However, if $\lambda_+ \gg \lambda_- \gg 0$, the smoothing takes into account corners. Also, the functions $f_+(\lambda_+, \lambda_-)$ and $f_-(\lambda_+, \lambda_-)$ are chosen carefully in order to satisfy the following smoothing constraints:

1. *Isotropic*: in the homogeneous areas the diffusion process should be isotropic.
2. *Anisotropic*: near of sharp edges, the diffusion process should be anisotropic.

Many choices have been proposed for the functions $f_{+/-}$, namely, the functions associated to the Perona–Malik process are defined as

$$f_{+/-}(\lambda_+, \lambda_-) = \exp\left(-\frac{(\lambda_+ + \lambda_-)^2}{k}\right).$$

Using this function, the diffusivity matrix D has no specific orientation and it stops diffusion near edges and corners. To outperform the defects of the Perona–Malik model, the model of Weickert [39] has been proposed choosing f_+ and f_- as follows:

$$\begin{cases} f_+(\lambda_+, \lambda_-) = \begin{cases} \alpha + (1-\alpha)\exp(-\dfrac{k}{(\lambda_+ - \lambda_-)^2}) & \text{if} \quad \lambda_+ \neq \lambda_-, \\ \alpha & \text{else} \quad (\alpha = 0.001), \end{cases} \\ f_-(\lambda_+, \lambda_-) = \alpha. \end{cases}$$

The Weickert coherence enhancing diffusion approach [44] leads to good results in the case of anisotropic process, particularly, the diffusion is oriented in one direction. However, the Weickert model destroys corners and singularities of the restored image. To overcome this issue, we propose new functions which leads to skipping the behaviour of Weickert coefficients and allows to diffuse in two directions θ_+ and θ_-. For that, we propose to keep the Weickert diffusion oriented in the direction of θ_+ and we modify the diffusivity coefficient f_- to consider also the direction θ_-. This choice takes into account the diffusion near a contour or a corner where the values of λ_+ or λ_- are large. For example, if λ_+ is high the diffusion is oriented in the direction of θ_-. In the same way, if λ_- is high the diffusion is oriented in the direction of θ_+. In contrast, since λ_+ and λ_- present high values near a contour, two threshold are used to identify the diffusion along the respective directions θ_+ and θ_-. A convenient choice that meets all considered requirements is the following one:

$$\begin{cases} f_+(\lambda_+, \lambda_-) = \exp\left(-\dfrac{\lambda_+}{k_1}\right), \\ f_-(\lambda_+, \lambda_-) = \exp\left(-\dfrac{\lambda_-}{k_2}\right)\left(1 - \exp\left(-\dfrac{\lambda_+}{k_1}\right)\right), \end{cases} \tag{10}$$

where k_1 and k_2 are two thresholds which define the diffusion along the directions θ_+ and θ_- respectively. Since the desired diffusion changes according to the image feature: contour or singularity, the parameters k_1 and k_2 are used to define the sensitivity threshold that determine the variation in a given image. With this choice, the PDE in (6) allows to reduce noise in uniform zones, smooth edges and preserve singularities correctly.

Since the restoration step is an ill-posed problem [45], we have to check the existence and uniqueness of the proposed PDE (6). We first recalls some interesting assumptions:

1. There exists $\eta_k > 0$ such that $\langle N_k Y, Y\rangle \leq \eta_k \parallel Y \parallel^2$, with
   $N_k = (WF_kH)^\top WF_kH, \quad \forall\, k = 1, \ldots n.$
2. The diffusion tensor D is infinitely continuous, i.e $D \in C^\infty(\mathbb{R}^{2\times 2}, \mathbb{R}^{2\times 2})$.
3. D is a symmetric and positive-definite matrix.

In the following theorem, we prove the existence and uniqueness of solution to (6).

Theorem 2.1. *Let $X_0 \in \mathbf{L}^\infty(\Omega)$, $\rho \geq 0$ and $\sigma, T > 0$. Under the assumptions above, there exists a unique function $X(t, x) \in C((0, T); \mathbf{L}^2(\Omega)) \cap \mathbf{L}^2((0, T); \mathbf{H}^1(\Omega))$ satisfying the equation (6). Moreover, the solution depends continuously on X_0 with respect to $\|.\|_{\mathbf{L}^2(\Omega)}$, and $X \in C((0, T) \times \overline{\Omega})$.*

Proof. See Appendix.□

Since we have demonstrated the existence and uniqueness of the solution to the proposed equation, the convergence of the proposed scheme is assured. Indeed, we solve the problem (6) using an explicit finite difference scheme [44]. We will not detail this part, since it is similar to previous work [36,37], where an appropriate scheme is used to discretize the proposed PDE. This scheme is computationally fast, efficient, and mathematically well posed as demonstrated in [37,44]. To compute the warping matrix F_k (to align the input images), we used a sub-pixel accurate optical flow method [46]. In this approach, a variational minimization problem is used with a Huber-Norm regularization and an L^1 based data in the optical flow constraint model.

To observe the robustness of the proposed regularizer in (6) with respect to noise, we consider a simple denoising problem. For that, we take m=1 and $WF_kH = I$ in (6), we select the image in Fig. 1(a) and added to it a zero mean Gaussian noise with standard deviation $\sigma_{noise} = 25$, then, we applied the nonlinear diffusion regularizer of Perona–Malik (PM) [41], Weickert et al. [38], BTV [34] and ours, to generate the results in Fig. 1. Concerning the parameter selection in Eq. (10), the two threshold k_1 and k_2 are chosen carefully with respect to the best obtained PSNR value as shown in Fig. 2. We use the same thing to select the parameter σ (see Fig. 2). For the parameter ρ, Weickert demonstrated in [38] that in general $\rho = 3\sigma$. We recall that the choice of these parameters depends strictly on the nature of the image, the level of noise and loss continuity of lines. We use the same techniques to determine k_1, k_2 and σ, after, in the simulation experiments within the SR framework. Figs. 3–5 show the evolution of the PSNR up to 120 iterations of the reconstruction results (with $\sigma_{noise} = 25$, $\sigma_{noise} = 30$ and $\sigma_{noise} = 35$ respectively), using different regularizers. If we look at the evolution of the PSNR associated with our approach, we can deduce that the curve is not too much affected by increasing noise levels compared with the other methods. Which once again demonstrates the effectiveness of the proposed PDE.

3. Numerical result

In this section various simulated and also real results for assessing the performance of the proposed multi-frame SR method will be presented. Comparison with some available and competitive multi-frame SR method is carried out: namely, TV [47], BTV [34], TV+BTV [36], and nonlinear diffusion regularizer [37]. The proposed super-resolution method is tested on a large benchmark image, we present only four of them. In the first simulated experiment, we degraded each of the original HR images of the Castle, EIAcen, Champignon, and Penguin to generate the corresponding sequence of twenty synthetic LR images. This sequence is obtained from the original images such as each frame is blurred by a Gaussian low-pass filter with a 3×3 and standard deviation of 1.5. Then, the blurred frames are down-sampled vertically and horizontally by a factor of r=4 and Gaussian noise was finally added, to down-sampled frames, with $\sigma_{noise} = 20$. Afterwards, we applied various multi-frame SR method chosen above, and the proposed method to reconstruct the respective original images of the degraded ones (represented in Figs. 6–9). To test and compare the performance of the proposed method with the others, a quantitative evaluation is needed, we have used two metrics: peak-signal-to-noise ratio (PSNR) [48], and mean structure similarity (SSIM) [49]. The

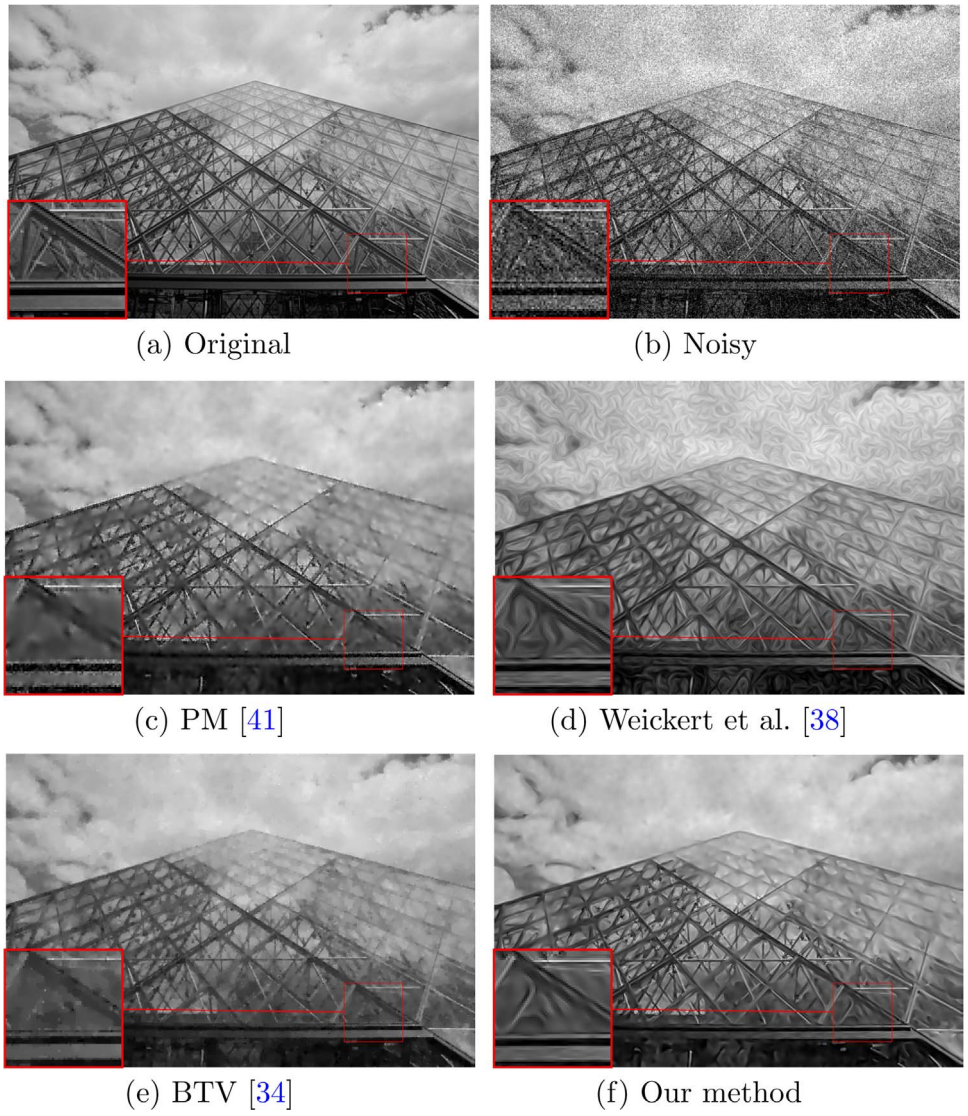


Fig. 1. The denoising process using different regularizers compared with our approach. (a) Original. (b) Noisy. (c) PM [41]. (d) Weickert et al. [38]. (e) BTV [34]. (f) Our method.

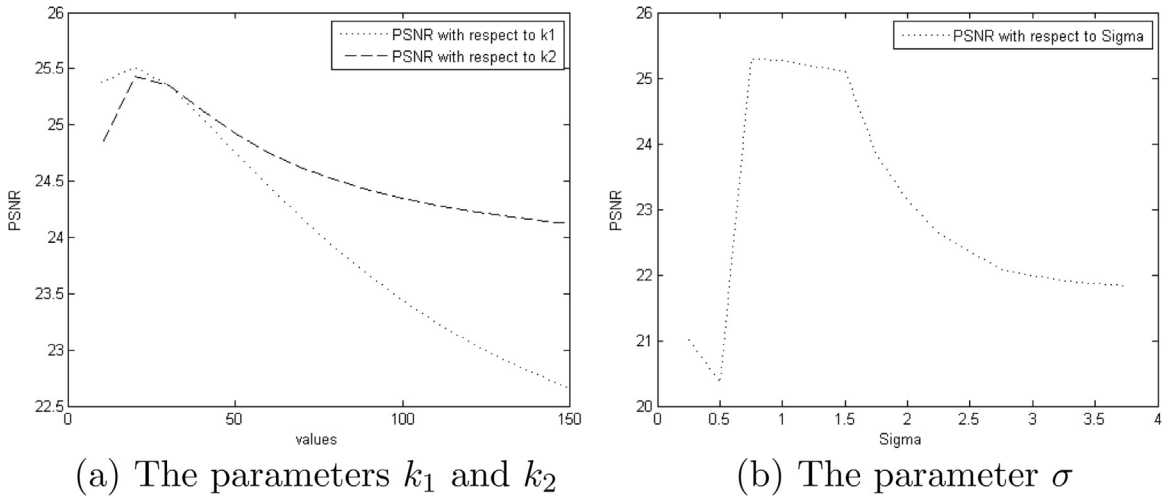
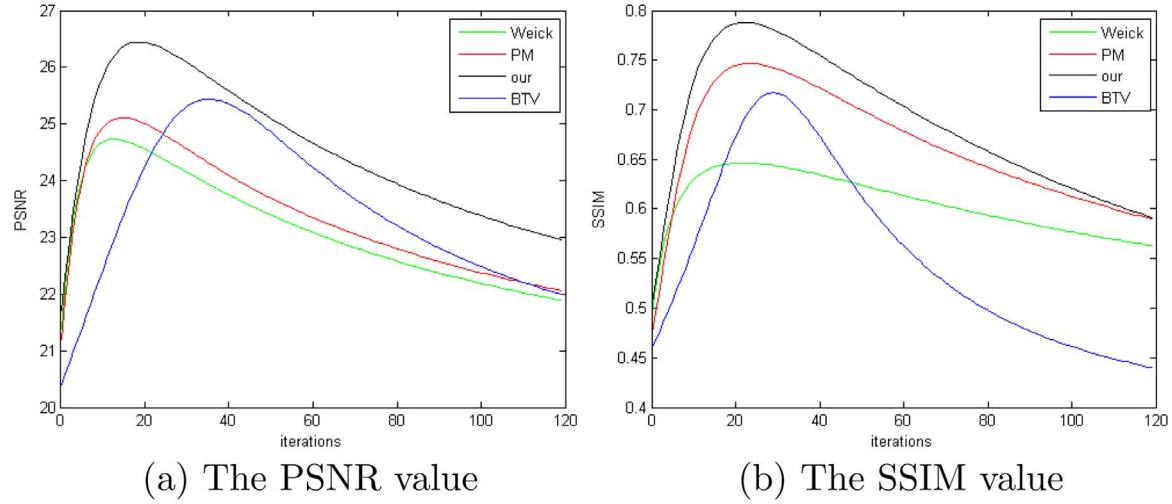
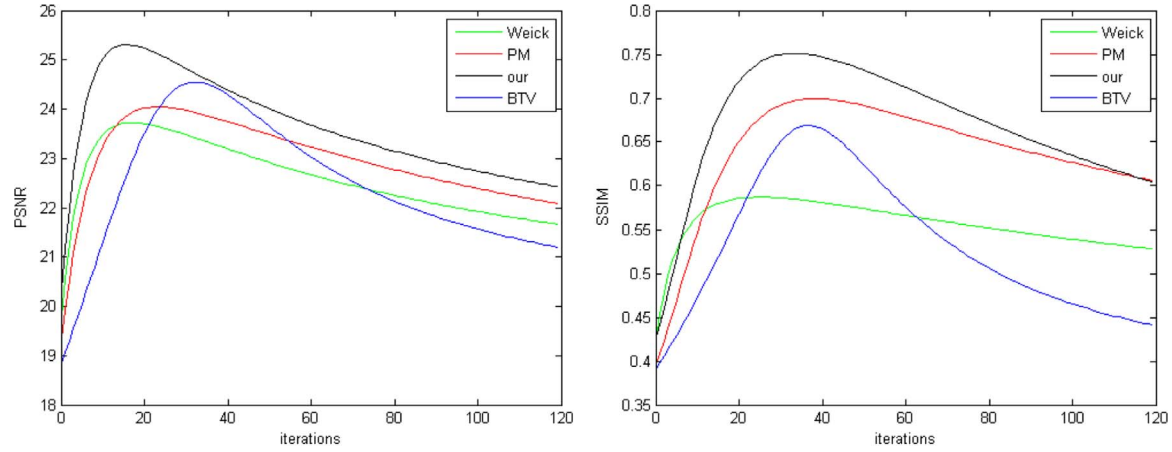


Fig. 2. The variation of the PSNR value with respect to the parameters k_1 (fixing $k_2 = 30$), k_2 (fixing $k_1 = 30$) and σ. (a) The parameter k_1 and k_2. (b) The parameter σ.



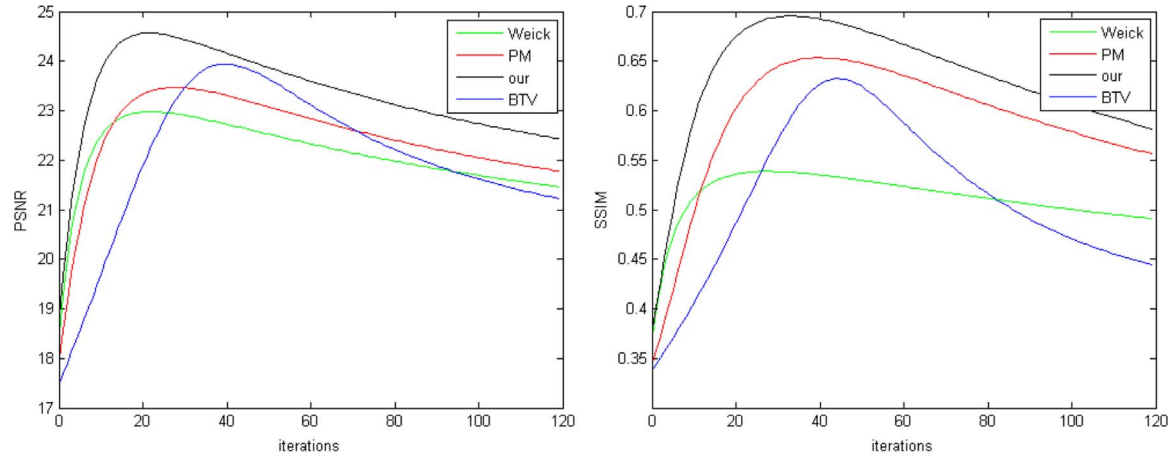
(a) The PSNR value (b) The SSIM value

Fig. 3. The variation of the PSNR and SSIM values (for the obtained images in Fig. 1) with respect to number of iterations for different regularizers. (a) The PSNR value. (b) The SSIM value.



(a) The PSNR value (b) The SSIM value

Fig. 4. The variation of the PSNR and SSIM values (with $\sigma_{noise} = 30$) with respect to number of iterations for different regularizers. (a) The PSNR value. (b) The SSIM value.



(a) The PSNR value (b) The SSIM value

Fig. 5. The variation of the PSNR and SSIM values (with $\sigma_{noise} = 35$) with respect to number of iterations for different regularizers. (a) The PSNR value. (b) The SSIM value.

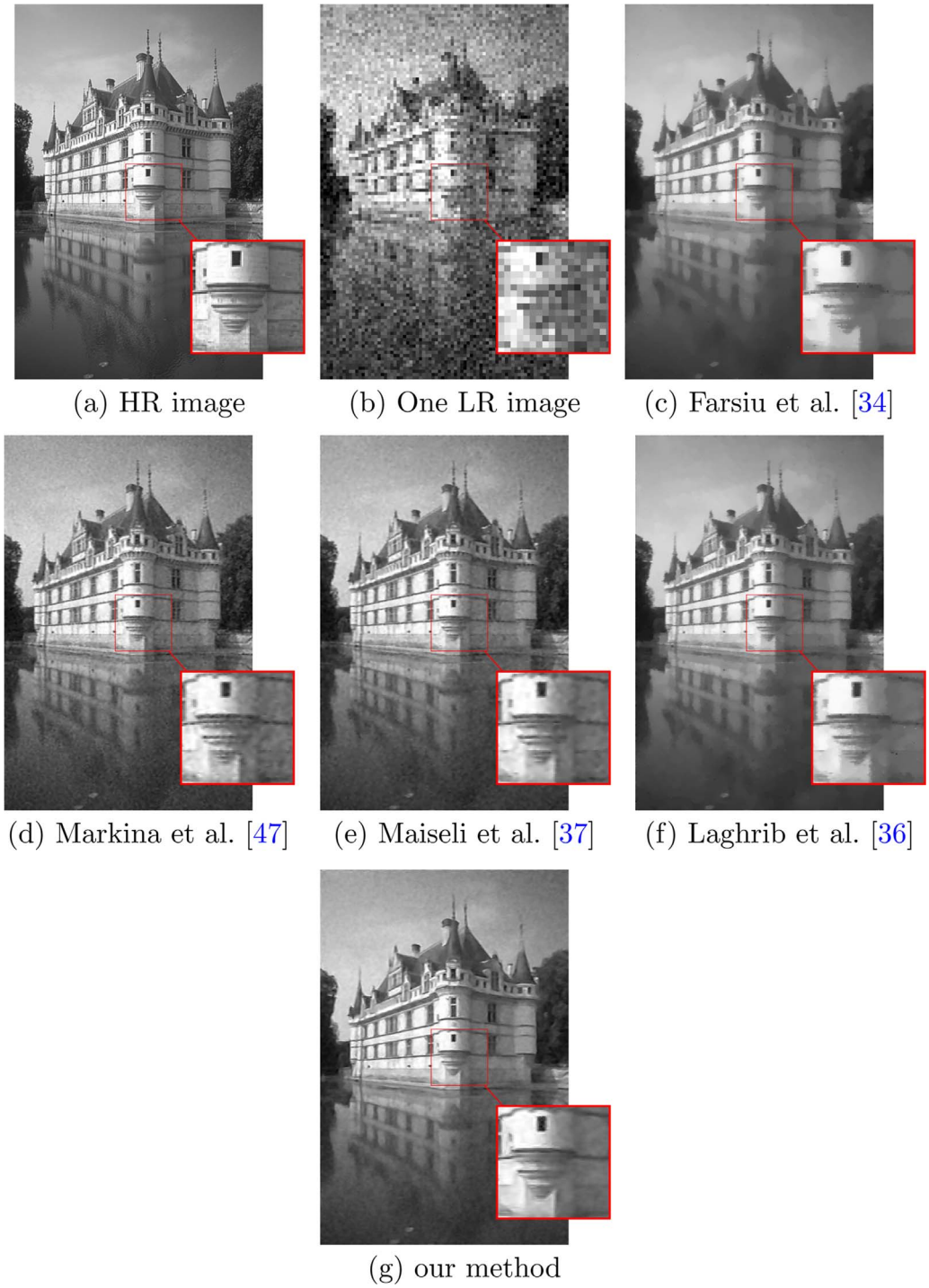


Fig. 6. The results obtained by applying different methods to LR (*Castle* sequence). (a) HR image (b) One LR image (c) Farsiu et al. [34] (d) Markina et al. [47] (e) Maiseli et al. [37] (f) Laghrib et al. [36] (g) our method.

PSNR measures signal strength relative to noise in the image and is defined by

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right),$$

where the *MSE* is the mean squared error defined by

$$MSE = \frac{1}{MN}\Sigma_{i=1}^{M}\Sigma_{j=1}^{N}(Y(i,j) - X(i,j))^2.$$

The *SSIM* is calculated on multiple windows of given image, i.e. the measurement between two windows x and y of size $N \times N$ is defined by

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_x\sigma_y + c_2)(2cov_{xy} + c_3)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)(\sigma_x\sigma_y + c_3)},$$

where the variables, respectively, defined for x and y as follows: μ_x and μ_y, mean; ${\sigma_x}^2$ and ${\sigma_y}^2$, variance; cov_{xy}, covariance; $c_1 = (k_1L)^2$, $c_2 = (k_2L)^2$ are two stabilizing constants; and L the dynamics of the pixel values, 255 for 8-bit encoded image. This metric gives an indication on the quality of the image based on the known characteristics of human visual system.

The obtained results in Figs. 6–9**?** demonstrate that, visually, the proposed method outperforms the others. Moreover, the quantita-

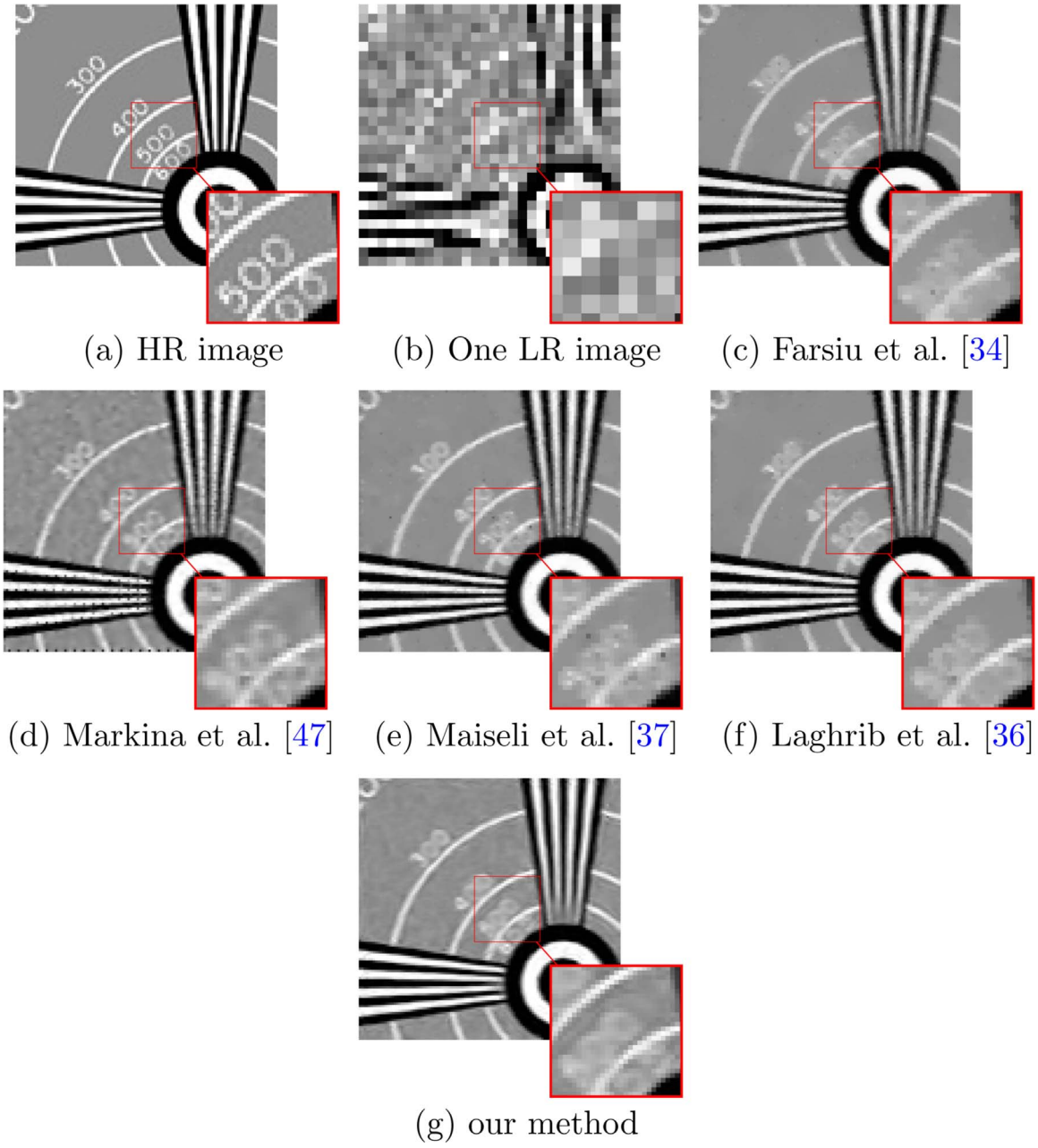


Fig. 7. The results obtained by applying different SR methods to LR *EIAcen* sequence. (a) HR image. (b) One LR image. (c) Farsiu et al. [34]. (d) Markina et al. [47]. (e) Maiseli et al. [37]. (f) Laghrib et al. [36]. (g) Our method.

tive results presented in Tables 1 and 2 confirm that the proposed model is usually with the higher PSNR and SSIM values. Note that we select the optimal parameters according to the best PSNR value in all the experiments for the other methods. Table 1 shows that our method is always with the best PSNR values in most cases, which affirm its robustness against noise and blur reduction. Also, the obtained SSIM values in Table 3 confirms the high quality of the obtained image in terms of details image preserving. Characteristically, the execution of the main implementation on MATLAB R2014 requires on the average about 7–14 s on a i7 3.4 GHz Quad-Core computer for 256×256 images. Effectively, the time becomes quite large for the color and large-size images. Furthermore, we resume the necessary execution time to increase the resolution by a factor of r=4 of the proposed algorithm compared with the others in Table 3.

To measure the robustness of the proposed approach against noise, we add another test which the Gaussian noise is added with $\sigma_{noise} = 50$ to all 100 generated LR Fish and Barbara images. Fig. 10 shows the obtained result (for the Fish image) using the proposed algorithm (for different choices of k_1 and k_2 values) compared with other approaches. By a visual evaluation, we can see the robustness of the proposed method in reducing noise with the capability of a better edge recovery compared to other methods. Moreover, as is shown in Table 4, quantitative results confirm that our proposed approach has, usually, the best performance in reconstruction in terms of the PSNR values. In addition, we can see that our approach is robust against noise compared with the others. Since, for example, we obtain the best PSNR value in the image of Barbara for $\sigma_{noise} = 50$, while for $\sigma_{noise} = 10$, Maiseli et al. [37] method was the best. Despite the promising results, the proposed method suffers from the weakness that it does not

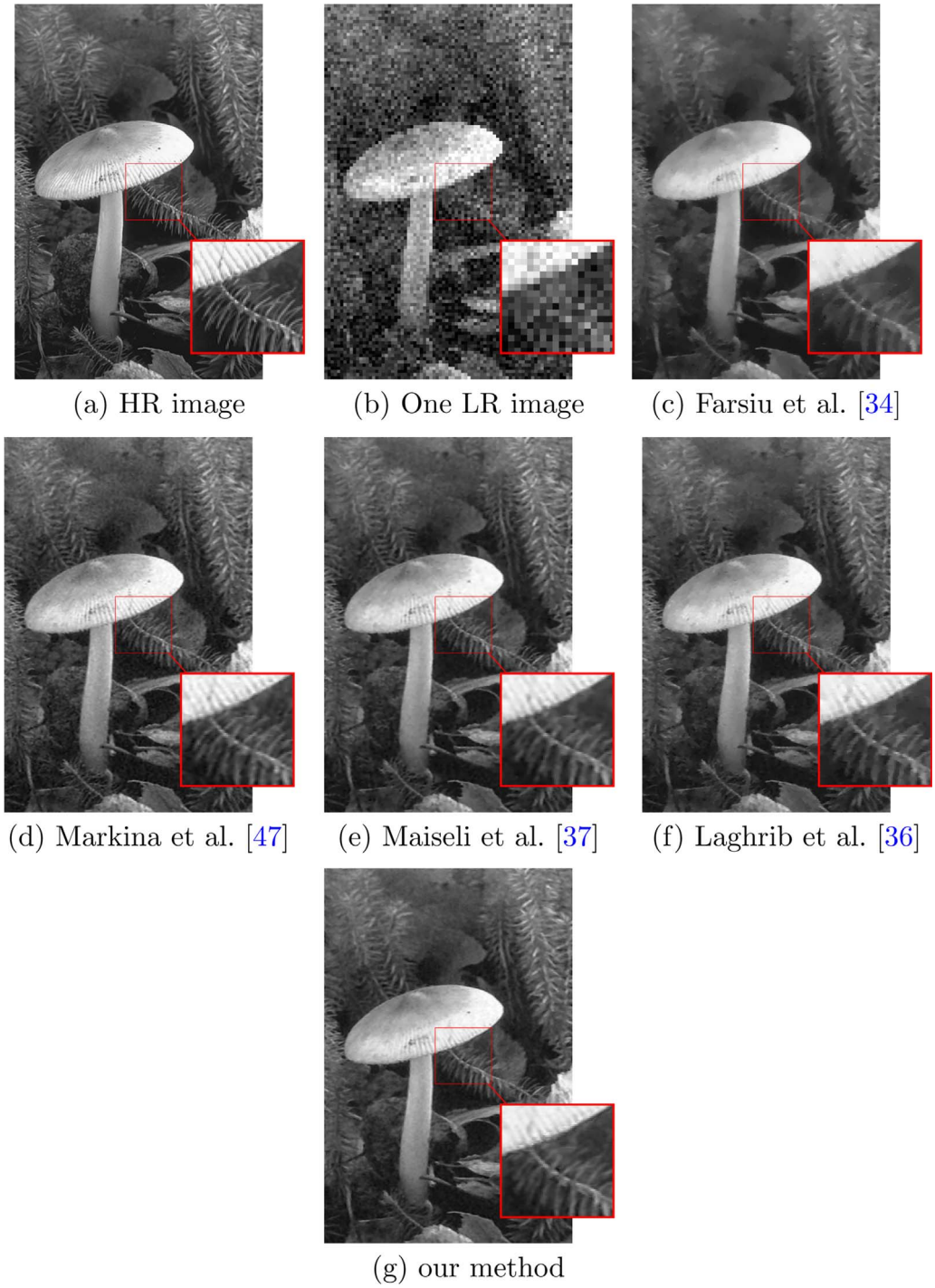


Fig. 8. The results obtained by applying different methods to LR (*Champignon* sequence). (a) HR image. (b) One LR image. (c) Farsiu et al. [34]. (d) Markina et al. [47]. (e) Maiseli et al. [37]. (f) Laghrib et al. [36]. (g) Our method.

preserve as much the texture. A deep research is thus needed to address this problem.

The final experiment is for a real video sequence called “Text” downloaded from the website,[1] which is a challenging example, since it contains a high level of outliers. We use the eight first images of the “Text” video as a LR sequence of size 49×57, and the size of the reconstructed images is 98×114. The reconstructed image by different methods are presented in the Fig. 11. Visually, the obtained results show a significant improvement in our method over the others.

Finally, we can summarise that both simulated and real results show that the new approach produces a sharper image with less blur compared with the other methods.

4. Conclusion

In this paper, we have presented an new multi-frame super-resolution approach that robustly restores image features from a much noised and blurred LR sequence. The new local method is based on a

[1] https://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html

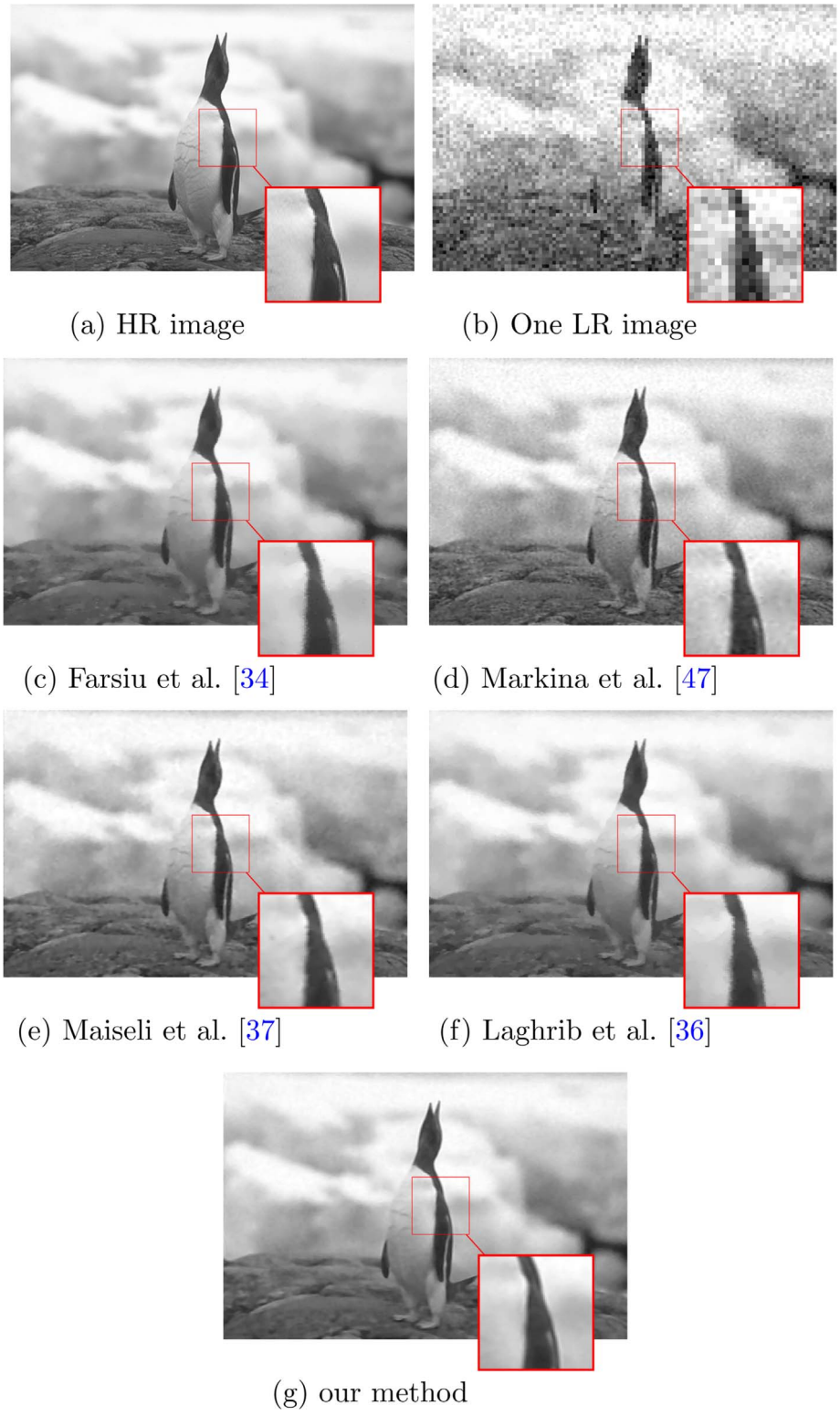


Fig. 9. The results obtained by applying different methods to (*Penguin* sequence). (a) HR image. (b) One LR image. (c) Farsiu et al. [34]. (d) Markina et al. [47]. (e) Maiseli et al. [37]. (f) Laghrib et al. [36]. (g) Our method.

new tensor diffusion regularization with a linear isotropic behaviour in flat regions and a nonlinear filter near sharp edges. To approve this model, the existence and uniqueness of the proposed equation is proved. Simulated and real results demonstrate, visually, the performance of the new method and reveals the robustness with respect to blur and noise reduction compared with competitive methods. Moreover, we have shown that quantitatively the method generates always the best PSNR and SSIM values.

Compliance with ethical standards

- Funding: This research was entirely funded by the respective

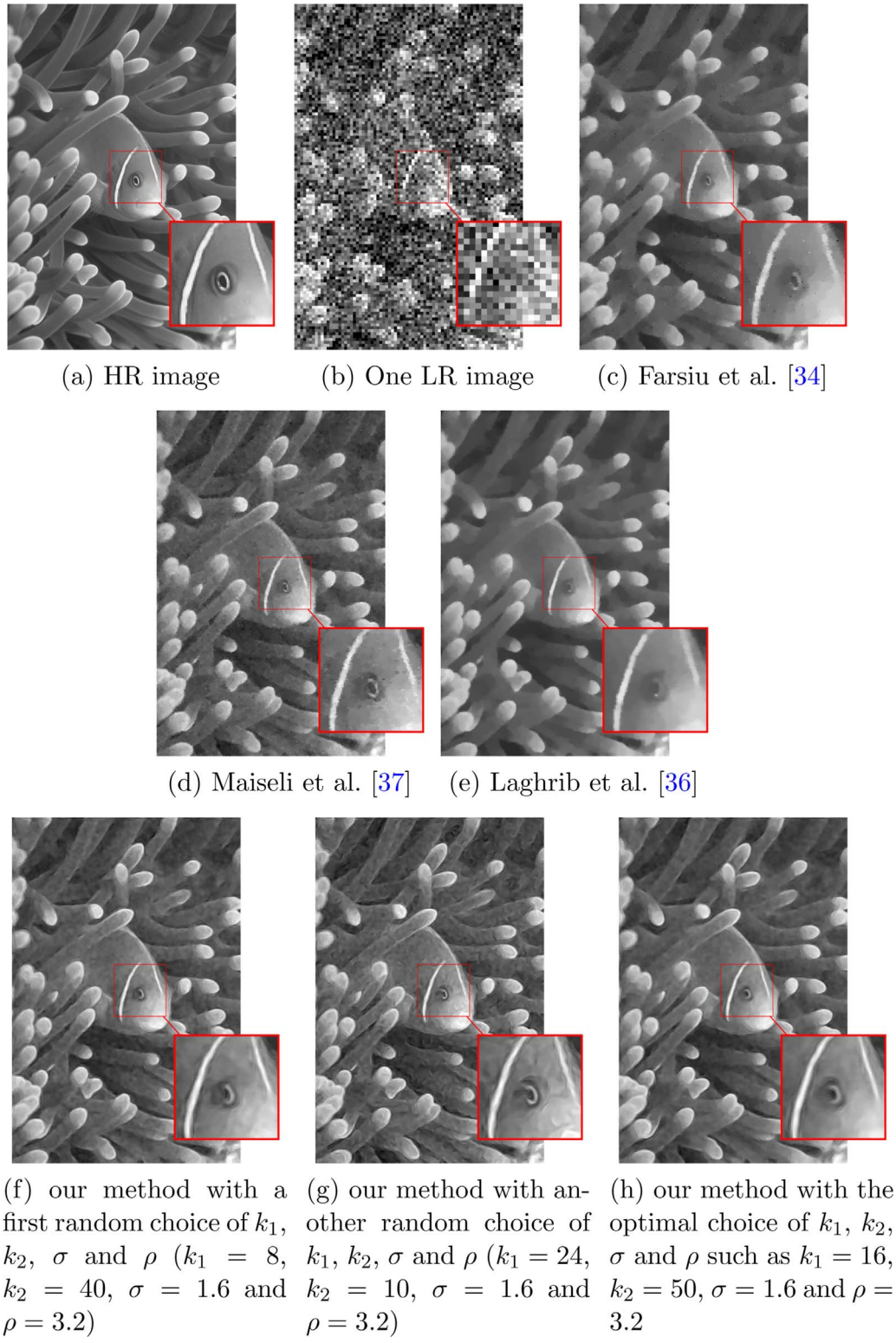


Fig. 10. The results obtained by applying different methods to LR (*Fish* sequence) with $\sigma = 50$ noise. (a) HR image. (b) One LR image. (c) Farsiu et al. [34]. (d) Maiseli et al. [37]. (e) Laghrib et al. [36]. (f) Our method with a first random choice of k_1, k_2, σ and ρ ($k_1 = 8$, $k_2 = 40$, $\sigma = 1.6$ and $\rho = 3.2$). (g) Our method with another random choice of k_1, k_2, σ and ρ ($k_1 = 24$, $k_2 = 10$, $\sigma = 1.6$ and $\rho = 3.2$). (h) Our method with the optimal choice of k_1, k_2, σ and ρ such as $k_1 = 16$, $k_2 = 50$, $\sigma = 1.6$ and $\rho = 3.2$.

institutions of the authors.

- Conflict of interest: The authors declare that they have no conflict of interest.
- Neither human participants nor animals are involved in this research.

Acknowledgments

We are grateful to the anonymous referee for the corrections and useful suggestions that have improved this paper.

Table 1
The PSNR table.

Image	PSNR				
	Farsiu et al. [34]	Markina and Osher [47]	Maiseli et al. [37]	Laghrib et. al [36]	Our method
Castle	25.70	27.02	27.13	26.56	**27.28**
EIAcen	20.63	20.55	21.12	20.87	**22.11**
Champignon	27.87	29.24	29.27	29.06	**29.90**
Penguin	32.04	31.57	32.37	32.18	**33.05**
Lena	29.08	29.15	29.84	29.11	**30.22**
Bird	33.96	33.92	34.60	34.84	**35.08**
Lake	30.64	30.88	31.04	31.73	**31.90**
Baboon	27.22	27.49	27.48	27.59	**27.98**
Peppers	29.94	30.66	**31.45**	31.18	31.06
Goldhill	29.40	29.34	30.11	29.93	**30.38**
Average	28.64	28.98	29.44	29.30	**29.99**

Table 2
The SSIM table.

Image	SSIM				
	Farsiu et al. [34]	Markina and Osher [47]	Maiseli et al. [37]	Laghrib et. al [36]	Our method
Castle	0.776	0.763	0.794	**0.806**	0.803
EIAcen	0.865	0.837	0.877	0.745	**0.880**
Champignon	0.738	0.805	0.815	0.802	**0.830**
Penguin	0.876	0.806	0.880	0.879	**0.884**
Lena	0.822	0.831	0.836	0.844	**0.852**
Bird	0.864	0.852	0.871	0.873	**0.892**
Lake	0.803	0.810	0.818	0.822	**0.828**
Baboon	0.721	0.728	0.736	0.745	**0.750**
Peppers	0.752	0.749	0.769	0.763	**0.774**
Goldhill	0.830	0.821	0.865	0.869	**0.878**
Average	0.804	0.800	0.826	0.814	**0.847**

Table 3
CPU times (in seconds) of different super-resolution methods and the proposed method when the magnification factor is 4.

Image	CPU				
	Farsiu et al. [34]	Markina and Osher [47]	Maiseli et al. [37]	Laghrib et. al [36]	Our method
Castle	7.24	7.26	8.02	7.22	7.96
EIAcen	6.22	6.64	7.62	6.34	7.61
Champignon	7.31	7.48	8.16	7.62	8.08
Penguin	7.66	7.75	8.84	7.92	8.93
Lena	9.05	9.18	10.86	9.22	10.80
Bird	8.22	8.19	9.67	8.26	9.54
Lake	8.16	8.14	10.32	8.22	10.24
Baboon	7.87	7.96	8.66	7.88	8.60
Peppers	8.02	8.16	9.22	8.36	9.76
Goldhill	8.12	8.59	9.98	8.41	10.02
Average time	7.78	7.93	9.13	7.94	9.15

Table 4
The PSNR table of the reconstructed Fish and Barbara images using different σ noise values.

Image	Method	$\sigma = 10$	$\sigma = 30$	$\sigma = 50$
Fish	Farsiu et al. [34]	34.57	30.88	28.47
	Maiseli et al. [37]	34.79	30.98	28.74
	Laghrib et al. [36]	34.12	30.87	29.12
	Our method	**34.95**	**31.70**	**30.47**
Barbara	Farsiu et al. [34]	28.16	25.08	23.66
	Maiseli et al. [37]	**29.13**	**25.86**	24.03
	Laghrib et al. [36]	28.38	25.16	23.91
	Our method	28.94	25.84	**24.54**

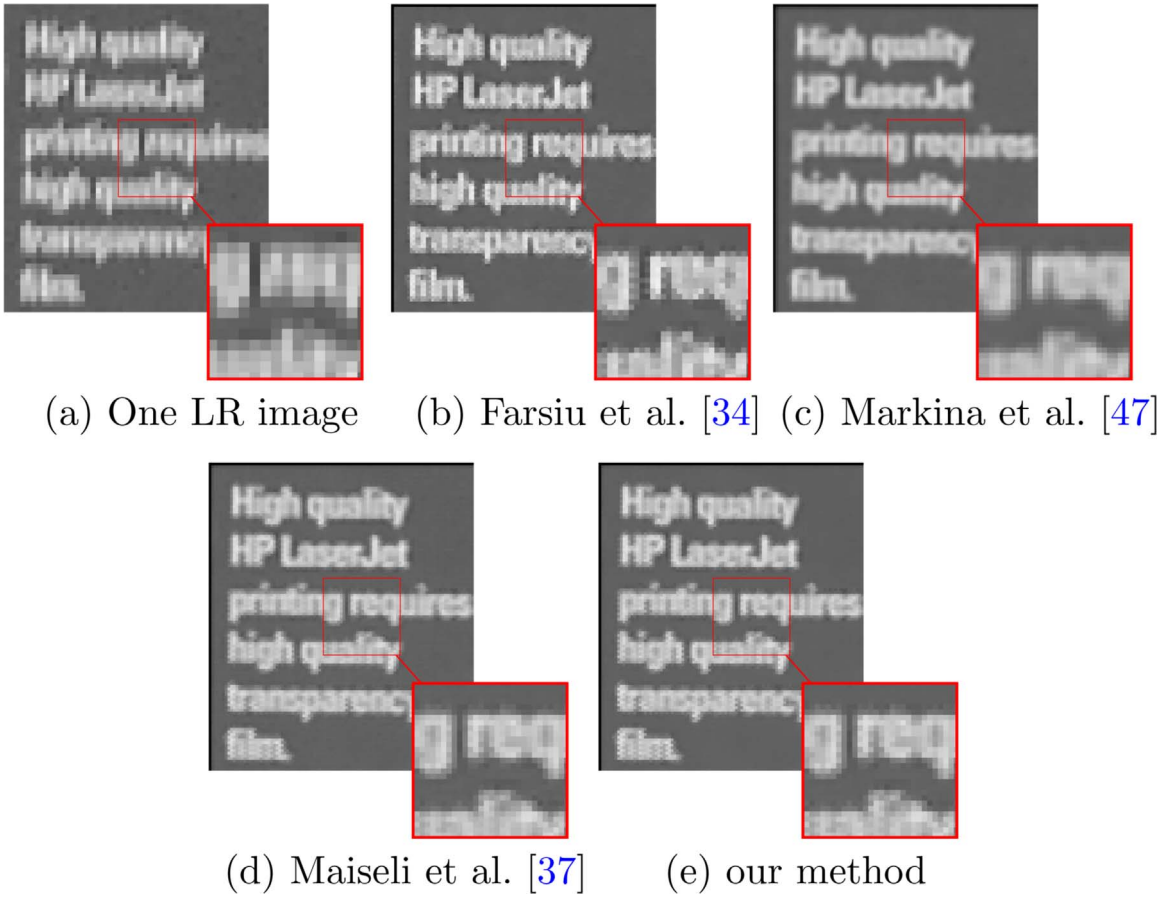


Fig. 11. The results obtained by applying different methods to LR (*Text* sequence). (a) One LR image. (b) Farsiu et al. [34]. (c) Markina et al. [47]. (c) Maiseli et al. [37]. (d) Our method.

Proof of Theorem 2.1

1. *Uniqueness*:

We follow the same technique used in [38], let X_1 and X_2 be two solutions of the problem (6), for every $t \in [0, T]$ we have

$$\frac{\partial X_1}{\partial t}(t, x) - div(D(J_\rho(\nabla X_{1\sigma}))\nabla X_1) - \frac{1}{m}\sum_{k=1}^{m}(WF_kH)^\top(WF_kHX_1 - X_0) = 0 \tag{11}$$

$$\frac{\partial X_2}{\partial t}(t, x) - div(D(J_\rho(\nabla X_{2\sigma}))\nabla X_2) - \frac{1}{m}\sum_{k=1}^{m}(WF_kH)^\top(WF_kHX_2 - X_0) = 0 \tag{12}$$

let $w(t) = X_1(t) - X_2(t)$, $D_1(t) = D(J_\rho(\nabla X_{1\sigma})$, $D_2(t) = D(J_\rho(\nabla X_{2\sigma})$ and $N_k = (WF_kH)^\top WF_kH$, using Eqs. (11) and (12), we obtain

$$\frac{\partial w}{\partial t}(t) - div(D_1 \nabla w(t)) - \frac{1}{m}\sum_{k=1}^{m} N_k w(t) = div((D_1 - D_2)\nabla X_2). \tag{13}$$

Then, multiplying the above inequality by $w(t)$, integrating over Ω, and using the Green formula [50]

$$\frac{1}{2}\frac{\partial}{\partial t}\| w \|^2_{\mathbf{L}^2(\Omega)} + \int_\Omega \langle D_1(t)\nabla w(t),\, \nabla w(t)\rangle dx - \frac{1}{m}\sum_{k=1}^{m}\int_\Omega \langle N_k w(t).\, w(t)\rangle dx = \int_\Omega \langle (D_2(t) - D_1(t))\nabla X_2,\, \nabla w(t)\rangle dx. \tag{14}$$

Since the matrices W, F_k and H are symmetric positive-definite, there exists a positive constant η_k such as

$$\langle N_k w.\, w\rangle \le \eta_k |w|^2. \tag{15}$$

On the other hand, there exists a positive constant $\nu = \nu(\sigma, \| u_0 \|_{\mathbf{L}^\infty(\Omega)})$ (see [51] for more details), such that

$$\nu|y|^2 \le \langle D_1(t)y,\, y\rangle \quad \forall\, y. \tag{16}$$

From (15) and (16), Eq. (14) becomes

$$\frac{1}{2}\frac{\partial}{\partial t}\| w \|^2_{\mathbf{L}^2(\Omega)} + \nu \|\nabla w(t)\|^2_{\mathbf{L}^2(\Omega)} \le \frac{1}{m}\sum_{k=1}^{m}\int_\Omega \langle N_k w(t).\, w(t)\rangle dx + \| D_1(t) - D_2(t)\|_{\mathbf{L}^\infty(\Omega)} \|\nabla X_2(t)\|_{\mathbf{L}^2(\Omega)} \|\nabla w(t)\|_{\mathbf{L}^2(\Omega)}. \tag{17}$$

Moreover, since the matrices $D_1(t)$ and $D_2(t)$ are smooth enough, we have

$$\| D_1(t) - D_2(t)\|_{\mathbf{L}^\infty(\Omega)} \le c \| w(t)\|_{\mathbf{L}^2(\Omega)}. \tag{18}$$

Using the Young's inequality [52] in (17), we have

$$\frac{\partial}{\partial t}\| w \|^2_{\mathbf{L}^2(\Omega)} + 2\nu \|\nabla w(t)\|^2_{\mathbf{L}^2(\Omega)} \le \frac{2}{m}\sum_{k=1}^{m}\eta_k \| w(t)\|_{\mathbf{L}^2(\Omega)} + \frac{4c^2}{\nu}\| w(t)\|^2_{\mathbf{L}^2(\Omega)} \|\nabla X_2(t)\|^2_{\mathbf{L}^2(\Omega)} + \nu \|\nabla w(t)\|^2_{\mathbf{L}^2(\Omega)}, \tag{19}$$

which is equivalent to

$$\frac{\partial}{\partial t}\| w \|^2_{\mathbf{L}^2(\Omega)} \le \left(\frac{2}{m}\sum_{k=1}^{m}\eta_k + \frac{4c^2}{\nu}\|\nabla X_2(t)\|^2_{\mathbf{L}^2(\Omega)}\right)\| w(t)\|^2_{\mathbf{L}^2(\Omega)}. \tag{20}$$

Using Gronwall's inequality [50], we deduce finally that

$$\| w \|^2_{\mathbf{L}^2(\Omega)} \le 0,$$

then $X_1 = X_2$. Let us prove the existence of this solution.

1. *Existence*:

To prove the existence of a weak solution to the problem (6), we use the fixed point theorem of Schauder [53]. Firstly, we introduce the following functional space:

$$\mathbf{H}(0, T) = \left\{ w \in \mathbf{L}^2((0, T); \mathbf{H}^1(\Omega));\, \frac{\partial w}{\partial t} \in \mathbf{L}^2((0, T); \mathbf{H}^1(\Omega)')\right\},$$

where $\mathbf{H}^1(\Omega)'$ is the dual of the Sobolev space $\mathbf{H}^1(\Omega)$. $\mathbf{H}(0, T)$ is a Hilbert space equipped with the norm

$$\| w \|_{\mathbf{H}(0,T)} = \| w \|_{\mathbf{L}^2((0,T);\mathbf{H}^1(\Omega))} + \| \frac{\partial w}{\partial t} \|_{\mathbf{L}^2((0,T);\mathbf{H}^1(\Omega)')}.$$

Let's $w \in \mathbf{H}(0, T) \cap \mathbf{L}^\infty((0, T); \mathbf{L}^2(\Omega))$, such that

$$\| w \|_{\mathbf{L}^\infty((0,T);\mathbf{L}^2(\Omega))} \le \| X_0 \|_{\mathbf{L}^2(\Omega)},$$

and let's define the variational problem associated to (6) with a fixed w

$$\left\langle \frac{\partial X}{\partial t}, v\right\rangle_{\mathbf{H}^1(\Omega)', \mathbf{H}^1(\Omega)} + \int_\Omega \langle D(J_\rho(\nabla w_\sigma))\nabla X,\, \nabla Y\rangle dx = \int_\Omega \frac{1}{m}\sum_{i=1}^{m} W^\top F_k^\top H^\top (WF_k Hw - X_0) Y dx, \quad \forall\, Y \in \mathbf{H}^1(\Omega),\ \ a.\, e \text{ in } [0, \mathrm{T}]. \tag{21}$$

Therefore, based on the existence of parabolic equations results [45], we can prove that the problem (21) admits a unique solution X_w in $\mathbf{H}(0, T)$, satisfying the following estimations:

$$\begin{cases} \| X_w \|_{\mathbf{L}^2((0,T);\mathbf{H}^1(\Omega))} \leq c_1, \\ \| X_w \|_{\mathbf{L}^\infty((0,T);\mathbf{L}^2(\Omega))} \leq \| X_0 \|_{\mathbf{L}^2(\Omega)}, \\ \| \dfrac{\partial X_w}{\partial t} \|_{\mathbf{L}^2((0,T);\mathbf{H}^1(\Omega)')} \leq c_2, \end{cases}$$

where the constants c_1 and c_2 depends on ρ, σ, W, F, H and X_0. These estimations lead to define the subset $\mathbf{H}_0$ of $\mathbf{H}(0, T)$ as

$$\mathbf{H}_0 = \{w \in \mathbf{H}(0, T),\ w(0) = X_0;\ \| w \|_{\mathbf{L}^2((0,T);\mathbf{H}^1(\Omega))} \leq c_1,\ \| w \|_{\mathbf{L}^\infty((0,T);\mathbf{L}^2(\Omega))} \leq \| X_0 \|_{\mathbf{L}^2(\Omega)},\ \| \frac{\partial w}{\partial t} \|_{\mathbf{L}^2((0,T);\mathbf{H}^1(\Omega)')} \leq c_2\} \tag{22}$$

Let U: $w \to U(w) = X_w$ be a mapping from $\mathbf{H}_0$ to $\mathbf{H}_0$. It's clear that $\mathbf{H}_0$ is a non empty, convex, and weakly compact subspace of $\mathbf{H}(0, T)$ [52]. In order to use the classical Schauder's fixed-point theorem, we need first to prove that the mapping $w \to U(w)$ is weakly continuous. Indeed, let $(w_n)_n$ be a sequence that converges weakly to w in $\mathbf{H}_0$ and $X_n = U(w_n)$. Using the compact inclusions of Sobolev space [50] and by the estimations (22), there exists a subsequence noted also (w_n) and (X_n) such that

$$\frac{\partial w_n}{\partial t} \rightharpoonup \frac{\partial w}{\partial t} \quad in \quad \mathbf{L}^2((0, T); \mathbf{H}^1(\Omega)'),\ X_n \to X \quad in \quad \mathbf{L}^2((0, T); \mathbf{L}^2(\Omega)),\ \nabla X_n \rightharpoonup \nabla X \quad in \quad (\mathbf{L}^2((0, T); \mathbf{L}^2(\Omega)))^2,\ w_n \to w \quad in \quad \mathbf{L}^2((0, T); \mathbf{L}^2(\Omega)),\ D(J_\rho(\nabla w_{n\sigma})) \to D(J_\rho(\nabla w_\sigma)) \quad in \quad \mathbf{L}^2((0, T); \mathbf{L}^2(\Omega)),\ X_n(0) \to X(0) \quad in \quad \mathbf{H}^1(\Omega)'.$$

Using the limit as $n \to \infty$, we obtain $X = X_w = U(w)$. Moreover, by the uniqueness of the solution of (6), the sequence $X_n = U(w_n)$ converge weakly to $X = U(w)$, this prove the continuity of U. Finally, from the Schauder's fixed-point theorem, there exists a function $w \in \mathbf{H}_0$ such that $w = S(w) = X_w$. Moreover, since $X \in \mathbf{L}^2((0, T); \mathbf{H}^1(\Omega))$ and $\frac{\partial X}{\partial t} \in \mathbf{L}^2((0, T); \mathbf{H}^1(\Omega)')$, by the Aubin's theorem [54], we can deduce that $X \in C((0, T); \mathbf{L}^2(\Omega))$.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.sigpro.2016.09.014.

References

[1] P. Milanfar, Super-Resolution Imaging, Digital Imaging and Computer Vision, CRC Press, New York, 2010.
[2] Yue Linwei, et al., A locally adaptive l1–l2 norm for multi-frame super-resolution of images with mixed noise and outliers, Signal Process 105 (2014) 156–174.
[3] D. Capel, A. Zisserman, Computer vision applied to super-resolution, IEEE Signal Process. Mag. 20 (2003) 75–86.
[4] M. Protter, M. Elad, H. Takeda, P. Milanfar, Generalizing the nonlocal-means to super-resolution reconstruction, IEEE Trans. Image Process. 18 (1) (2009) 36–51.
[5] R.M. Bahy, G.I. Salama, T.A. Mahmoud, Adaptive regularization-based super resolution reconstruction technique for multi-focus low-resolution images, Signal Process. 103 (2014) 155–167.
[6] X. Li, Y. Hu, X. Gao, D. Tao, B. Ning, A multi-frame image super-resolution method, Signal Process. 90 (2) (2010) 405–414.
[7] A. Laghrib, A. Hakim, S. Raghay, M.E.L. Rhabi, Robust super resolution of images with non-parametric deformations using an elastic registration, Appl. Math. Sci. 8 (179) (2014) 8897–8907.
[8] E. Sardis, A. Voulodimos, V. Anagnostopoulos, C. Lalos, A. Doulamis, D. Kosmopoulos, An industrial video surveillance system for quality assurance of a manufactory assembly, in: Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments, Samos, ACM, 2010, p. 66.
[9] D.I. Kosmopoulos, N.D. Doulamis, A.S. Voulodimos, Bayesian filter based behavior recognition in workflows allowing for user feedback, Comput. Vis. Image Underst. 116 (3) (2012) 422–434.
[10] K. Makantasis, K. Karantzalos, A. Doulamis, N. Doulamis, Deep supervised learning for hyperspectral data classification through convolutional neural networks, in: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, Milan, 2015, pp. 4959–4962.
[11] A.J. Tatem, H.G. Lewis, P.M. Atkinson, M.S. Nixon, Super-resolution land cover pattern prediction using a Hopfield neural network, Remote Sens. Environ. 79 (1) (2002) 1–14.
[12] L. Zhang, H. Zhang, H. Shen, P. Li, A super-resolution reconstruction algorithm for surveillance images, Signal Process. 90 (3) (2010) 848–859.
[13] F.C. Lin, C.B. Fookes, V. Chandran, S. Sridharan, Investigation into optical flow super-resolution for surveillance applications, The University of Queensland, 2005, pp. 73–78.
[14] F.Stanco, S. Battiato, G. Gallo, Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks, New York, CRC Press, 2011.
[15] A. Doulamis, N. Doulamis, C. Ioannidis, C. Chrysouli, N. Grammalidis, K. Dimitropoulos, C. Potsiou, E.K. Stathopoulou, M. Ioannides, 5d modelling: an efficient approach for creating spatiotemporal predictive 3d maps of large-scale cultural resources, ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. 2 (5) (2015) 61.
[16] J. Yang, T. Huang, Image super-resolution: historical overview and future challenges, Super-Resolution Imag. (2010) 20–34.
[17] A.J. Tatem, H.G. Lewis, P.M. Atkinson, M.S. Nixon, Super-resolution target identification from remotely sensed images using a hopfield neural network, IEEE Trans. Geosci. Rem. Sens. 39 (4) (2001) 781–796.
[18] S.C. Park, M.K. Park, M.G. Kang, Super-resolution image reconstruction: a technical overview, IEEE Signal Process. Mag. 20 (3) (2003) 21–36.
[19] R.Y. Tsai, T.S. Huang, Multiframe image restoration and registration, In: T.S. Huang (Editor). Advances in Computer Vision and Image Processing, JAI Press, Greenwich, CT.
[20] S. Borman, R.L. Stevenson, Super-resolution from image sequences-a review, in: MWSCAS, IEEE, Notre Dame, Indiana, 1998, p. 374.
[21] M. Elad, Y. Hel-Or, A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur, IEEE Trans. Image Process. 10 (8) (2001) 1187–1193.
[22] N. Nguyen, P. Milanfar, G. Golub, A computationally efficient superresolution image reconstruction algorithm, IEEE Trans. Image Process. 10 (4) (2001) 573–583.
[23] S. Villena, M. Vega, R. Molina, A.K. Katsaggelos, Bayesian super-resolution image reconstruction using an l1 prior, in: Proceedings of 6th International Symposium on Image and Signal Processing and Analysis, ISPA 2009, IEEE, Salzburg, Austria, 2009, pp. 152–157.
[24] S.D. Babacan, R. Molina, A.K. Katsaggelos, Variational Bayesian super resolution, IEEE Trans. Image Process. 20 (4) (2011) 984–999.
[25] O.A. Omer, T. Tanaka, Region-based weighted-norm with adaptive regularization for resolution enhancement, Digital Signal Process. 21 (4) (2011) 508–516.
[26] S. Zhao, H. Liang, M. Sarem, A generalized detail-preserving super-resolution method, Signal Process. 120 (2016) 156–173.
[27] S.D. Babacan, R. Molina, A.K. Katsaggelos, Parameter estimation in tv image restoration using variational distribution approximation, IEEE Trans. Image Process. 17 (3) (2008) 326–339.
[28] A. Panagiotopoulou, V. Anastassopoulos, Regularized super-resolution image reconstruction employing robust error norms, Opt. Eng. 48 (11) (2009) 117004.
[29] V.Patanavijit, S. Jitapunkul, A robust iterative multiframe super-resolution reconstruction using a Huber Bayesian approach with Huber–Tikhonov regularization, in: International Symposium on Intelligent Signal Processing and Communications, 2006. ISPACS'06. IEEE, Tottori, Japan, 2006, pp. 13–16.
[30] N.A. El-Yamany, P.E. Papamichalis, Robust color image superresolution: an adaptive m-estimation framework, J. Image Video Process. 2008 (2008) 16.
[31] T.Q. Pham, L.V. Vliet, K. Schutte, Robust super-resolution by minimizing a gaussian-weighted l2 error norm, J. Phys.: Conf. Ser. 124 (2008) 012037 (IOP Publishing).

[32] S. Tourbier, X. Bresson, P. Hagmann, J.-P. Thiran, R. Meuli, M.B. Cuadra, An efficient total variation algorithm for super-resolution in fetal brain mri with adaptive regularization, NeuroImage 118 (2015) 584–597.
[33] V. Patanavijit, S. Jitapunkul, A Lorentzian stochastic estimation for a robust iterative multiframe super-resolution reconstruction with Lorentzian–Tikhonov Regularization, EURASIP J. Adv. Signal Process. 2007 (1) (2007) 1–21.
[34] S. Farsiu, D. Robinson, M. Elad, P. Milanfar, Advances and challenges in super-resolution, Int. J. Imag. Syst. Technol. 14 (2) (2004) 47–57.
[35] X. Zeng, L. Yang, A robust multiframe super-resolution algorithm based on half-quadratic estimation with modified btv regularization, Digital Signal Process. 23 (1) (2013) 98–109.
[36] A. Laghrib, A. Hakim, S. Raghay, A combined total variation and bilateral filter approach for image robust super resolution, EURASIP J. Image Video Process. 2015 (1) (2015) 1–10.
[37] B.J. Maiseli, N. Ally, H. Gao, A noise-suppressing and edge-preserving multiframe super-resolution image reconstruction method, Signal Process.: Image Commun. 34 (2015) 1–13.
[38] J. Weickert, Anisotropic diffusion in image processing, vol. 1, Teubner, Stuttgart, 1998.
[39] J. Weickert, B.T.H. Romeny, M. Viergever, et al., Efficient and reliable schemes for nonlinear diffusion filtering, IEEE Trans. Image Process. 7 (3) (1998) 398–410.
[40] J. Weickert, C. Schnörr, A theoretical framework for convex regularizers in pde-based computation of image motion, Int. J. Comput. Vis. 45 (3) (2001) 245–264.
[41] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Trans. Pattern Anal. Mach. Intell. 12 (7) (1990) 629–639.
[42] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D: Nonlinear Phenom. 60 (1) (1992) 259–268.
[43] Y.-L. You, W. Xu, A. Tannenbaum, M. Kaveh, Behavioral analysis of anisotropic diffusion in image processing, IEEE Trans. Image Process. 5 (11) (1996) 1539–1553.
[44] J. Weickert, Coherence-enhancing diffusion of colour images, Image Vis. Comput. 17 (3) (1999) 201–212.
[45] G. Aubert, P. Kornprobst, Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations, vol. 147, Springer Science & Business Media, New York, 2006.
[46] M. Werlberger, W. Trobin, T. Pock, A. Wedel, D. Cremers, H. Bischof, Anisotropic Huber-l1 optical flow, in: BMVC, vol. 1, 2009, p.3.
[47] A. Marquina, S.J. Osher, Image super-resolution by tv-regularization and Bregman iteration, J. Sci. Comput. 37 (3) (2008) 367–382.
[48] Z. Wang, A.C. Bovik, Mean squared error: love it or leave it? a new look at signal fidelity measures, IEEE Signal Process. Mag. 26 (1) (2009) 98–117.
[49] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[50] H. Attouch, G. Buttazzo, G. Michaille, Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization, vol. 17, SIAM, New York, 2014.
[51] J. Weickert, Scale-space properties of nonlinear diffusion filtering with a diffusion tensor, Universität Kaiserslautern. Fachbereich Mathematik. Arbeitsgruppe Technomathematik (1994).
[52] H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer Science & Business Media, New York, 2010.
[53] J.T. Schwartz, Nonlinear Functional Analysis, CRC Press, New York, 1969.
[54] J.-P. Aubin, Un théoreme de compacité, C.R. Acad. Sci. Paris 256 (24) (1963) 5042–5044.

Signal Processing 131 (2017) 502–513

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Blind noisy mixture separation for independent/dependent sources through a regularized criterion on copulas

A. Ghazdali [a], M. El Rhabi [b], H. Fenniri [c], A. Hakim [a], A. Keziou [d,*]

[a] LAMAI, FSTG, Université Cadi-Ayyad, Marrakech, Maroc
[b] Ecole des Ponts ParisTech (ENPC), France
[c] CReSTIC, Université de Reims Champage-Ardenne, France
[d] LMR EA 4535 and ARC-Mathématiques CNRS 3399, Université de Reims Champagne-Ardenne, France

ARTICLE INFO

Article history:
Received 18 September 2015
Received in revised form
30 July 2016
Accepted 7 September 2016
Available online 13 September 2016

Keywords:
Blind source separation
Noisy instantaneous mixtures
Copulas
Total variation
Mutual information
Kullbak–Leibler divergence between copulas

ABSTRACT

The paper introduces a new method for Blind Source Separation (BSS) in noisy instantaneous mixtures of both independent or dependent source component signals. This approach is based on the minimization of a regularized criterion. Precisely, it consists in combining the total variation method for denoising with the Kullback–Leibler divergence between copula densities. The latter takes advantage of the copula to model the structure of the dependence between signal components. The obtained algorithm achieves separation in a noisy context where standard BSS methods fail. The efficiency and robustness of the proposed approach are illustrated by numerical simulations.

1. Introduction

The blind source separation (BSS) problem is a fundamental issue in applications of many different fields such as signal and image processing, biological and medical data analysis, communications, etc. The BSS aims to recover unknown source signals from a set of observations which are unknown mixtures of source signals. In order to separate the mixtures, different assumptions on the sources have to be made. In the literature, the most common assumptions are statistical independence of the source components and the condition that at most one of the components is Gaussian. Under these assumptions, the BSS problem is linked to the well known problem of Independent Component Analysis (ICA), see for instance [1]. Using the above conditions, particularly, for noise-free mixtures, many methods of BSS have been proposed in the literature, see e.g. [2–5]. Some of these procedures use second or higher order statistics [6,7], maximize likelihood [8], maximize nongaussianity [9], minimize the mutual information [10] or φ-divergences [11]. An interesting overview on the subject can be found in [12]. Recently, under the assumption that the source components are independent, based on estimation of mutual information and total variation (TV) regularization, Ref. [13] provided a new BSS algorithm to separate noisy mixtures of instantaneous or covolutive mixtures of independent source components. Further, it has been shown in [14] that, based on copula models, without the assumption of independence of the source components, we can still identify both mixing matrix and sources uniquely (up to scale and permutation indeterminacies) of (noise-free) mixtures of both independent and dependent source components. Motivated by various applications, we investigate, in the present paper, models of noisy linear instantaneous mixtures of independent/dependent sources, for which we propose, based on the previews works [13,14], a new BSS procedure. We successfully combined the two methods of [13,14] to obtain a new BSS algorithm able to separate noisy mixtures of independent/dependent source components. Our methodology consists in denoising the observed signals through the minimization of their total variation, and then minimizing Kullback–Leibler divergence, between copula densities, penalized by the total variation of the estimated source signals. The outline of this paper is organized as follows. In Section 2, we state the BSS problem in a noisy mixture context. In Section 3, we give a brief introduction on copulas and some of their main properties. In Section 4, we describe our approach. Section 5 presents how to implement the proposed approach using both numerical and statistical techniques. Section 6 gives some

* Corresponding author.
E-mail addresses: a.ghazdali@gmail.com (A. Ghazdali), elrhabi@gmail.com (M. El Rhabi), hassan.fenniri@univ-reims.fr (H. Fenniri), abdelilah.hakim@gmail.com (A. Hakim), amor.keziou@univ-reims.fr (A. Keziou).

http://dx.doi.org/10.1016/j.sigpro.2016.09.006
0165-1684/

numerical results, illustrating the efficiency and robustness of the proposed method. Conclusions are drawn in Section 7.

2. The BSS model

2.1. BSS model in the noise-free case

BSS problem can be modeled as follows. Denoting $\boldsymbol{A}[\cdot]$ the (unknown) mixing operator, the relationship between the observed and source signals can be written as

$$\boldsymbol{x}(t) := \boldsymbol{A}[\boldsymbol{s}(t)] \in \mathbb{R}^p, \tag{1}$$

where $\boldsymbol{s}(t) \in \mathbb{R}^q$ is the unknown vector of source signals to be estimated, and $\boldsymbol{x}(t)$ represents the observed signal vector at time $t \in [0, T]$. The goal of BSS is therefore to estimate the unknown sources $\boldsymbol{s}(t)$ from the observed mixtures $\boldsymbol{x}(t)$. The estimation is performed with no prior information about either the sources or the mixing operator $\boldsymbol{A}[\cdot]$. Specific restrictions are made on both the mixing model and the source signals in order to limit the generality. We will restrict ourselves to the case where the number of source components and the number of observed mixture ones are equal ($p=q$), and we assume, in the present paper, that the mixtures are linear and instantaneous, so that the mixing operator $\boldsymbol{A}$ can be considered as a $p \times p$ matrix. In this case, supposing in addition that $\boldsymbol{A}$ is invertible, the candidate estimates of the sources will be obviously of the form

$$\boldsymbol{y}(t) := \boldsymbol{B}\,\boldsymbol{x}(t) \in \mathbb{R}^p, \tag{2}$$

where $\boldsymbol{B} \in \mathbb{R}^{p\times p}$ represents an appropriate demixing matrix. In other words, the problem is to obtain an estimate, denote it $\hat{\boldsymbol{B}}$, "closing" as much as possible to the ideal solution $\boldsymbol{B} = \boldsymbol{A}^{-1}$, by the use of only the observed vector signal $\boldsymbol{x}(t)$, leading to accurate estimate $\hat{\boldsymbol{s}}(t)$ of the source vector signal $\boldsymbol{s}(t)$,

$$\hat{\boldsymbol{s}}(t) := \hat{\boldsymbol{B}}\,\boldsymbol{x}(t). \tag{3}$$

2.2. BSS model for noisy mixtures

In the present paper, we will focus on the BSS problem for noisy linear instantaneous mixtures

$$\overline{\boldsymbol{x}}(t) := \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t) \in \mathbb{R}^p, \quad t \in [0, T], \tag{4}$$

where $\overline{\boldsymbol{x}}(t) \in \mathbb{R}^p$ is the vector of noisy observations, $\boldsymbol{s}(t) \in \mathbb{R}^p$ is the unknown vector of sources to be estimated, $\boldsymbol{n}(t) \in \mathbb{R}^p$ is the (unknown) noise vector, and $\boldsymbol{A} \in \mathbb{R}^{p\times p}$ is the unknown mixing $p \times p$-matrix. We assume that the source signal $\boldsymbol{s}(t)$ and the noise $\boldsymbol{n}(t)$ are statistically independent. The mixing model (4) can also be written as

$$\overline{\boldsymbol{x}}(t) = \boldsymbol{x}(t) + \boldsymbol{n}(t), \tag{5}$$

where

$$\boldsymbol{x}(t) := \boldsymbol{A}\,\boldsymbol{s}(t) \tag{6}$$

is the inaccessible noise-free mixed vector signals. The aim here is to estimate the sources $\boldsymbol{s}(t)$ using only the observed noisy signal $\overline{\boldsymbol{x}}(t)$. The presence of noise $\boldsymbol{n}(t)$ within the mixing model (4), as well as the possible dependency structure of the source components, complicate significantly the BSS problem. Unlike the above case of noise-free mixtures, see Section 2.1, the separating system of the form

$$\overline{\boldsymbol{y}}(t) := \boldsymbol{B}\overline{\boldsymbol{x}}(t) \tag{7}$$

does not provide a valid estimation for the source signals $\boldsymbol{s}(\cdot)$. In fact, the obtained signals according to the above system, can be written as

$$\overline{\boldsymbol{y}}(t) := \boldsymbol{B}\overline{\boldsymbol{x}}(t) = \boldsymbol{B}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{n}(t) =: \boldsymbol{y}(t) + \overline{\boldsymbol{n}}(t), \tag{8}$$

where $\overline{\boldsymbol{y}}(t) \in \mathbb{R}^p$ is the "noisy estimate" of $\boldsymbol{s}(t)$. That is the noisy estimated source $\overline{\boldsymbol{y}}(t)$ is the sum of

$$\boldsymbol{y}(t) := \boldsymbol{B}\boldsymbol{x}(t) \tag{9}$$

the "noisy-free" estimate of the sources, which is not feasible, and the unknown "noise"

$$\overline{\boldsymbol{n}}(t) := \boldsymbol{B}\boldsymbol{n}(t). \tag{10}$$

Ideally, we would like to retrieve $\boldsymbol{y}(t)$ by denoising $\overline{\boldsymbol{y}}(t)$, but it is rather difficult since the noise $\overline{\boldsymbol{n}}(t)$ is unknown. During last years, several algorithms have been proposed to tackle the noisy BSS problem; In [15], the authors proposed a two-step approach by combining the fractional lower order statistics for the mixing matrix estimation and minimum entropy criterion for noise-free source components estimation with gradient-based BSS algorithms. In [16], a whitening procedure is proposed to reduce the noise effect. At our knowledge, the noisy BSS problem for possibly dependent sources has not been considered in the literature. In [14], the authors proposed a new criterion to successfully separate noise-free mixtures of both independent and dependent sources, based on estimation of Kullback–Leibler divergence between copula densities. Ref. [13] provides a BSS procedure for noisy mixtures using TV-regularization of mutual information, which applies exclusively under the assumption that the source components are independent. Based on these last two works, the present paper introduces a new procedure for estimating sources in the context of noisy mixtures of possibly dependent source components. It proceeds on two stages: (i) denoising the observed signal $\overline{\boldsymbol{x}}(t)$ before demixing; (ii) a simultaneous BSS-denoisy procedure via minimizing a TV-regularized measure between copula densities.

3. Brief recall on copula, model selection and statistical estimation

When modeling multivariate distributions, one has to take into account the effects of the marginal distributions as well as the dependence between them. This can be achieved by using the copula approach, which allows to deal with the margins and the dependency structure separately. The concept of copulas was introduced in [17] as a function which couples a joint distribution function with its univariate margins. Copula is considered as a multivariate distribution function on the unit cube $[0, 1]^p$, with uniformly distributed marginals. Precisely, it can be presented as follows. Consider any random vector

$$\mathbf{Y} := (Y_1, \ldots, Y_p)^\top \in \mathbb{R}^p, \quad p \geq 1,$$

with joint distribution function (d.f.)

$$\mathbb{F}_{\mathbf{Y}}(\cdot)\colon \mathbf{y} \in \mathbb{R}^p \mapsto \mathbb{F}_{\mathbf{Y}}(\mathbf{y}) := \mathbb{F}_{\mathbf{Y}}(y_1, \ldots, y_p) := \mathbb{P}(Y_1 \leq y_1, \ldots, Y_p \leq y_p),$$

and continuous marginal d.f.'s

$$F_{Y_j}(\cdot)\colon y_j \in \mathbb{R} \mapsto F_{Y_j}(y_j) := \mathbb{P}(Y_j \leq y_j), \quad \forall\, j = 1, \ldots, p.$$

The characterization theorem of Sklar [17] shows that there exists a unique p-variate function $\mathbb{C}_{\mathbf{Y}}(\cdot)\colon [0, 1]^p \mapsto [0, 1]$, such that,

$$\mathbb{F}_{\mathbf{Y}}(\mathbf{y}) = \mathbb{C}_{\mathbf{Y}}(F_{Y_1}(y_1), \ldots, F_{Y_p}(y_p)), \quad \forall\, \mathbf{y} := (y_1, \ldots, y_p)^\top \in \mathbb{R}^p.$$

The function $\mathbb{C}_{\mathbf{Y}}(\cdot)$ is called a copula, and it is in itself a joint d.f. on $[0, 1]^p$, with uniform margins. We have

$\forall\, \mathbf{u}:=(u_1, \ldots, u_p)^\top \in [0,1]^p,\ \mathbb{C}_{\mathbf{Y}}(\mathbf{u}) = \mathbb{P}\Big(F_{Y_1}(Y_1) \le u_1, \ldots, F_{Y_p}(Y_p) \le u_p\Big).$

Conversely, for any marginal d.f.'s $F_1(\cdot), \ldots, F_p(\cdot)$, and any copula function $\mathbb{C}(\cdot)$, the function $\mathbb{C}(F_1(\cdot), \ldots, F_p(\cdot))$ is a multivariate d.f. on $\mathbb{R}^p$. On the other hand, since the marginal d.f.'s $F_{Y_j}(\cdot)$, $j = 1, \ldots, p$, are assumed to be continuous, the random variables $U_1:=F_{Y_1}(Y_1), \ldots, U_p:=F_{Y_p}(Y_p)$ are uniformly distributed on the interval $[0, 1]$, and the distribution function of the random vector $\boldsymbol{U}:=(U_1, \ldots, U_p)^\top \in [0,1]^p$ is a copula. Moreover, we have that the components $Y_1, \ldots, Y_p$ are statistically independent if and only if (iff) the copula $\mathbb{C}_{\mathbf{Y}}(\cdot)$ of the random vector $\mathbf{Y}:=(Y_1, \ldots, Y_p)^\top \in \mathbb{R}^p$ is of the form

$$\mathbb{C}_{\mathbf{Y}}(\mathbf{u}) = \prod_{j=1}^{p} u_j =: \mathbb{C}_{\Pi}(\mathbf{u}),\ \forall\, \mathbf{u} \in [0,1]^p,$$

which will be denoted $\mathbb{C}_{\Pi}(\cdot)$. It is called the copula of independence. Define, when it exists, the copula density

$$c_{\mathbf{Y}}(\mathbf{u}) := \frac{\partial^p \mathbb{C}_{\mathbf{Y}}(\mathbf{u})}{\partial u_1 \cdots \partial u_p},\quad \forall\, \mathbf{u} \in [0,1]^p.$$

Hence, the copula density of independence, denote it $c_{\Pi}(\cdot)$, is the function taking the value 1 on $[0, 1]^p$ and zero otherwise, namely,

$$c_{\Pi}(\mathbf{u}) := \mathbb{1}_{[0,1]^p}(\mathbf{u}),\quad \forall\, \mathbf{u} \in [0,1]^p. \tag{11}$$

Let $f_{\mathbf{Y}}(\cdot)$, if it exists, be the probability density of the random vector $\mathbf{Y}:=(Y_1, \ldots, Y_p)^\top$, and, respectively, $f_{Y_1}(\cdot), \ldots, f_{Y_p}(\cdot) \in \mathbb{R}^p$, the marginal probability densities of the random variables $Y_1, \ldots, Y_p$. Then, a straightforward computation shows that, for all vector $\mathbf{y}:=(y_1, \ldots, y_p)^\top \in \mathbb{R}^p$, we have the relation

$$f_{\mathbf{Y}}(\mathbf{y}) = \left(\prod_{j=1}^{p} f_{Y_j}(y_j)\right) c_{\mathbf{Y}}\Big(F_{Y_1}(y_1), \ldots, F_{Y_p}(y_p)\Big), \tag{12}$$

In the literature, many models have been proposed for copulas. An important class of those models is that of semiparametric copula models, which has been widely used in survival analysis, where modeling and estimating the dependency structure, between survival variables, are fundamental problems. Models of this class are based on parametric copulas $\mathbb{C}(\cdot, \boldsymbol{\theta})$, indexed by a parameter $\theta \in \Theta \subset \mathbb{R}^d$, with nonparametric margins. We recall, in Table 1, a description of three copula models that we used in Section 6 in our simulation study, namely, the copulas of Clayton [18], Ali–Mikhail–Haq (AMH) [19] and Farlie–Gumbel–Morgenstern (FGM) [20]. We provide for each model the corresponding parameter space $\boldsymbol{\Theta}$, and the particular value θ_0, of the parameter θ, corresponding to the independence hypothesis of margins, i.e., the value θ_0 satisfying

Table 1
Examples of semiparametric copulas: AMH, Clayton and FGM.

	$\mathbb{C}(u_1, u_2, \ldots, u_p, \theta)$	Θ	θ_0
AMH	$\dfrac{\prod_{i=1}^{p} u_i}{1-\theta\left(\prod_{i=1}^{p}(1-u_i)\right)}$	$[-1, 1]$	0
Clayton	$\max\left[\left(\sum_{i=1}^{p} u_i^{-\theta} - p + 1\right), 0\right]^{-\frac{1}{\theta}}$	$[-1, +\infty[\setminus\{0\}$	0
FGM	$\left[\prod_{i=1}^{p} u_i\right]\left[1 + \sum_{i=1}^{p-1}\sum_{j=i+1}^{p}\theta(1-u_i)(1-u_j)\right]$	$[-1, 1]$	0

$$\mathbb{C}\big(u_1, \ldots, u_p; \boldsymbol{\theta}_0\big) = \mathbb{C}_{\Pi}\big(u_1, \ldots, u_p\big) := \prod_{i=1}^{p} u_i,\quad \forall\, \big(u_1, \ldots, u_p\big)^\top \in [0,1]^p. \tag{13}$$

For more details on modeling theory as well as surveys of the commonly semiparametric used copulas, we can refer to [21,22]. In the following lines, we briefly expose some copula model selection procedures as well as method of estimating the parameter θ from the data. Let $\boldsymbol{S}$ be a random vector with values in $\mathbb{R}^p$, and assume that a training sample of S is available, that is, we have at hand i.i.d. realizations $\boldsymbol{s}(1), \ldots, \boldsymbol{s}(N)$ of $\boldsymbol{S}$. The aim is to select, from the data, the "best" copula model, among a list of candidate models, modeling the dependency structure of the components of $\boldsymbol{S}$, and to estimate the parameter θ of the selected model. Let $\Big\{C_1(\cdot, \theta_1);\ \theta_1 \in \Theta_1 \subset \mathbb{R}^{d_1}\Big\}, \ldots, \Big\{C_K(\cdot, \theta_K);\ \theta_K \in \Theta_K \subset \mathbb{R}^{d_K}\Big\}$ be a list of candidate copula models. Model selection can be made for instance using the classical Akaike information criterion (AIC) [23] or Bayesian information criterion (BIC) [24], derived from the semiparametric log-likelihood, see e.g. [25,26]. Denote by $c_k(\cdot, \theta_k)$ the density of the copula $\mathbb{C}_k(\cdot, \theta_k)$, for all k. The AIC of a given model k, $k = 1, \ldots, K$, can be written as

$$AIC(k) = -2 \sup_{\theta_k \in \Theta_k} \sum_{n=1}^{N} \log c_k\Big(\hat{F}_{S_1}(s_1(n)), \ldots, \hat{F}_{S_p}(s_p(n)), \theta_k\Big) + 2\, d_k, \tag{14}$$

where d_k is the dimension of the parameter space $\boldsymbol{\Theta}_k$, and $\hat{F}_{S_i}(\cdot)$ is the rescaled empirical distribution function of the random variable (the component) S_i, $i = 1\ldots, p$, defined by

$$\forall\, x \in \mathbb{R},\quad \hat{F}_{S_i}(x) := \frac{1}{N+1} \sum_{n=1}^{N} \mathbb{1}_{]-\infty, x]}(s_i(n))$$

The rescaling by $N + 1$ instead of N is used in order to avoid difficulties of possible unboundedness of the considered copula density $c_k(\boldsymbol{u}, \theta_k)$ when u_i tends to 1.

The BIC, of a given model k, is defined by

$$BIC(k) = -2 \sup_{\theta_k \in \Theta_k} \sum_{n=1}^{N} \log c_k\Big(\hat{F}_{S_1}(s_1(n)), \ldots, \hat{F}_{S_p}(s_p(n)), \theta_k\Big) + d_k \log(N). \tag{15}$$

The optimal model in term of AIC (respectively, BIC) is the one that minimizes the AIC (respectively, BIC) values, namely, the density copula model $\Big\{c_{k^*}(\cdot, \theta_{k^*});\ \theta_{k^*} \in \Theta_{k^*} \subset \mathbb{R}^{d_{k^*}}\Big\}$ where

$$k^* = \operatorname*{argmin}_{k \in \{1, \ldots, K\}} AIC(k)$$

or

$$k^* = \operatorname*{argmin}_{k \in \{1, \ldots, K\}} BIC(k).$$

Denote, simply, $\Big\{c(\cdot, \theta);\ \theta \in \Theta \subset \mathbb{R}^d\Big\}$ a selected model according to one of the above procedures, then the parameter θ of the considered copula model can be estimated by maximizing the semiparametric log-likelihood

$$\hat{\theta} = \operatorname*{argsup}_{\theta \in \Theta} \sum_{n=1}^{N} \log c\Big(\hat{F}_{S_1}(s_1(n)), \ldots, \hat{F}_{S_p}(s_p(n)), \theta\Big).$$

4. A combined TV-copula approach for BSS

Our TV-Copula approach for BSS proceeds in two steps: (i) it uses the TV regularization technique for denoising the observed signals; (ii) it separates the mixtures by minimizing Kullback–Leibler divergence between copulas, penalized by a TV-regularization term. In the following, we describe these two steps.

4.1. Step 1: denoising the observed signals

As in [13], we would obtain the denoised mixture signal

$$\boldsymbol{x}(t) := (x_1(t), \ldots, x_p(t))^\top, \quad t \in [0, T], \tag{16}$$

from the (noisy) observed one $\overline{\boldsymbol{x}}(t)$, see (4), by means of the following variational minimization problem

$$x_i(\cdot) = \arg\min_{w_i \in \mathcal{X}_c} \left\{ \frac{1}{T} \int_0^T \frac{1}{2} \big(w_i(t) - \overline{x}_i(t)\big)^2 \, dt + \lambda \frac{1}{T} \int_0^T \phi\big(w_i(t)\big) dt \right\}, \quad i = 1, \ldots, p, \tag{17}$$

where $\lambda > 0$ is the parameter of penalization, $\phi(\cdot)\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a well chosen function, $w_i(t)$ is the first derivative of the signal $w_i(t)$ with respect to time t, and $\mathcal{X}_c$ is some appropriate space of signals. In practice, generally, in order to reduce the noise, the function $\phi(\cdot)$ is chosen to encourage smoothing in regions where the variations of the signal are weak, that is $w_i(t) \approx 0$, and to preserve discontinuities where $w_i(t)$ is strong. The total variation case corresponds to the choice $\phi(x) = |x|$, which will be adopted here. One can make other choices like $\phi(x) = \sqrt{1 + x^2}$. A natural space when dealing with the above continuous variational problem (17), with $\phi(x) = |x|$, is $\mathcal{X}_c = BV([0, T])$, the space of all real valued functions on the interval $[0, T]$ with bounded variation; see e.g. [27] and [28]. Here T is the observation time of the signals. The mathematical treatment of the above problem, in its continuous form, is beyond the scope of this paper. We will consider only the corresponding discrete version which will be described in Section 5.1.

4.2. Step 2: BSS via minimizing TV-regularized KL-divergence between copulas

The aim of the following step is to construct an estimated source signal

$$\hat{\boldsymbol{s}}(t) := \widehat{\boldsymbol{B}} \boldsymbol{x}(t)$$

from the denoised signal $\boldsymbol{x}(t) := (x_1(t), \ldots, x_p(t))^\top$ obtained as solution of (17). Our approach consists in minimizing, with respect to $\boldsymbol{B}$ (on the demixing matrices space), an estimate (to be defined below) with some criterion of the form

$$\mathcal{J}\colon \boldsymbol{B} \mapsto \mathcal{J}(\boldsymbol{B}) := \mathcal{J}_{\text{sep}}(\boldsymbol{B}) + \mathcal{J}_{\text{reg}}(\boldsymbol{y}), \tag{18}$$

where

$$\boldsymbol{y}(t) := \boldsymbol{B}\boldsymbol{x}(t) \in \mathbb{R}^p,$$

and $\boldsymbol{x}(t) = (x_1(t), \ldots, x_p(t))^\top \in \mathbb{R}^p$ is obtained from (17); $\mathcal{J}_{\text{sep}}(\cdot)$ is some separating criterion, while $\mathcal{J}_{\text{reg}}(\boldsymbol{y})$ is a regularization term devoted to denoising. For the denoising phase, for both cases, of independent or dependent source components, we propose to use as regularization term the following one:

$$\mathcal{J}_{\text{reg}}(\boldsymbol{y}) := \gamma \frac{1}{2T} \int_0^T \| \boldsymbol{y}(t) - \overline{\boldsymbol{y}}(t) \|^2 \, dt + \mu \frac{1}{T} \int_0^T |\nabla \boldsymbol{y}(t)| dt, \quad \gamma > 0, \ \mu > 0, \tag{19}$$

where $\overline{\boldsymbol{y}} = \boldsymbol{B}\,\overline{\boldsymbol{x}}$; the real numbers γ and μ are regularization parameters to be suitably chosen by the user. In the above display, $\|\cdot\|$ is used to denote the Euclidian norm on $\mathbb{R}^p$, and $|\nabla \boldsymbol{y}(t)| := \sum_{i=1}^p |y_i(t)|$, for all t. The first term $\mathcal{J}_{\text{sep}}(\cdot)$, in (18), is a measure of dependence, or similarity between copula densities. Its form depends on the fact that the source components are independent or dependent. Hence, to describe the criterion $\mathcal{J}_{\text{sep}}(\cdot)$, we will consider separately, the case where the source components are independent and the case where they are dependent. First, we will consider the following stochastic modeling of the continuous time signals $\boldsymbol{s}(t)$, $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t) := \boldsymbol{B}\boldsymbol{x}(t)$, $t \in [0, T]$. We assume that the above last random processes are stationary, so that the corresponding sampled versions, with certain time period, say T_e,

$$\boldsymbol{s}(n) := \boldsymbol{s}(nT_e), \quad \boldsymbol{x}(n) := \boldsymbol{x}(nT_e) \quad \text{and} \quad \boldsymbol{y}(n) := \boldsymbol{y}(nT_e) = \boldsymbol{B}\boldsymbol{x}(n), \quad n = 1, \ldots, N, \tag{20}$$

can be considered as realizations of random vectors in $\mathbb{R}^p$, which will be denoted

$$\boldsymbol{S}, \quad \boldsymbol{X} \quad \text{and} \quad \boldsymbol{Y} := \boldsymbol{B}\boldsymbol{X}, \tag{21}$$

respectively. It has been shown, in [14], for the noisy-free mixtures case, that if we have available some prior information about the copula density of the random source vector $\boldsymbol{S}$, one can detect, under some assumptions, both the mixing matrix and the sources uniquely (up to scale and permutation indeterminacies) for both cases of independent or dependent source components. Denote by

$$\mathbb{F}_{\boldsymbol{Y}}(\cdot)\colon \boldsymbol{y} \in \mathbb{R}^p \mapsto \mathbb{F}_{\boldsymbol{Y}}(\boldsymbol{y}) := \mathbb{F}_{\boldsymbol{Y}}(y_1, \ldots, y_p) := \mathbb{P}(Y_1 \le y_1, \ldots, Y_p \le y_p), \tag{22}$$

the joint distribution function of the random vector $\boldsymbol{Y} := (Y_1, \ldots, Y_p)^\top \in \mathbb{R}^p$, with continuous marginal distribution functions

$$F_{Y_i}(\cdot)\colon y_i \in \mathbb{R} \mapsto F_{Y_i}(y_i) := \mathbb{P}(Y_i \le y_i), \quad \forall\, i = 1, \ldots, p.$$

Let, if it exists, $f_{\boldsymbol{Y}}(\cdot)$ be the probability density of the random vector $\boldsymbol{Y}$, and, respectively, $f_{Y_1}, \ldots, f_{Y_p}$, the marginal probability densities of the components $Y_1, \ldots, Y_p$. The mutual information (MI) of $\boldsymbol{Y}$ is defined by

$$\mathbb{I}(\boldsymbol{Y}) := \int_{\mathbb{R}^p} \log \frac{f_{\boldsymbol{Y}}(\boldsymbol{y})}{\prod_{i=1}^p f_{Y_i}(y_i)} f_{\boldsymbol{Y}}(\boldsymbol{y}) \, d\boldsymbol{y} = \mathbb{E}\left(\log \frac{f_{\boldsymbol{Y}}(\boldsymbol{Y})}{\prod_{i=1}^p f_{Y_i}(Y_i)} \right) \tag{23}$$

which equals the Kullback–Leibler divergence, denote it $\mathbb{K}(\cdot, \cdot)$, between the joint density $f_{\boldsymbol{Y}}(\cdot)$ and the product density $\prod_{i=1}^p f_{Y_i}(\cdot)$ of the margins $f_{Y_i}(\cdot)$, $i = 1 \ldots, p$, that is

$$\mathbb{I}(\boldsymbol{Y}) = \mathbb{K}\left(f_{\boldsymbol{Y}}, \prod_{i=1}^p f_{Y_i} \right). \tag{24}$$

In (23), the notation $\mathbb{E}(\cdot)$ is used to denote the mathematical expectation. Note also that the above criterion $\mathbb{I}(\boldsymbol{Y})$ is nonnegative and achieves its minimum value zero iff $f_{\boldsymbol{Y}}(\cdot) = \prod_{i=1}^p f_{Y_i}(\cdot)$, i.e., iff the components of the vector $\boldsymbol{Y}$ are statistically independent. We give now, the form of the separating term $\mathcal{J}_{\text{sep}}(\cdot)$ of the criterion (18).

4.2.1. The case of independent source components

Assume that the source components are independent. Using the relation (12), and applying the change variable formula for multiple integrals, one can show that $\mathbb{I}(\boldsymbol{Y})$ can be written, in terms of copula densities, as follows

$$\mathbb{I}(\boldsymbol{Y}) = \int_{[0,1]^p} \log\big(c_{\boldsymbol{Y}}(\boldsymbol{u})\big) c_{\boldsymbol{Y}}(\boldsymbol{u}) d\boldsymbol{u} \tag{25}$$

$$\mathbb{I}(\boldsymbol{Y}) = \mathbb{E}\Big(\log c_{\boldsymbol{Y}}\Big(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p) \Big) \Big) =: -\,\mathbb{H}(c_{\boldsymbol{Y}}) \tag{26}$$

$$\mathbb{I}(\boldsymbol{Y}) = \int_{[0,1]^p} \log\left(\frac{c_{\boldsymbol{Y}}(\boldsymbol{u})}{1} \right) c_{\boldsymbol{Y}}(\boldsymbol{u}) d\boldsymbol{u} \tag{27}$$

$$\mathbb{I}(\boldsymbol{Y}) = \int_{[0,1]^p} \log\left(\frac{c_{\boldsymbol{Y}}(\boldsymbol{u})}{c_{\Pi}(\boldsymbol{u})}\right) c_{\boldsymbol{Y}}(\boldsymbol{u}) d\boldsymbol{u} \tag{28}$$

$$\mathbb{I}(\boldsymbol{Y}) = \mathbb{E}\left(\log \frac{c_{\boldsymbol{Y}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}{c_{\Pi}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}\right) = \mathbb{K}\left(c_{\boldsymbol{Y}}, c_{\Pi}\right), \tag{29}$$

where $c_{\boldsymbol{Y}}(\cdot)$ is the copula density of $\boldsymbol{Y}$, $c_{\Pi}(\cdot) := \mathbb{1}_{[0,1]^p}(\cdot)$ is the copula density of independence, and $\mathbb{E}(\cdot)$ the mathematical expectation operator. The above equations mean that the MI of a random vector $\boldsymbol{Y}$ can be seen as the opposite of the Shannon entropy $\mathbb{H}(\cdot)$ of the copula density $c_{\boldsymbol{Y}}(\cdot)$ of $\boldsymbol{Y}$, or as the Kullback–Leibler divergence between the copula density $c_{\boldsymbol{Y}}(\cdot)$ of $\boldsymbol{Y}$ and the copula density $c_{\Pi}(\cdot)$ of independence. Moreover, $\mathbb{I}(\boldsymbol{Y}) = \mathbb{K}\left(c_{\boldsymbol{Y}}, c_{\Pi}\right)$ is nonnegative and achieves its minimum value zero iff $c_{\boldsymbol{Y}}(\boldsymbol{u}) = c_{\Pi}(\boldsymbol{u})$, $\forall\, \boldsymbol{u} \in [0,1]^p$, namely, iff the components of the vector $\boldsymbol{Y}$ are independent. The separating term will be chosen then to be

$$\mathcal{J}_{\text{sep}}(\boldsymbol{B}) := \mathcal{J}^{\text{ind}}_{\text{sep}}(\boldsymbol{B}) := \mathbb{K}\left(c_{\boldsymbol{Y}}, c_{\Pi}\right) = \mathbb{E}\left(\log c_{\boldsymbol{Y}}\left(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p)\right)\right). \tag{30}$$

We have that the function $\boldsymbol{B} \mapsto \mathcal{J}^{\text{ind}}_{\text{sep}}(\boldsymbol{B})$ is nonnegative and attains its minimum value zero iff $\boldsymbol{B} = \boldsymbol{A}^{-1}$ (up to scale and permutation indeterminacies of rows).

4.2.2. The case of dependent source components

In the case where the source components are dependent, we assume that we have at hand some prior information about the copula density of the random source vector $\boldsymbol{S}$. Note that this is possible for many practical problems, it can be done, see Section 3, from training sample $\boldsymbol{S}$, by a model selection procedure, see e.g. [29], among semiparametric copula density models $\{c_\theta(\cdot);\ \theta \in \Theta \subset \mathbb{R}^d\}$, typically indexed by a multivariate parameter θ; see e.g. [22,21] for many examples of such models. The parameter θ can be estimated using maximum semiparametric likelihood, see e.g. [25,26]. Denote then by $\hat{\theta}$, the obtained value of θ and $c_{\hat{\theta}}(\cdot)$ the copula density modeling the dependency structure of the source components. Obviously, since the source components are assumed to be dependent, $c_{\hat{\theta}}(\cdot)$ should be different from the density copula of independence $c_{\Pi}(\cdot)$. Hence, we naturally replace in (29), $c_{\Pi}(\cdot)$ by $c_{\hat{\theta}}(\cdot)$, and propose the following separating criterion $\mathcal{J}_{\text{sep}}(\cdot)$

$$\begin{aligned}\mathcal{J}_{\text{sep}}(\boldsymbol{B}) &:= \mathcal{J}^{\text{dep}}_{\text{sep}}(\boldsymbol{B}) := \int_{[0,1]^p} \log\left(\frac{c_{\boldsymbol{Y}}(\boldsymbol{u})}{c_{\hat{\theta}}(\boldsymbol{u})}\right) c_{\boldsymbol{Y}}(\boldsymbol{u}) d\boldsymbol{u} \\ &:= \mathbb{E}\left(\log \frac{c_{\boldsymbol{Y}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}{c_{\hat{\theta}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}\right),\end{aligned} \tag{31}$$

which equals $\mathbb{K}(c_{\boldsymbol{Y}}, c_{\hat{\theta}})$, the Kullback–Leibler divergence between the copula densities $c_{\boldsymbol{Y}}(\cdot)$ and $c_{\hat{\theta}}(\cdot)$. Moreover, one can show that the function $\boldsymbol{B} \mapsto \mathcal{J}^{\text{dep}}_{\text{sep}}(\boldsymbol{B})$ is nonnegative, and attains its minimum value iff $\boldsymbol{B} = \boldsymbol{A}^{-1}$ (up to scale and permutation indeterminacies of rows), provided that the copula density $c_{\hat{\theta}}(\cdot)$ of the random source vector $\boldsymbol{S}$ satisfies the following assumption: for any regular matrix $\boldsymbol{M}$, if the copula density of the random vector $\boldsymbol{M}\,\boldsymbol{S}$ equals $c_{\hat{\theta}}(\cdot)$, then $\boldsymbol{M} = \boldsymbol{D}\,\boldsymbol{P}$, for some diagonal matrix $\boldsymbol{D}$ and permutation matrix $\boldsymbol{P}$.

Remark 1. The above criterion (31) supposes the knowledge of the copula density model $c_\theta(\cdot)$ (with known parameter θ), or the availability of training sample of the random vector source $\boldsymbol{S}$, from which we can obtain the copula density $c_{\hat{\theta}}(\cdot)$ modeling the dependency structure of the source components, using for example the model selection criteria described in Section 4.2.2. This case can be applied, for example, for the following real problem. In separating recto and verso sides from two-sides scanned archival documents, the observed signals (the two scans of an ancient document for example) are unknown mixtures of sources (the recto and verso sides of the document). In this case, the training samples of the source vector $\boldsymbol{S}$ can be obtained by scanning opaque documents for which the sources are of the same nature as those of the ancient documents.

When only the model is known (with unknown parameter), which is the case for example in separating rotating machinery vibration signals, where we may have an idea about the dependency structure of the source components, we can adapt the above criterion as follows. Denote by $\left\{c_\theta(\cdot);\ \theta \in \Theta\right\}$ the density copula model of the dependency structure of the source components. The criterion (31) can be in this case replaced by the following one

$$\begin{aligned}\mathcal{J}_{\text{sep}}(\boldsymbol{B}) &:= \inf_{\theta \in \Theta} \int_{[0,1]^p} \log\left(\frac{c_{\boldsymbol{Y}}(\boldsymbol{u})}{c_{\theta}(\boldsymbol{u})}\right) c_{\boldsymbol{Y}}(\boldsymbol{u}) \mathrm{d}\boldsymbol{u} \\ &:= \inf_{\theta \in \Theta} \mathbb{E}\left(\log \frac{c_{\boldsymbol{Y}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}{c_{\theta}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}\right).\end{aligned} \tag{32}$$

Note that the above separation criterion function $\mathbf{B} \mapsto \inf_{\theta \in \Theta} \mathbb{E}\left(\log \frac{c_{\boldsymbol{Y}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}{c_{\theta}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}\right)$ is nonnegative and achieves its minimum value zero iff $\mathbf{B} = \mathbf{A}^{-1}$ (up to scale and permutation indeterminacies), provided that the copula density model $\left\{\mathbf{c}_{\boldsymbol{\theta}}(\cdot);\ \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d\right\}$ of $\mathbf{S}$ satisfies the following identifiability assumption: for any regular matrix $\mathbf{M}$, if the copula density of $\mathbf{MS} \in \left\{\mathbf{c}_{\boldsymbol{\theta}}(\cdot);\ \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d\right\}$, then $\mathbf{M} = \mathbf{DP}$, where $\mathbf{D}$ is diagonal and $\mathbf{P}$ is a permutation.

In the more general case, when both the model and the parameter are unknown, the separation criterion (31) may be adapted, in a similar way as in our previous work [14] Section 5.3, as follows. We consider a list of a number L of copula density models $M_\ell := \left\{c_{\theta_\ell}(\cdot);\ \theta_\ell \in \Theta_\ell\right\}$, $\ell = 1, \ldots, L$, for modeling the dependency structure of the source components, with parameter spaces $\Theta_1, \ldots, \Theta_L$, with the same dimension. Assume that each model M_ℓ, $\ell = 1, \ldots, L$, satisfies the above identifiability condition. Then, the separation criterion function $\mathbf{B} \mapsto \inf_{\ell=1,\ldots,L} \inf_{\boldsymbol{\theta}_\ell \in \boldsymbol{\Theta}_\ell} \mathbb{E}\left(\log \frac{c_{\boldsymbol{Y}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}{c_{\theta_\ell}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}\right)$ is nonnegative and achieves its minimum value zero iff $\mathbf{B} = \mathbf{A}^{-1}$ (up to scale and permutation indeterminacies). Hence, we propose to apply the method described above for each model M_ℓ, with separation criterion of the form $\mathcal{J}_{\text{sep}}(\boldsymbol{B}) = \inf_{\boldsymbol{\theta}_\ell \in \boldsymbol{\Theta}_\ell} \mathbb{E}\left(\log \frac{c_{\boldsymbol{Y}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}{c_{\theta_\ell}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}\right)$, and then choose the solution that minimizes the criterion over all considered models, i.e.,

$$\widehat{\mathbf{B}} := \arg\inf_{\mathbf{B}} \inf_{\theta \in \Theta_{\ell^*}} \mathbb{E}\left(\log \frac{c_{\boldsymbol{Y}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}{c_{\theta^*_{\ell}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}\right),$$

where

$$\ell^* := \arg\min_{\ell=1,\ldots,L} \inf_{\mathbf{B}} \inf_{\theta \in \Theta_\ell} \mathbb{E}\left(\log \frac{c_{\boldsymbol{Y}}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}{c_{\theta_\ell}(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p))}\right).$$

5. Discretization and statistical estimation

In this section, we describe how to make operational the above methodology. Precisely, using both numerical and statistical techniques, we give hereafter the discrete versions of the

regularization terms (17) and (19), as well as statistical estimates of the separation terms (30) and (31). In all the sequel, the continuous time signals are sampled with a certain period T_e. We need to define additional notation. To each scalar signal, say $u(t)$, $t \in [0, T]$, one associate a vector $\mathbb{U} := (u(1), \ldots, u(N))^\top \in \mathcal{X}_d := \mathbb{R}^N$ such that $u(k) := u(kT_e)$ for all $k = 1, \ldots, N$. The vector space $\mathcal{X}_d := \mathbb{R}^N$ will be equipped with the Euclidian inner product

$$\langle \mathbb{U}, \mathbb{V} \rangle := \sum_{k=1}^{N} u(k)v(k),$$

for all $\mathbb{U}, \mathbb{V} \in \mathcal{X}_d$. For any $\mathbb{U} \in \mathcal{X}_d$, denote by $\mathbb{U}' := (u'(1), \ldots, u'(N))^\top \in \mathcal{X}_d$, the vector obtained by numerical differentiation

$$u'(k) := \frac{u(k+1) - u(k)}{T_e}, \quad \text{for all } k = 1, \ldots, N-1, \tag{33}$$

and $u'(N) = 0$. Due to the discretization of the first derivatives, and in order to avoid the *edge effect*, we define the *backward derivative* of any $\mathbb{U} \in \mathcal{X}_d$, denoted $\mathbb{U}^* := (u^*(1), \ldots, u^*(N))^\top$, by

$$u^*(1) := \frac{u(1)}{T_e}, \quad u^*(N) := -\frac{u(N-1)}{T_e}$$

and

$$u^*(k) := \frac{u(k) - u(k-1)}{T_e}, \quad \text{for all } k = 2, \ldots, N-1.$$

Here, the backward derivatives $(\cdot)^*$ is the discrete adjoint operator of $-(\cdot)'$. That is, for all $\mathbb{U}, \mathbb{V} \in \mathcal{X}_d$, we have

$$\langle \mathbb{U}^*, \mathbb{V} \rangle = -\langle \mathbb{U}, \mathbb{V}' \rangle. \tag{34}$$

All these definitions can be extended in a natural way to elements of the Cartesian product space $\mathcal{X}_d^p := \mathbb{R}^p \times \mathbb{R}^N$. In the sequel, elements of the vector space $\mathcal{X}_d = \mathbb{R}^N$ will be denoted by double capital symbols, while the elements of the vector space $\mathcal{X}_d^p$ are denoted by bold double capital symbols, which can be considered as $p \times N$-matrices.

5.1. Denoising the discrete observed signal

In this section, we show how to compute the denoised observations

$$\mathbb{X} := \Big(\boldsymbol{x}(n) = \big(x_1(n), \ldots, x_p(n)\big)^\top, \ n = 1, \ldots, N \Big)$$

from the observed noisy ones

$$\overline{\mathbb{X}} := \Big(\overline{\boldsymbol{x}}(n) = \big(\overline{x}_1(n), \ldots, \overline{x}_p(n)\big)^\top, \ n = 1, \ldots, N \Big).$$

This will be obtained by solving the discrete version

$$x_i(\cdot) = \arg\min_{\mathbb{W}_i \in \mathcal{X}_d} \left\{ \frac{1}{N} \sum_{n=1}^{N} \frac{1}{2} \big(w_i(n) - \overline{x}_i(n)\big)^2 + \lambda \frac{1}{N} \sum_{n=1}^{N} |w_i(n)| \right\}, \quad i = 1, \ldots, p, \tag{35}$$

of the optimization problem (17). We start with the following proposition which gives a simple characterization of the solution; see [30] for a proof.

Proposition 1 (*[30]*). *The discrete version (35) of the optimization problem (17) has a unique solution given by*

$$\mathbb{X} = \overline{\mathbb{X}} - \prod_{\lambda G} \overline{\mathbb{X}}, \tag{36}$$

where $\prod_{\lambda G}(\cdot)$ is the Euclidian orthogonal projection operator on the convex set λG with

$$G := \{ \mathbb{V}^* \text{ such that } \mathbb{V} \in \mathcal{X}_d^p \text{ with } |v_i(n)| \le 1, \ \forall\, (i, n) \in \{1, \ldots, p\} \times \{1, \ldots, N\} \}. \tag{37}$$

Thus, to compute $\mathbb{X}$, we are led to compute the projection operator $\prod_{\lambda G}(\cdot)$ on the convex set λG. For the denoising step, we then propose Algorithm 1; the reader may refer to [13] for more details, particularly, for a proof on convergence of this algorithm.

Algorithm 1. The denoising step.

Data: $\overline{\mathbb{X}}$ the observed noised signal
Result: $\mathbb{X}$ the denoised signal
Initialization: $q = 0$, $\mathbb{X}^{(0)} = \overline{\mathbb{X}}$. Given $\epsilon > 0$, $\lambda > 0$ and $\rho > 0$ suitably chosen
Do: $q = q + 1$
  for $j = 1, \ldots, p$
    $\mathbb{X}_j^{(q)} = \prod_{\lambda G}(\mathbb{X}_j^{(q-1)} + 2\lambda\rho(\lambda(\mathbb{X}_j^{(q-1)})^* - \overline{\mathbb{X}}_j)'$
  end
Until $\| \mathbb{X}^{(q)} - \mathbb{X}^{(q-1)} \| < \epsilon$
$\mathbb{X} = \mathbb{X}^{(q)}$.

In all the sequel, the filtered observed signal $\mathbb{X}$ will be considered as the denoised version of the observed signal $\overline{\mathbb{X}}$.

5.2. BSS by minimizing an estimate of TV-regularized KL-divergence between copulas

5.2.1. The case of independent source components

Recall that, when the source components are independent, the criterion function (18) is defined by

$$\mathcal{J}^{\text{ind}}(\cdot) : \boldsymbol{B} \mapsto \mathcal{J}^{\text{ind}}(\boldsymbol{B}) := \mathcal{J}^{\text{ind}}_{\text{sep}}(\boldsymbol{B}) + \mathcal{J}_{\text{reg}}(\boldsymbol{y}), \tag{38}$$

where $\mathcal{J}^{\text{ind}}_{\text{sep}}(\boldsymbol{B})$ is given by, see (30),

$$\mathcal{J}^{\text{ind}}_{\text{sep}}(\boldsymbol{B}) := \mathbb{K}\Big(c_{\boldsymbol{Y}}, c_{\prod}\Big) = \mathbb{E}\Big(\log c_{\boldsymbol{Y}}\big(F_{Y_1}(Y_1), \ldots, F_{Y_p}(Y_p)\big)\Big)$$

and

$$\mathcal{J}_{\text{reg}}(\boldsymbol{y}) := \gamma \frac{1}{2T} \int_0^T \| \boldsymbol{y}(t) - \overline{\boldsymbol{y}}(t) \|^2 \, dt + \mu \frac{1}{T} \int_0^T |\nabla \boldsymbol{y}(t)| \, dt, \quad \gamma > 0, \ \mu > 0,$$

with $\boldsymbol{y}(t) = \boldsymbol{B}\boldsymbol{x}(t)$ and $\overline{\boldsymbol{y}} = \boldsymbol{B}\overline{\boldsymbol{x}}(t)$. Using the stochastic modeling (20) and the relation (30), we propose to approximate the criterion (38) by

$$\boldsymbol{B} \mapsto \widehat{\mathcal{J}^{\text{ind}}}(\boldsymbol{B}) := \widehat{\mathcal{J}^{\text{ind}}_{\text{sep}}}(\boldsymbol{B}) + \mathcal{J}_{\text{reg,d}}(\boldsymbol{y}), \tag{39}$$

where

$$\mathcal{J}_{\text{reg,d}}(\boldsymbol{y}) := \frac{\gamma}{2N} \sum_{i=1}^{N} \| \boldsymbol{y}(i) - \overline{\boldsymbol{y}}(i) \|^2 + \frac{\mu}{N} \sum_{i=1}^{N} |\nabla \boldsymbol{y}(i)| \tag{40}$$

is the discrete version of $\mathcal{J}_{\text{reg}}(\boldsymbol{y})$, and $\boldsymbol{B} \mapsto \widehat{\mathcal{J}^{\text{ind}}_{\text{sep}}}(\boldsymbol{B})$ is the statistical estimate, of the separating criterion $\boldsymbol{B} \mapsto \mathcal{J}^{\text{ind}}_{\text{sep}}(\boldsymbol{B})$, which we define as follows

$$\boldsymbol{B} \mapsto \widehat{\mathcal{J}^{\text{ind}}_{\text{sep}}}(\boldsymbol{B}) := \frac{1}{N} \sum_{i=1}^{N} \log\left(\hat{c}_Y\big(\hat{F}_{Y_1}(y_1(i)), \ldots, \hat{F}_{Y_p}(y_p(i))\big) \right), \tag{41}$$

where

$$\hat{c}_{\boldsymbol{Y}}(\hat{F}_{Y_1}(y_1(i)),\ \ldots,\ \hat{F}_{Y_p}(y_p(i))) := \frac{1}{NH_1\cdots H_p}\sum_{\ell=1}^{N}\prod_{j=1}^{p} k\left(\frac{\hat{F}_{Y_j}(y_j(i)) - \hat{F}_{Y_j}(y_j(\ell))}{H_j}\right),$$

is the kernel estimate of the copula density $c_{\boldsymbol{Y}}(\cdot)$, and $\hat{F}_{Y_j}(\cdot)$, $\forall\, j = 1, \ldots, p$, is the smooth estimate of the marginal distribution function $F_{Y_j}(\cdot)$ of the random variable Y_j, defined for any real number $r \in \mathbb{R}$, by

$$\hat{F}_{Y_j}(r) := \frac{1}{N}\sum_{\ell=1}^{N} K\left(\frac{r - y_j(\ell)}{h_j}\right),$$

where $K(\cdot)$ is the primitive of a kernel $k(\cdot)$, a symmetric centered probability density. In our forthcoming simulation study, we will take as kernel $k(\cdot)$ a standard Gaussian probability density. A more appropriate choice of the kernel $k(\cdot)$, for estimating the copula density, can be done according to [31], which copes with the boundary effect. The bandwidth parameters $H_1, \ldots, H_p$ and $h_1, \ldots, h_p$ will be chosen according to Silverman's rule of thumb, see e.g. [32], i.e., for all $j = 1, \ldots, p$, we take $H_j = \left(\frac{4}{p+2}\right)^{\frac{1}{p+4}} N^{\frac{-1}{p+4}}\,\widehat{\Sigma}_j$, and $h_j = \left(\frac{4}{3}\right)^{\frac{1}{5}} N^{\frac{-1}{5}}\,\hat{\sigma}_j$, where $\widehat{\Sigma}_j$ and $\hat{\sigma}_j$ are, respectively, the empirical standard deviation of the data $\hat{F}_{Y_j}(y_j(1)), \ldots, \hat{F}_{Y_j}(y_j(N))$ and $y_j(1), \ldots, y_j(N)$. The source signal vector $\boldsymbol{s}(i)$, $i = 1, \ldots, N$, will then be estimated by

$$\hat{\boldsymbol{s}}(i) = \widehat{\boldsymbol{B}}\,\boldsymbol{x}(i), \quad i = 1, \ldots, N,$$

where

$$\widehat{\boldsymbol{B}} := \underset{\boldsymbol{B}}{\operatorname{arginf}}\ \widehat{\mathcal{J}^{\mathrm{ind}}}(\boldsymbol{B}),$$

which can be computed using a gradient descent type algorithm. In fact, straightforward computation shows that, the gradient in $\boldsymbol{B}$ of the estimated criterion $\boldsymbol{B} \mapsto \widehat{\mathcal{J}^{\mathrm{ind}}}(\boldsymbol{B})$, can be written as

$$\frac{d\,\widehat{\mathcal{J}^{\mathrm{ind}}}(\boldsymbol{B})}{d\boldsymbol{B}} = \frac{1}{N}\sum_{i=1}^{N}\frac{\frac{d}{d\boldsymbol{B}}\hat{c}_{\boldsymbol{Y}}(\boldsymbol{u}(i))}{\hat{c}_{\boldsymbol{Y}}(\boldsymbol{u}(i))} + \frac{\gamma}{N}\sum_{i=1}^{N}(\boldsymbol{y}(i) - \overline{\boldsymbol{y}}(i))(\boldsymbol{x}(i) - \overline{\boldsymbol{x}}(i))^{\top} + \frac{\mu}{N}\sum_{i=1}^{N}\left(\frac{\nabla \boldsymbol{y}(i)}{|\nabla \boldsymbol{y}(i)|}\right)^{*}\boldsymbol{x}(i)^{\top}, \tag{42}$$

where $\frac{d}{d\boldsymbol{B}} := \left(\frac{\partial}{\partial \boldsymbol{B}_{l,j}}\right)$, $l, j = 1, \ldots, p$, $\boldsymbol{u}(i) := (\hat{F}_{Y_1}(y_1(i)), \ldots, \hat{F}_{Y_p}(y_p(i)))^{\top}$ and,

$$\frac{\partial \hat{c}_{\boldsymbol{Y}}(\hat{F}_{Y_1}(y(i)),\ \ldots,\ \hat{F}_{Y_p}(y_p(i)))}{\partial \boldsymbol{B}_{l,j}} = \frac{1}{NH_1\cdots H_p}\sum_{m=1}^{N}\prod_{j=1, j\neq l}^{p} k\left(\frac{\hat{F}_{Y_j}(y_j(m)) - \hat{F}_{Y_j}(y_j(i))}{H_j}\right) \times k'\left(\frac{\hat{F}_{Y_l}(y_l(m)) - \hat{F}_{Y_l}(y_l(i))}{H_l}\right)\frac{1}{H_l}\frac{\partial(\hat{F}_{Y_l}(y_l(m)) - \hat{F}_{Y_l}(y_l(i)))}{\partial \boldsymbol{B}_{l,j}}, \tag{43}$$

with

$$\frac{\partial(\hat{F}_{Y_l}(y_l(m))}{\partial \boldsymbol{B}_{l,j}} = \frac{1}{Nh_l}\sum_{i=1}^{N} k\left(\frac{y_l(i) - y_l(m)}{h_l}\right)(x_j(i) - x_j(m)) \tag{44}$$

and

$$\frac{\partial \hat{F}_{Y_l}(y_l(i))}{\partial \boldsymbol{B}_{l,j}} = \frac{1}{N}h_l\sum_{n=1}^{N} k\left(\frac{y_l(n) - y_l(i)}{h_l}\right)(x_j(n) - x_j(i)). \tag{45}$$

We can then derive the following Algorithm 2.

Algorithm 2. BSS algorithm for independent source components.

Data: $\overline{\boldsymbol{x}}$ the observed signal
Result: $\hat{\boldsymbol{s}}$ the estimated source signal
Initialization: Compute $\boldsymbol{x} = \overline{\boldsymbol{x}} - \prod_{\lambda G}\overline{\boldsymbol{x}}$ from Algorithm 1,
$\boldsymbol{B}^{(0)} = \boldsymbol{I}_p$, $\boldsymbol{y}^{(0)} = \boldsymbol{B}^{(0)}\,\boldsymbol{x}$. Given $\varepsilon > 0$, $\nu > 0$.
Do: Update $\boldsymbol{B}$ and $\boldsymbol{y}$:

$$\boldsymbol{B}^{(q+1)} = \boldsymbol{B}^{(q)} - \nu\frac{d\,\widehat{\mathcal{J}^{\mathrm{ind}}}(\boldsymbol{B})}{d\boldsymbol{B}}$$

$$\boldsymbol{y}^{(q+1)} = \boldsymbol{B}^{(q+1)}\,\boldsymbol{x}.$$

Until $\| \boldsymbol{B}^{(q+1)} - \boldsymbol{B}^{(q)} \| < \epsilon$
$\hat{\boldsymbol{s}} = \boldsymbol{y}^{(q+1)}$.

5.2.2. The case of dependent source components

Recall that, in the case where the source components are dependent, the criterion function (18) is defined by

$$\mathcal{J}^{\mathrm{dep}}(\cdot)\colon \boldsymbol{B} \mapsto \mathcal{J}^{\mathrm{dep}}(\boldsymbol{B}) := \mathcal{J}^{\mathrm{dep}}_{\mathrm{sep}}(\boldsymbol{B}) + \mathcal{J}_{\mathrm{reg}}(\boldsymbol{y}), \tag{46}$$

where $\mathcal{J}^{\mathrm{dep}}(\boldsymbol{B})$ is defined by, see (31),

$$\mathcal{J}^{\mathrm{dep}}_{\mathrm{sep}}(\boldsymbol{B}) := \int_{[0,1]^p}\log\left(\frac{c_{\boldsymbol{Y}}(\boldsymbol{u})}{c_{\hat{\theta}}(\boldsymbol{u})}\right)c_{\boldsymbol{Y}}(\boldsymbol{u})\,d\boldsymbol{u} := \mathbb{E}\left(\log\frac{c_{\boldsymbol{Y}}(F_{Y_1}(Y_1),\ \ldots,\ F_{Y_p}(Y_p))}{c_{\hat{\theta}}(F_{Y_1}(Y_1),\ \ldots,\ F_{Y_p}(Y_p))}\right).$$

and

$$\mathcal{J}_{\mathrm{reg}}(\boldsymbol{y}) := \gamma\,\frac{1}{2\,T}\int_0^T \|\,\boldsymbol{y}(t) - \overline{\boldsymbol{y}}(t)\|^2\ dt + \mu\,\frac{1}{T}\int_0^T |\nabla \boldsymbol{y}(t)|\ dt, \quad \gamma > 0,\ \mu > 0.$$

Using the stochastic modeling (20) and the relation (31), we propose to approximate the criterion (46) by

$$\boldsymbol{B} \mapsto \widehat{\mathcal{J}^{\mathrm{dep}}}(\boldsymbol{B}) := \widehat{\mathcal{J}^{\mathrm{dep}}_{\mathrm{sep}}}(\boldsymbol{B}) + \mathcal{J}_{\mathrm{reg,d}}(\boldsymbol{y}), \tag{47}$$

where

$$\mathcal{J}_{\mathrm{reg,d}}(\boldsymbol{y}) := \frac{\gamma}{2N}\sum_{i=1}^{N}\|\,\boldsymbol{y}(i) - \overline{\boldsymbol{y}}(i)\|^2 + \frac{\mu}{N}\sum_{i=1}^{N}|\nabla \boldsymbol{y}(i)| \tag{48}$$

is the discrete version of $\mathcal{J}_{\mathrm{reg}}(\boldsymbol{y})$, and $\widehat{\mathcal{J}^{\mathrm{dep}}_{\mathrm{sep}}}(\cdot)\colon \boldsymbol{B} \mapsto \widehat{\mathcal{J}^{\mathrm{dep}}_{\mathrm{sep}}}(\boldsymbol{B})$ is the statistical estimate, of the criterion $\mathcal{J}^{\mathrm{dep}}_{\mathrm{sep}}(\cdot)\colon \boldsymbol{B} \mapsto \mathcal{J}^{\mathrm{dep}}_{\mathrm{sep}}(\boldsymbol{B})$, which we define as follows:

$$\widehat{\mathcal{J}^{\mathrm{dep}}_{\mathrm{sep}}}(\cdot)\colon \boldsymbol{B} \mapsto \widehat{\mathcal{J}^{\mathrm{dep}}_{\mathrm{sep}}}(\boldsymbol{B}) := \frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{\hat{c}_{\boldsymbol{Y}}(\hat{F}_{Y_1}(y(i)),\ \ldots,\ \hat{F}_{Y_p}(y_p(i)))}{\hat{c}_{\hat{\theta}}(\hat{F}_{Y_1}(y(i)),\ \ldots,\ \hat{F}_{Y_p}(y_p(i)))}\right). \tag{49}$$

The source signal vector $\boldsymbol{s}(i)$, $i = 1, \ldots, N$, will be then estimated by

$$\hat{\boldsymbol{s}}(i) = \widehat{\boldsymbol{B}}\boldsymbol{x}(i), \quad i = 1, \ldots, N,$$

where

$$\widehat{\boldsymbol{B}} := \underset{\boldsymbol{B}}{\operatorname{arginf}}\ \widehat{\mathcal{J}^{\mathrm{dep}}}(\boldsymbol{B}),$$

which can be computed using a gradient descent type algorithm. In fact, straightforward computation shows that, the gradient in $\boldsymbol{B}$ of the estimated criterion $\boldsymbol{B} \mapsto \widehat{\mathcal{J}^{\mathrm{dep}}}(\boldsymbol{B})$, can be written as

$$\frac{d\,\widehat{\mathcal{J}^{\mathrm{dep}}}(\boldsymbol{B})}{d\boldsymbol{B}} = \frac{1}{N}\sum_{i=1}^{N}\frac{\frac{d}{d\boldsymbol{B}}\hat{c}_{\boldsymbol{Y}}(\boldsymbol{u}(i))}{\hat{c}_{\boldsymbol{Y}}(\boldsymbol{u}(i))}\frac{\hat{c}_{\hat{\theta}}(\boldsymbol{u}(i))}{\frac{d}{d\boldsymbol{B}}\hat{c}_{\hat{\theta}}(\boldsymbol{u}(i))} + \frac{\gamma}{N}\sum_{i=1}^{N}(\boldsymbol{y}(i) - \overline{\boldsymbol{y}}(i))(\boldsymbol{x}(i) - \overline{\boldsymbol{x}}(i))^{\top} + \frac{\mu}{N}\sum_{i=1}^{N}\left(\frac{\nabla \boldsymbol{y}(i)}{|\nabla \boldsymbol{y}(i)|}\right)^{*}\boldsymbol{x}(i)^{\top}, \tag{50}$$

where $\boldsymbol{u}(i) := (\widehat{F}_{Y_1}(y(i)), \ldots, \widehat{F}_{Y_p}(y_p(i)))$; the gradients $\frac{d}{d\boldsymbol{B}}\widehat{C}_{\boldsymbol{Y}}(\boldsymbol{u}(i))$ and $\frac{d}{d\boldsymbol{B}}\widehat{C}_{\hat{\theta}}(\boldsymbol{u}(i))$ can be explicitly computed, in a similar way as in Section 5.2.1. We then obtain the following Algorithm 3.

Algorithm 3. BSS algorithm for dependent source components.

Data: $\overline{\boldsymbol{x}}$ the observed signal
Result: $\hat{\boldsymbol{s}}$ the estimated source signal
Initialization: Compute $\boldsymbol{x} = \overline{\boldsymbol{x}} - \prod_{\lambda G}\overline{\boldsymbol{x}}$ from Algorithm 1,
$\boldsymbol{B}^{(0)} = \boldsymbol{I}_p$, $\boldsymbol{y}^{(0)} = \boldsymbol{B}^{(0)}\boldsymbol{x}$. Given $\varepsilon > 0$, $\nu > 0$.
Do: Update $\boldsymbol{B}$ and $\boldsymbol{y}$:

$$\boldsymbol{B}^{(q+1)} = \boldsymbol{B}^{(q)} - \nu\frac{d\ \widehat{\mathcal{J}^{\mathrm{dep}}}(\boldsymbol{B})}{d\boldsymbol{B}}$$

$$\boldsymbol{y}^{(q+1)} = \boldsymbol{B}^{(q+1)}\boldsymbol{x}.$$

Until $\| \boldsymbol{B}^{(q+1)} - \boldsymbol{B}^{(q)} \| < \epsilon$
$\hat{\boldsymbol{s}} = \boldsymbol{y}^{(q+1)}$.

6. Simulation results

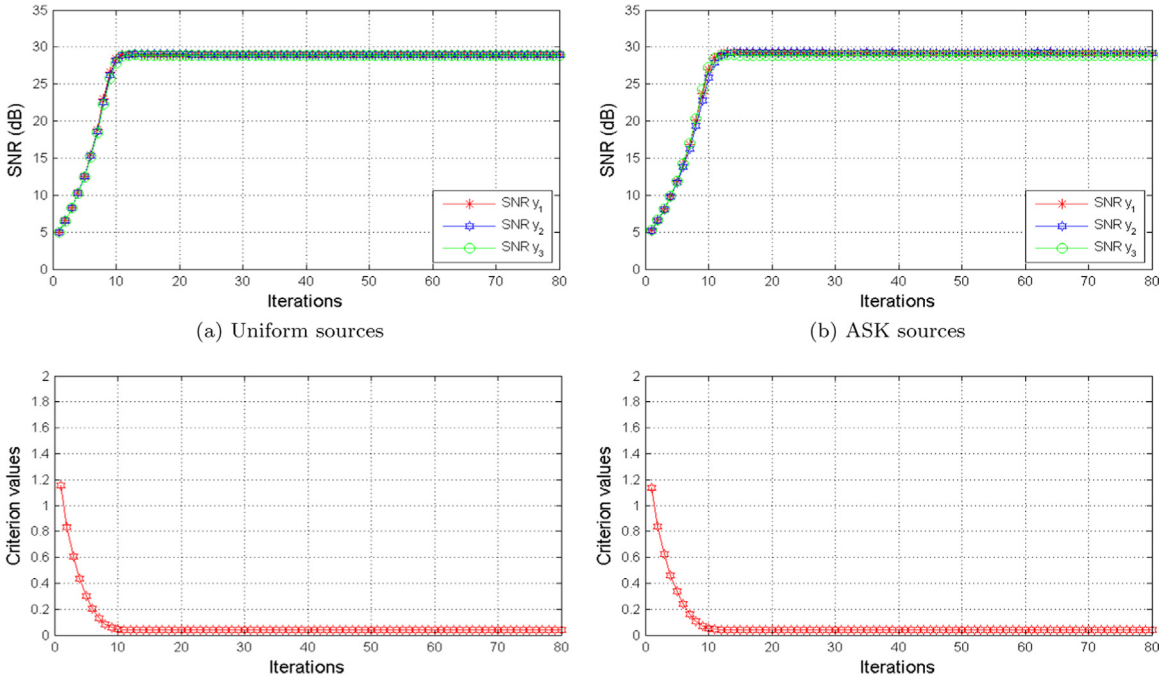
In this section, we present simulation results illustrating the performance of our proposed method. We will deal with the case of tree mixtures/tree sources, and consider many examples of independent/dependent source signal components. The simulation results will be compared with those obtained in [13], and those obtained by a TV-regularized versions of [33] (JADE) and [4] (FastICA), under the same conditions. Recall that the algorithm proposed in [13], for separating noisy mixtures of independent source signal components, is based on minimizing a kernel type estimate of mutual information between probability densities, penalized by the same TV-regularization term. In all examples, the number of samples is $N=2000$. The tree sources are mixed with the matrix $\boldsymbol{A}$:=[1 0.5 0.5; 0.5 1 0.5; 0.5 0.5 1]. A centered Gaussian noise with standard deviation 0.01 was added to the normalized mixtures, so that the signal-to-noise ratio (SNR) equals −20 dB. The gradient descent parameter is taken to be $\nu=0.1$. In the denoising step, see Algorithm 1, we take $\rho = 0.1$, $\lambda = 0.01$, and $\epsilon = 0.001$. In the second step, see Algorithm 2 or 3, we chose $\gamma = 0.001$ and $\mu = 0.01$. All simulations are repeated 100 times, and the accuracy of sources estimation is evaluated through the following signal-to-residue-ratio criterion, called again signal-to-noise-ratio, defined by

$$SNR_i := 10\log_{10}\frac{\sum_{n=1}^{N} s_i(n)^2}{\sum_{n=1}^{N}(\hat{s}_i(n) - s_i(n))^2}, \quad i = 1, 2, 3. \tag{51}$$

6.1. Examples where the source components are independent

We consider noisy mixed signals of two kinds of independent source components: uniform i.i.d on the interval [0, 1] in (Fig. 1a); i.i.d sources with components drawn from the 4-ASK (Amplitude Shift Keying) alphabets (−3, −1,1,3 with the same weights 0.25) in (Fig. 1b) at which was added i.i.d. realizations of centered Gaussian random variable with variance equals 0.25. We observe, from Fig. 1a and b, that our proposed method (Algorithm 2) gives accurate estimation of sources in the standard case of independent components, with performance around 30 dB in term of SNR. Fig. 1c and d shows that the criterion values (of $\widehat{\mathcal{J}^{\mathrm{ind}}}(\cdot)$) converge to 0 when the separation is achieved. In Fig. 2a and b, we compare Algorithm 2 with the MI-TV algorithm presented in [13], to separate the same instantaneous mixtures of independent source components. The results of our method are also compared with those of [33] (JADE) and [4] (FastICA), penalized by the same TV-regularization term. Table 2 gives the output SNR's obtained by each method for each source.

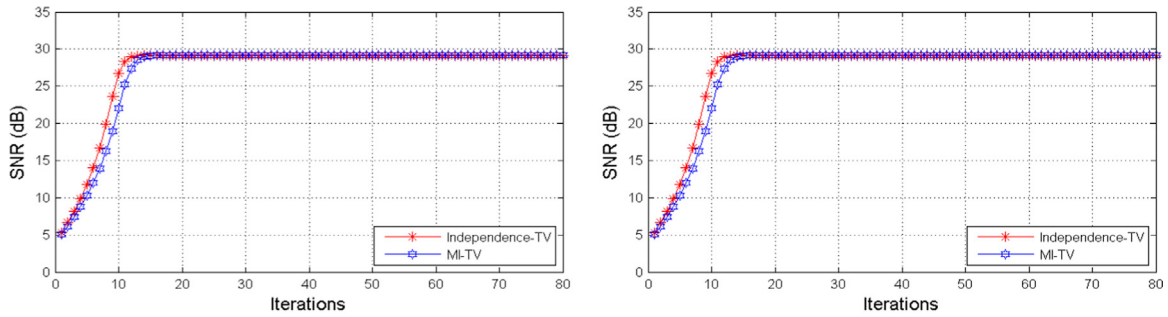
We can see that the approaches are equivalent, with small superiority of our method, in this standard case of independent components, as expected.



(a) Uniform sources (b) ASK sources

(c) Uniform sources (d) ASK sources

Fig. 1. SNRs and criterion values versus iteration number for independent source components.



(a) Uniform margins (b) ASK margins

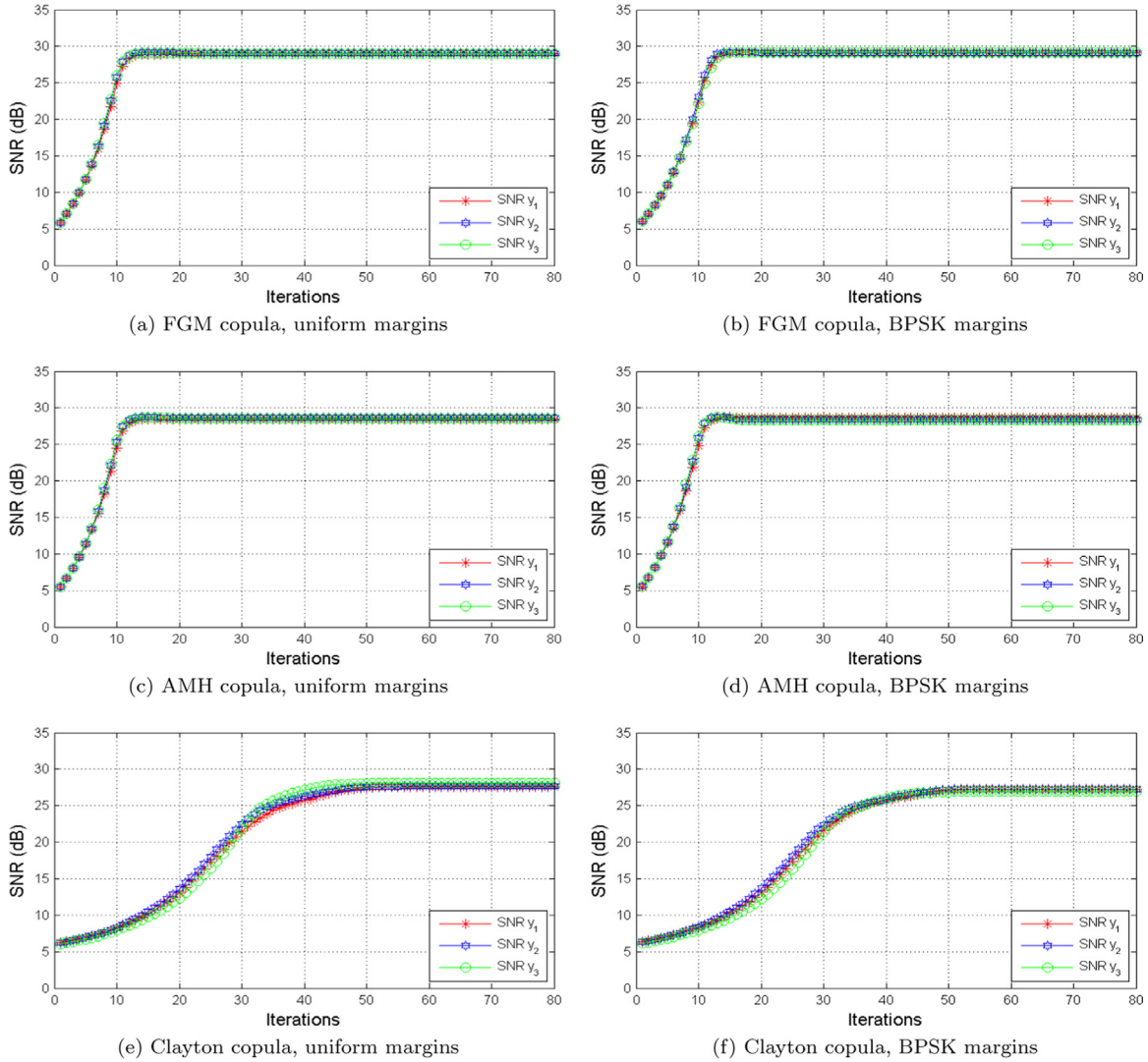
Fig. 2. SNRs versus iteration number for independent source components.

Table 2
Output SNR's.

Types	Sources	Our method	MI-TV	FastICA-TV	JADE-TV
Independent	S_1	28.9828	28.7778	28.8531	28.2541
uniform	S_2	29.0191	28.7478	27.9820	28.4857
sources	S_3	28.8840	28.7437	27.8602	28.2422
Independent	S_1	29.1942	28.5193	28.8531	28.0541
ASK	S_2	29.2374	28.8211	27.9820	28.2857
Sources	S_3	28.8734	28.7586	27.8602	28.0422

6.2. *Examples where the source components are dependent*

In this subsection, we show the capability of our proposed Algorithm 3 to successfully separate tree noisy mixtures of tree dependent source components. We dealt with noisy instantaneous mixtures of six kinds of sources: i.i.d. random source vectors (with uniform marginals in Fig. 3a, and binary phase-shift keying (BPSK) marginals in Fig. 3b), with dependent components generated from Fairlie–Gumbel–Morgenstern (FGM) copula model, where $\hat{\theta} = 0.9$; i.i.d. random source vectors (with uniform marginals in Fig. 3c, and



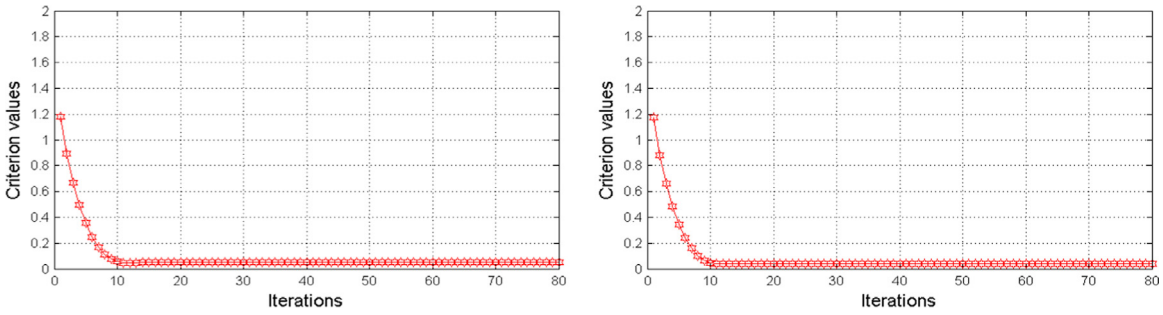
(a) FGM copula, uniform margins (b) FGM copula, BPSK margins

(c) AMH copula, uniform margins (d) AMH copula, BPSK margins

(e) Clayton copula, uniform margins (f) Clayton copula, BPSK margins

Fig. 3. SNRs values versus iteration number for dependent source components.

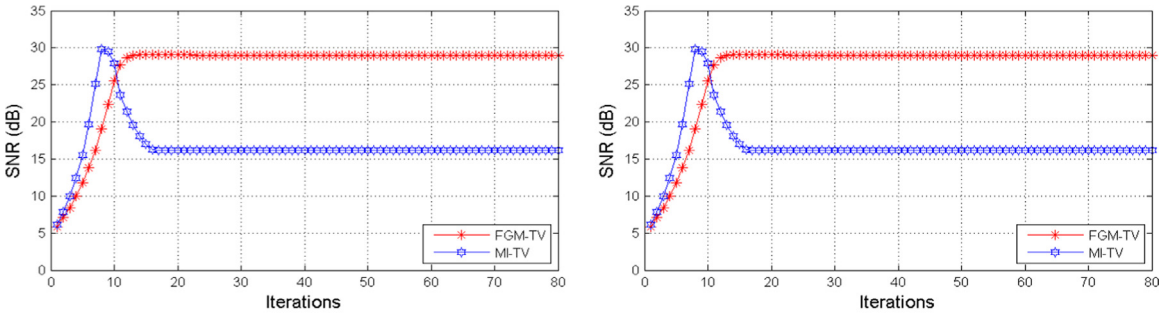
BPSK-marginals in Fig. 3d), with dependent components generated from Ali–Mikhail–Haq (AMH) copula model, where $\hat{\theta} = 0.75$; i.i.d. source vectors (with uniform marginals in Fig. 3e, and BPSK-marginals in Fig. 3f), where the components are dependent, generated from Clayton copula, with $\hat{\theta} = 1.4$. It can be seen from the results, in Fig. 3a–f, that the proposed method is able to separate mixtures of dependent source components, with good performance, equivalent to that of independent sources case, around 30 dB. Fig. 4a–d shows the convergence of the criterion values $\widehat{\mathcal{I}^{\text{dep}}}(\cdot)$ to 0 when the separation is achieved for the last example. We compare, through the following four examples of mixtures of dependent sources, the present approach (Algorithm 3) and the MI-TV of [13]. We consider i.i.d. vector sources (with uniform marginals in Fig. 5a and BPSK marginals in Fig. 5b), where the component are dependent, generated from FGM copula (with $\hat{\theta} = 0.9$). In Fig. 5c and d, the source components are dependent, generated from AMH copula ($\hat{\theta} = 0.75$), with uniform marginals (Fig. 5c), and from Clayton copula ($\hat{\theta} = 1.4$) with BPSK marginals in



(a) FGM copula, uniform margins (b) AMH copula, BPSK margins

(c) Clayton copula, uniform margins (d) Clayton copula, BPSK margins

Fig. 4. Criterion values versus iteration number for dependent source components.



(a) FGM copula, uniform margins (b) FGM copula, BPSK margins

(c) AMH copula, uniform margins (d) Clayton copula, BPSK margins

Fig. 5. SNRs versus iteration number for dependent source components.

Table 3
Output SNR's.

Types	Sources	Our method	MI-TV	FastICA-TV	JADE-TV
FGM	S_1	28.9305	15.0718	22.2844	14.4222
uniform	S_2	29.1057	15.1936	10.8547	14.0230
margins	S_3	29.0067	15.1124	9.6262	14.3052
FGM	S_1	29.1623	13.9554	26.4400	14.2968
BPSK	S_2	29.1161	14.1996	10.9654	13.5315
margins	S_3	29.1457	13.9930	9.8486	13.6773
AMH	S_1	28.4675	11.1730	18.7844	10.9222
uniform	S_2	28.6419	11.2599	7.3547	10.5230
margins	S_3	28.5433	11.2020	6.1262	10.8052
AMH	S_1	28.6486	11.1730	18.2844	11.1222
BPSK	S_2	28.4069	11.2599	8.8547	10.7230
margins	S_3	28.3088	11.2020	7.6262	11.0052
Clayton	S_1	27.6178	8.0160	20.2837	7.7411
uniform	S_2	27.6568	8.8441	4.9813	7.9196
margins	S_3	28.1373	8.6272	3.1241	8.1979
Clayton	S_1	27.6178	7.8098	21.8394	7.9147
BPSK	S_2	27.6568	8.4815	6.0928	7.8160
margins	S_3	28.1373	7.9975	4.1785	7.9973

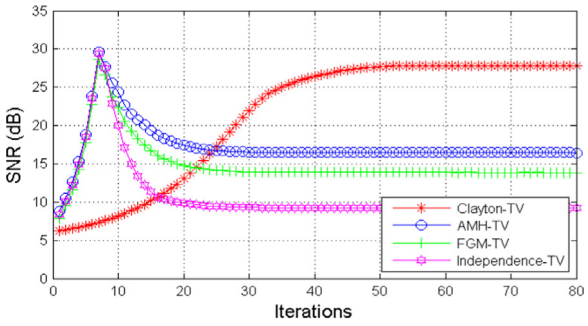


Fig. 6. SNRs versus iteration number for dependent source components.

Fig. 5d. Our results are also compared with those of [33] (JADE) and [4] (FastICA), penalized by the same TV-regularization term. Table 3 gives the output SNR's obtained by each method for each source.

We can see, clearly, from figures Fig. 4a–d and from Table 3, that separating mixtures of dependent source components, classical criterion devoted to the independent case, fails, and that our method is able to deal with the dependent case with good performance, if the copula modeling the dependency of the source components is correctly chosen. In Fig. 6, we present simulation results under misspecified model, that, when a model other than the correct one is used. We have considered noisy tree mixed signals of tree dependent source components, generated from Clayton copula model with $\theta = 1.5$, and BPSK margins. We can see, from the figure, that the use of an inappropriate model, leads to poor results, but better then MI-based BSS for independent sources.

7. Conclusion and discussion

We have presented a new BSS approach in the presence of noise, by minimizing a new regularized criterion. The approach is able to separate noisy instantaneous linear mixtures of independent/dependent source components. It proceeds in two steps: a first preprocessing step which consists in reducing the noise, and a second one minimizing an appropriate separation criterion based on Kullback–Leibler divergence between copula densities regularized by total variation term. The accuracy and the consistency of the obtained algorithms have been illustrated by simulation. It should be mentioned, however, that our proposed algorithms based on copula densities, rather than the classical ones based on probability densities, are more time consuming, since we estimate both copulas density of the vector and the marginal distribution function of each component. Moreover, the number of mixtures/sources should be small, since the accuracy estimation of the sources decreases when the number of mixtures/sources increases, caused by the curse of dimensionality in non-parametric multivariate density estimation. The present approach can be extended to deal with convolutive mixtures of independent/dependent sources, that will be addressed in future communications.

Acknowledgements

The authors wish to thank two anonymous reviewers for their constructive comments and criticisms leading to improvement of this paper.

References

[1] P. Comon, Independent component analysis, a new concept? Signal Process. 36 (3) (1994) 287–314, higher order statistics.
[2] A. Mansour, C. Jutten, A direct solution for blind separation of sources, IEEE Trans. Signal Process. 44 (3) (1996) 746–748.
[3] A. Taleb, C. Jutten, Entropy optimization, Artif. Neural Netw. ICANN'97 (1997) 529–534.
[4] A. Hyvärinen, E. Oja, A fast fixed-point algorithm for independent component analysis, Neural Comput. 9 (1997) 1483–1492.
[5] D. Pham, Blind separation of instantaneous mixture of sources based on order statistics, IEEE Trans. Signal Process. 48 (2) (2000) 363–375.
[6] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, E. Moulines, A blind source separation technique using second-order statistics, IEEE Trans. Signal Process. 45 (2) (1997) 434–444.
[7] J.-C. Pesquet, E. Moreau, Cumulant-based independence measures for linear mixtures, IEEE Trans. Inform. Theory 47 (5) (2001) 1947–1956, http://dx.doi.org/10.1109/18.930929.
[8] J.-F. Cardoso, Blind signal separation: statistical principles, Proc. IEEE 86 (10) (1998) 2009–2025.
[9] M. Novey, T. Adali, ICA by maximization of nongaussianity using complex functions, Proc. MLSP.
[10] D. Pham, Mutual information approach to blind separation of stationary sources, IEEE Trans. Inf. Theory 48 (7) (2002) 1935–1946.
[11] A. Keziou, H. Fenniri, M. Ould Mohamed, G. Delaunay, Séparations aveugle de sources par minimisation des α-divergences, in: XXIIe colloque GRETSI (traitement du signal et des images), Dijon (FRA), 8–11 septembre 2009, GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2009.
[12] P. Comon, C. Jutten, Handbook of Blind Source Separation: Independent Component Analysis and Applications, Academic Press, 2010.
[13] M. El Rhabi, H. Fenniri, A. Keziou, E. Moreau, A robust algorithm for convolutive blind source separation in presence of noise, Signal Process. 93 (4) (2013) 818–827.
[14] A. Keziou, H. Fenniri, A. Ghazdali, E. Moreau, New blind source separation method of independent/dependent sources, Signal Process. 104 (2014) 319–324.
[15] M. Sahmoudi, H. Snoussi, M.G. Amin, Robust approach for blind source separation in non-gaussian noise environments, Proceedings of ISCCSP, Marrakesh, Morocco, IEEE/EURASIP.
[16] A. Belouchrani, A. Cichocki, Robust whitening procedure in blind source separation context, Electron. Lett. 36 (24) (2000) 2050–2051.
[17] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris 8 (1959) 229–231.
[18] D.G. Clayton, A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, Biometrika 65 (1) (1978) 141–151.
[19] M.M. Ali, N.N. Mikhail, M.S. Haq, A class of bivariate distributions including the bivariate logistic, J. Multivar. Anal. 8 (3) (1978) 405–412.
[20] D. Morgenstern, Einfache Beispiele zweidimensionaler Verteilungen, Mitteilungsbl. Math. Statist. 8 (1956) 234–235.
[21] R.B. Nelsen, An introduction to copulas, Springer Series in Statistics, 2nd Edition, Springer, New York, 2006.
[22] H. Joe, Multivariate models and dependence concepts, Monographs on Statistics and Applied Probability, vol. 73, Chapman & Hall, London 1997 http://dx.doi.org/10.1201/b13150.
[23] H. Akaike, A new look at the statistical model identification, IEEE Trans.

Autom. Control 19 (6) (1974) 716–723.
[24] G. Schwarz, Estimating the dimension of a model, Ann. Stat. 6 (2) (1978) 461–464.
[25] C. Genest, K. Ghoudi, L.-P. Rivest, A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, Biometrika 82 (3) (1995) 543–552, http://dx.doi.org/10.1093/biomet/82.3.543.
[26] H. Tsukahara, Semiparametric estimation in copula models, Can. J. Stat. 33 (3) (2005) 357–375, http://dx.doi.org/10.1002/cjs.5540330304.
[27] L. Evans, R. Gariepy, Measure Theory and Fine Properties of Function, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL.
[28] A. Cohen, D. Wolfgang, I. Daubechies, R. DeVore, Harmonic analysis of the space BV, Rev. Mat. Iberoamericana 19 (1) (2003) 235–263.
[29] X. Chen, Y. Fan, Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification, J. Econometrics 135 (1–2) (2006) 125–154.
[30] A. Chambolle, An algorithm for total variation minimization and applications, J. Math. Imaging Vis. 20 (1–2) (2004) 89–97.
[31] M. Omelka, I. Gijbels, N. Veraverbeke, Improved kernel estimation of copulas: weak convergence and goodness-of-fit testing, Ann. Statist. 37 (5B) (2009) 3023–3058.
[32] B.W. Silverman, Density Estimation for Statistics and Data Analysis, in: Monographs on Statistics and Applied Probability, Chapman & Hall, London, 1986.
[33] J.-F. Cardoso, A. Souloumiac, Blind signal beamforming for non gaussian signals, Proc. IEEE 140 (6) (1993) 362–370.

B Dynamique des foules et modélisation numérique [1, 4]

Cette partie est dédiée à la présentation de nos résultats en modélisation mathématique et numérique de mouvement de foules (un article de revue soumis et un article accepté). Comme précédemment, pour chaque article, il est d'abord proposé un résumé en anglais, puis l'article dans son intégralité.

B.1 Estimating contact forces/pressure in a dense crowd from the non-smooth approach view for dynamics of solids [1]

This paper deals with the non-smooth dynamics of a rigid body system. The proposed theory is inspired by the formalism of J.J. Moreau and that of M. Frémond, and relies on the notion of percussion which is the integral of the contact force during the duration of the collision. Contrary to classical discrete element models, it is here assumed that percussions can be only expressed as a function of the velocity before the impact. This assumption is checked for the usual mechanical constitutive laws for collisions derived from a pseudo-potential of dissipation or the Coulomb friction law. Motion equations are then reformulated taking into account simultaneous collisions of solids. A mathematical study of the new model is presented : the existence and uniqueness of the solution are discussed according to the regularity of both the forces (Lebesgue-density occurring during the regular evolution of the system) and the percussions (Dirac-density describing the collision). In the light of the principles of thermodynamics, a condition on the internal percussion assuring that the collision is thermodynamically admissible, is established. Finally, an application of this new model to the motion of a system of rigid disks, including simultaneous collisions, is presented.

B.2 Discrete kinetic theory for 2D modeling of a moving crowd : Application to the evacuation of a non-connected bounded domain [4]

This paper concerns the mathematical modeling of the motion of a crowd in a non connected bounded domain, based on kinetic and stochastic game theories. The proposed model is a mesoscopic probabilistic approach that retains features obtained from both micro- and macro-scale representations ; pedestrian interactions with various obstacles being managed from a probabilistic perspective. A proof of the existence and uniqueness of the proposed mathematical model's solution is given for large times. A numerical resolution scheme based on the splitting method is implemented and then applied to crowd evacuation in a non connected bounded domain with one rectangular obstacle. The evacuation time of the room is then calculated by our technique, according to the dimensions and position of a square-shaped obstacle, and finally compared to the time obtained by a deterministic approach by means of randomly varying some of its parameters.

Estimating contact forces/pressure in a dense crowd from the non-smooth approach view for dynamics of solids

A.Jebrane[a,e], P.Argoul[b,d], M. El Rhabi[a,c], A.Hakim[a]

[a] *LAMAI, FST Marrakech, Université Cadi Ayyad , Maroc*
[b] *Université Paris-Est , LVMT (UMR_T 9403), Ecole des Ponts ParisTech, IFSTTAR, UPEMLV, F-77455 Marne la Vallée, France*
[c] *IMI, Ecole des Ponts ParisTech, Université Paris Est, France*
[d] *Université Paris-Est, Laboratoire EMGCU-MAST, IFSTTAR, Marne-la-Vallée, France*
[e] *Laboratoire Navier (UMR 8205), CNRS, ENPC, IFSTTAR, Université Paris-Est, 77455, Marne-la-Vallée, France*

Abstract

This paper deals with estimation of pressure during the movement of a dense crowd. Based on the non-smooth approach of contact dynamics both, for rigid and deformable solids, proposed by M.Frémond and his collaborators, we estimate the pressure/ contacts force generated through congestion or a panic situation. Firstly, we propose a second-order microscopic model, in which the crowd is treated as a system of rigid solids (discrete medium). Contact forces are rigorously defined taking into account multiple, simultaneous contact and the non overlapping condition between pedestrians. We show that for a dense crowd the percussions (momentum jump corresponding to instantaneous contact) become contact forces. Secondly, a continuous equivalent approach is proposed where the crowd is modeled as a deformable solid (continuous medium), and the pressure is calculated according to volume and surface constraints. This approach makes it possible to retain an admissible right-velocity (after impact), including both the non local interactions (at a distance interactions) between

*Corresponding author: A.Jebrane
Email addresses: aissam.jebrane@enpc.fr (A.Jebrane), pierre.argoul@ifsttar.fr (P.Argoul), mohammed.el-rhabi@enpc.fr (M. El Rhabi), abdelilah.hakim@gmail.com (A.Hakim)

non neighbor pedestrians and the choice of displacement strategy of each pedestrian.

Finally, two applications are presented: a one-dimensional simulation of an aligned pedestrian chain crashing into an obstacle, and a two-dimensional simulation corresponding to the evacuation of a room.

Keywords: `Crowd modeling, Non-local interactions, Contact forces/Pressure, Panic situation, Non-smooth dynamics, Simultaneous collisions`

1. Introduction

The modeling of crowd dynamics has been deeply studied in the 25 past years, in particular, approaches drawn from granular modeling [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16] offer both a rich mathematical and physical framework to simulate and comprehend the dynamics of a crowd.

Predicting and anticipating the different scenarios that can occur when a crowd moves is a challenge that must be taken into account for infrastructure planning (airport, concert hall, station ...), it is a necessity to avoid the tragedy dramas in a dense crowd in panic, and an essential tool to achieve this goal is the prediction of the pressure and the contact forces generated in a moving crowd.

At first, we present a brief review of pressure estimation or contact forces techniques existing (in the knowledge of the author). We distinguish two major approaches that have dealt with this issue:

- A microscopic approach proposed in [10, 11, 13] where pedestrians are identified with rigid disks that must obey the non-overlapping constraint. The right- velocity (after collision) is the projection of the left-velocity (corresponding to the desired velocity) over the set of admissible velocities (with regard to the non-overlapping constraint). This constrained optimization problem is rewritten in a saddle-point form, and the corresponding Lagrange multipliers are interpreted as the pressure between pedestrians in contact. In this first-order microscopic model (the un-

known is the motion q), pedestrian mass is not taken into account and the dimension of the Lagrange multipliers is incompatible with that of a pressure. In addition, it is impossible to estimate the elementary percussions between pedestrians because of the non-uniqueness of the Lagrange multipliers.

To overcome theses problems, we propose a second-order model mainly developed by M.Frémond [17, 18, 19, 20, 21] for the modeling of granular media. The application of this approach for a crowd has been the subject of our previous work [2, 4, 5, 6, 7, 8, 9, 22]: simulation of panic situations, reproducing alignment phenomenon at the meeting of two groups, contact management, interaction between pedestrians and structures and management of pedestrian traffic in train stations. By adopting this approach, we will rigorously define the contact forces. A comparison with the first-order microscopic model will be presented.
We cite also the social force model [23, 24, 25, 26, 27], in which repulsive forces can be considered as contact forces. This force is introduced primarily as a penalty term to keep a certain distance and consequently prevent inter-penetration between pedestrians. The presence of many modeling parameters makes it difficult to estimate contact forces with this approach.

- Macroscopic approaches: On the other hand, macroscopic models have been developed, describing the evolution of the crowd as a continuum medium , we cite between them, a classical fluid approach [28] where the motion of crowd and the generated pressure are described by Navier-Stokes equations. In general, the question of estimating of pressure in a moving crowd is not addressed in most of the continuous approaches [30, 31, 32, 33].

  Similarly to the first-order microscopic model, where pressure is discussed, an equivalent macroscopic approach which aims to treat the pressures is proposed in [11, 15], the pressure is the solution of the continuous Laplace problem with a second positive member and a homogeneous Dirichlet con-

dition on the boundary of the saturated zone (with regard to the density). The model makes it possible to take into account the non-local character of the pressure, however, it has been shown [15] that this model in its current version reproduces a contradictory effect to that observed experimentally; congestion accelerates the evacuation of a convex room through a single door. Note also that several approaches that have recently developed like [34] and kinetic model [35, 36, 37, 38] have not yet dealt pressure issue.

Contribution and paper content:

In this work, we devote interest to the estimation of contact forces/ pressure, in a moving dense crowd, the theory is inspired from the non-smooth approach developed by M.Frémond and his collaborators both for the evolution of rigid and deformable bodies. To fix the ideas, we consider the configuration of a dense crowd at a given moment (ie a stationary problem), and we ask how to estimate the contact forces or the generated pressure. Therefore, two classes of models will be developed: In section 2, a microscopic second-order model, in which the crowd is identified with a system of rigid bodies is presented. In previous works [6, 7], we have evoked this question, and we have arrived at an estimation of the percussion which is the integral of the contact force during the duration of the collision. Now, we propose to define the contact forces rigorously and with regard to physical proprieties. We start with the presentation of a brief overview on the mathematical and physical framework of the non-smooth dynamics approach, particularly, the collision equation, then, we show that in the case of a dense crowd where the contacts remain maintained that percussions (jumps of the momentum) qualify as contact forces, ie, they will have the dimension of a force. A one-dimensional simulation corresponding to a pedestrians chain lining up, and a two-dimensional simulation representing the evacuation of a room are presented, including the interpretation and study of the influence of dissipation parameter involved in the model.

Since we are interested in the case of a dense crowd, it is quite normal to con-

sider the crowd as a continuous medium, which is the object of section 4 where we present a first continuous version of the second-order model. This approach makes it possible to take into account the deformability property by modeling a dense crowd as a deformable solid. In a panic situation (for example, congestion during evacuation of a room), an arch formation is observed, the domain occupied by the crowd is considered as a moving deformable solid with the boundaries in contact with obstacles (walls...) subjected to surface constraints to avoid inter-penetration with obstacles, and volume constraints defining the crowd deformation. Using constitutive laws, pressure is deduced from the volume and surface constraints. We perform simulations similar to those of the discrete case. Generally, a coherence between the two approaches is seen. Section 5 is devoted to the discussion of the results obtained and the perspectives that have emerged.

2. A second-order microscopic approach: Crowd is modeled to a system of rigid bodies

2.1. An overview of mathematical and physical framework: The non-smooth view of contact dynamics by M.Frémond

We consider a crowd as a system of rigid bodies, using the theory of non-smooth contact dynamics by M.Frémond, we obtain the contact equation. This approach makes it possible to manage multiple and simultaneous contact and to take into account the corresponding local and non-local (at a distance) interactions. Moreover, the capacity to use different constitutive laws make it possible to account for diverse experimental results. The first element of the theory is to identify and quantify the deformation of the system. In the following, we retain a two-dimensional representation, ie, a pedestrian p^i is modeled as a rigid disk of mass m^i, inertia I^i, center of mass $\mathbf{q}^i$ and velocity $\mathbf{u}^i$.

We note $\mathbf{u} = \left(\mathbf{u}^1, ..., \mathbf{u}^j, ..., \mathbf{u}^{n_p}\right)$, the vectorial function formed by velocities $\mathbf{u}^i$, for i varying from 1 to n_p, the number of pedestrian, and $\mathbf{q} = \left(\mathbf{q}^1, ..., \mathbf{q}^j, ..., \mathbf{q}^{n_p}\right)^T$

the vectorial function defines the motion of a pedestrian by:

$$\mathbf{q}(t) = \mathbf{q}(0) + \int_0^t \mathbf{u}(\tau)d\tau. \tag{1}$$

$\mathbf{q}$ belongs to $\mathcal{C}(I, \mathbf{E})$, $I = [0, T], T > 0$, $\mathbb{E} = \mathbb{R}^n$, $n = n_p \times n_d$, where n_d the degrees of freedom, here $n_d = 3$. $\mathbf{u}$ is a function of bounded variation and continuous to the right: $\mathbf{u} \in \tilde{\mathbb{BV}}(I, \mathbb{E}) = \{\mathbf{u} \in \mathbb{BV}(I, \mathbf{E}), \mathbf{u}(t)$ Continuous to the right$\}$, this allows us to include the discontinuity of velocity.
The problem describing simultaneous and multiple contact is :

$$\mathbf{M}d\mathbf{u} = d\mathbf{p}, \tag{2}$$

$$\Rightarrow \mathbf{M}(\mathbf{u}^+ - \mathbf{u}^-) = -\mathbf{p}^{int} + \mathbf{p}^{ext}, \tag{3}$$

Where,

- $\mathbf{M}$ is the inertia matrix.
- $d\mathbf{u}$ is the differential measure (the measure of Stieltjes) associated to the velocity $\mathbf{u}$.
- $d\mathbf{p}$ is the differential measure defined by : $d\mathbf{p}\{t\} = \mathbf{p}(t)\delta(t - t_c), where t_c$ is the instant of contact. $\mathbf{p}$ is the density function with respect to the Dirac measure, and represents physically the jump of the momentum, as well as quantifying the percussion corresponding to a contact. $\mathbf{p}^{int}$ (respectively $\mathbf{p}^{ext}$) are internal (respectively external) percussions.

The internal percussion $\mathbf{p}^{int}$ is defined according to the laws of thermodynamics and takes into account the condition of non-overlapping. The existence and uniqueness of solutions of the contact equation are discussed in [17, 18, 39, 40] with regard to chosen constitutive law. In our case, internal percussions derive from a pseudo potential.

Let's introduce the velocity of deformation $D(\mathbf{u})$ corresponding to contact of two pedestrians p^i and p^j at point A is given by:

$$D^{i,j}(\mathbf{u}) = \mathbf{u}^i(A) - \mathbf{u}^j(A) \tag{4}$$

In 2D:

$$D^{i,j}(\mathbf{u}) = \mathbf{u}^i(\mathbf{q}^i) + w^i \times G^i A - \mathbf{u}^j(\mathbf{q}^j) - w^j \times G^j A, \tag{5}$$

where w^i is the rotation velocity of pedestrian p^i, and $\mathbf{D}(\mathbf{u}) = \left\{\mathbf{D}^{i,j}(\mathbf{u}), 1 \leq i < j \leq n_p\right\}$ is a column vector of size n_c, where n_c is number of contact.

Local/non local interactions are introduced by percussion $\mathbf{p}^{i,j}, 1 \leq i < j \leq n_p$, with:

$$\mathbf{p}^{i,j} = \mathbf{p}^{i,j}_{reac} + \mathbf{p}^{i,j}_{diss}. \tag{6}$$

The percussion of reaction $\mathbf{p}^{i,j}_{reac}$ is defined by:

$$\mathbf{p}^{i,j}_{reac} \in \partial \mathbb{1}_{\mathbb{R}^+} \left(D^{i,j}(\mathbf{u}^+).\mathbf{n}^{i,j}_q\right), \tag{7}$$

where $\mathbf{n}^{i,j}_q$ is the normal oriented from p^i to p^j, and $\mathbb{1}_{\mathbb{R}+}(x) = 0 \quad if \quad x \geq 0, +\infty$ otherwise.

The percussion of dissipation $\mathbf{p}^{i,j}_{diss}$ is defined by:

$$\mathbf{p}^{i,j}_{diss} \in \partial \phi^{i,j} \left(\frac{D^{i,j}(\mathbf{u}^+ + \mathbf{u}^-)}{2}\right) \tag{8}$$

where $\phi^{i,j}$ is a pseudo-potential of dissipation. We choose a quadratic form, ie

$$\phi^{i,j}(\mathbf{x}) = k\left(\mathbf{x}\right)^2, \tag{9}$$

k is a coefficient of dissipation. This choice allows one to find the classical results when the coefficient of restitution is used.

Introducing the set $\mathcal{U}$ of admissible velocities:

$$\mathcal{U}_q = \left\{\mathbf{u} \in \mathbf{R}^{2n_p}, D^{i,j}(\mathbf{u}).\mathbf{n}^{i,j}_q \leq 0, 1 \leq i < j \leq n_p\right\}, \tag{10}$$

and the normal cone $\mathbf{N}_{\mathcal{U}_q}(\mathbf{u})$ of $\mathcal{U}_q$ at point $\mathbf{u}$, the non-overlapping condition is expressed by:

$$\mathbf{p}_{reac} \in \partial \mathbb{1}_{\mathcal{U}}\left(\mathbf{u}^+\right) = \mathbf{N}_{\mathcal{U}}(\mathbf{u}^+) \tag{11}$$

where $\mathbf{p}_{reac} = \left\{\mathbf{p}^{i,j}_{reac}, 1 \leq i < j \leq n_p\right\}$.

Finally, the collision equation is :

$$\mathbf{M}\left(\mathbf{u}^+ - \mathbf{u}^-\right) + \partial\phi\left(\mathbf{D}\frac{(\mathbf{u}^+ + \mathbf{u}^-)}{2}\right) + \mathbf{N}_{\mathcal{U}_q}(\mathbf{u}^+) \ni 0, \tag{12}$$

where ϕ regroups the elementary pseudo-potential of dissipation $\phi^{i,j}$ so that $\phi(\mathbf{D}(\mathbf{x})) = k(\mathbf{D}(\mathbf{x}))^2$.

2.2. Characterization of the percussions and contact forces

Wes consider Eq.12, for:

$$\bar{\mathbf{x}} = \frac{\mathbf{u}^+ + \mathbf{u}^-}{2}, \tag{13}$$

$$\tilde{\mathcal{U}}_q = \left\{\mathbf{x} \in \mathbf{R}^{2n_p}, D^{i,j}(\mathbf{x}).\mathbf{n}_q^{i,j} \leq D^{i,j}(\mathbf{u}^-).\mathbf{n}_q^{i,j}, 1 \leq i < j \leq n_p\right\}, \tag{14}$$

$$= \left\{\mathbf{x} \in \mathbf{R}^{2n_p}, \mathbf{C}\mathbf{x} \leq \frac{1}{2}\mathbf{C}\mathbf{u}^-, \right\}, \tag{15}$$

where $\mathbf{C}$ the matrix defining the constraints on the normal deformation velocities.

$$\text{and} \quad f(\mathbf{x}) = \mathbf{x}^T M\mathbf{x} - 2\mathbf{x}^T M\mathbf{u}^- + \phi\left(\mathbf{D}(\mathbf{x})\right). \tag{16}$$

$\mathbf{u}^+$ the solution of Eq.12 is equivalent to

$$\bar{\mathbf{x}} = Argmin_{\mathbf{x}\in\tilde{\mathcal{U}}} f(\mathbf{x}), \tag{17}$$

which has a unique solution [17].
We introduce the associated Lagrangian $\mathcal{L}$ of the minimization problem

$$\mathcal{L}(x, \lambda) = f(\mathbf{x}) + \lambda.\left(\mathbf{C}.\left(\mathbf{x} - \frac{\mathbf{u}^-}{\mathbf{2}}\right)\right), \tag{18}$$

where $\lambda = \left\{\lambda^1, ..., \lambda^{nc}\right\} \in (\mathbb{R}^+)^{nc}$, where n_c number of contact.
The solution $\mathbf{u}^+$ and Lagrange multipliers λ verifying:

$$\begin{cases} \mathbf{M}\left(\mathbf{u}^+ - \mathbf{u}^-\right) + k\mathbf{D}^T\mathbf{D}\left(\mathbf{u}^+ + \mathbf{u}^-\right) + \mathbf{C}^T\lambda = 0 \\ \mathbf{C}\mathbf{u}^+.\lambda = 0 \end{cases} \tag{19}$$

For a dense crowd, contacts are maintained, we characterize this by the fact that the constraints are saturated, ie:

$$\mathbf{C}\mathbf{u}^+ = 0, \tag{20}$$

which means that contact is purely inelastic, and there is no dissipation ($k = 0$). Consequently, λ is a solution of:

$$\mathbf{W}\lambda = \mathbf{C}\mathbf{u}^{-}, \quad \mathbf{W} = \mathbf{C}\mathbf{M}^{-1}\mathbf{C}^{T}. \tag{21}$$

$\mathbf{W}$ is the Delassus matrix.

The first-order microscopic model [10, 11] does not take into account the mass of pedestrians, λ is solution of the discrete Poisson's equation,

$$\mathbf{C}\mathbf{C}^{T}\lambda = \mathbf{C}\mathbf{u}^{-}, \tag{22}$$

where $\mathbf{C}\mathbf{C}^{T}$ is considered as the discrete Laplacian matrix. It is a special case of the developed approach where the matrix of inertia equals the identity matrix. They are implicitly assumed to have the same mass of $1kg$ and represented by disks of the same radius, this leads to the non-uniqueness of the Lagrange multipliers for equilibrium situations and produce particular scenarios that are specially related to the geometry of pedestrians (disks) and their size (radius of disks) such as the complete congestion of pedestrians during the evacuation of a room. As shown in equation Eq.21, neglecting the influence of pedestrian masses is a restrictive hypothesis, an analytical solution in the case of an aligned pedestrian chain shows the dependence on the mass and the Lagrange multipliers. The comparison between an aligned pedestrians chain of identical masses and an aligned pedestrian chain of different masses is presented in Figures 1 and 2 .

According to the collision equation Eq.(12), the percussion has the dimension of momentum, and hence it is perfectly normal for instant contact for an instantaneous collision. For a dense crowd (this is the case that interests us), this does not remain true, indeed: the contact is maintained and the duration of contact is important, because of these considerations, the motion of the crowd is considered smooth and Eq.(12) can be rewritten as a second-order spewing process:

$$-\mathbf{M}\frac{d\mathbf{u}}{dt} \in \mathbf{N}_{\tilde{\mathcal{U}}_q(t_c)}(\mathbf{u}(t_c)), \tag{23}$$

where $t \longrightarrow \tilde{\mathcal{U}}_{q(t)}$ is the moving set of admissible velocities. The contact forces $\mathbf{f}_{contact}$ correspond to a selector in the outward normal cone. ie:

$$-\mathbf{M}\frac{d\mathbf{u}}{dt} = \mathbf{f}_{contact}(t_c), \text{where} \mathbf{f}_{contact}(t_c) \in \mathbf{N}_{\mathcal{V}_{ad}(t_c)}(\mathbf{u}(t_c)). \tag{24}$$

The existence and uniqueness of the solution of Eq.23 is widely discussed in [41, 42, 43, 44] both for smooth and non-smooth (ie for a measure differential inclusion) motion. In this formalism the acceleration has a sens and consequently the contact force $\mathbf{f}_{contact}$ which approached numerically by $\mathbf{C}^T\lambda$.

3. Numerical simulations

3.1. 1 D Simulation: Pedestrian chain lining up

A first simulation is presented in Fig. 1 simulating the Lagrange multipliers λ corresponding to contacts in a chain of 20 pedestrians who stand in line. We point out that the columns of the constraint matrix are unit vectors hence λ corresponds to the generated contact force modules. The first example Fig. 1a presents jostling in a free queue, ie; in an unconstrained pedestrian chain. The second example Fig. 1b illustrates crushes facing an obstacle, ie; for a constrained pedestrian chain.

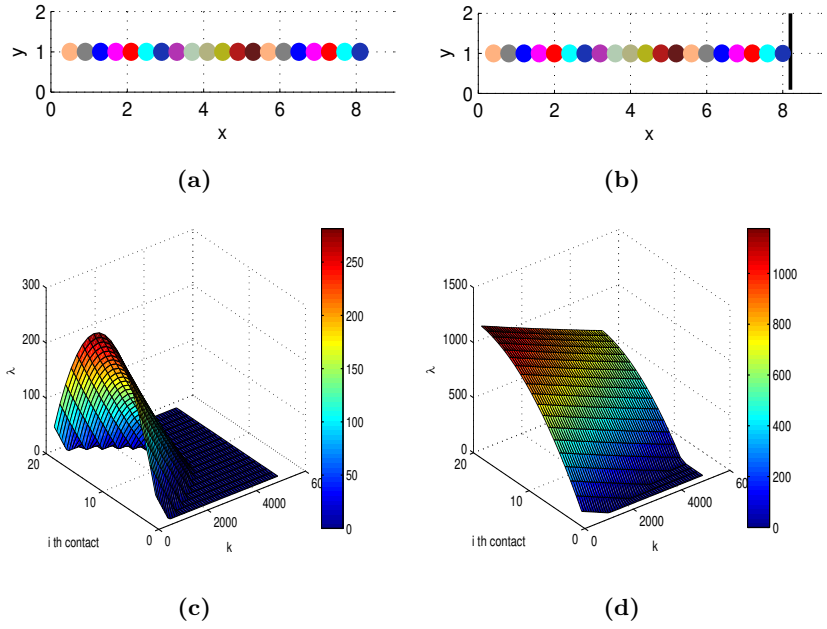


Figure 1: Lagrange multipliers $\lambda^{i,j}$ (corresponding to the contact force $\mathbf{f}^{i,j}_{contact}$ between pedestrians i and j) as a function of the dissipation parameter k for 20 pedestrians of identical masses $m = 75kg$ and increasing left-velocities (non-admissible). (a) A free pedestrian chain, (ie there is no obstacle ahead), (c) the associated Lagrange multipliers. (b) A constrained pedestrian chain, (ie there is an obstacle ahead (black bar)), (d) the associated Lagrange multipliers.

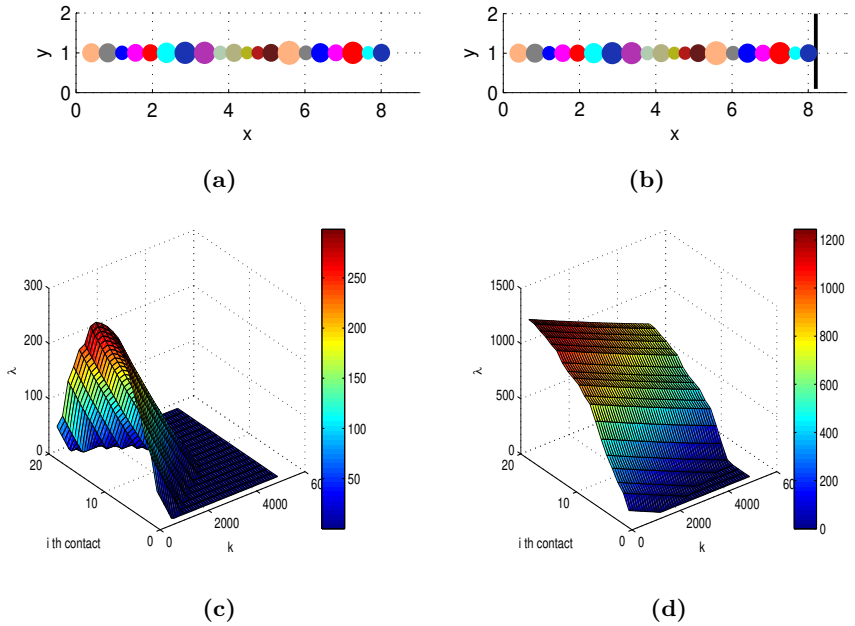


Figure 2: Lagrange multipliers $\lambda^{i,j}$ as a function of the dissipation parameter k for 20 pedestrians of different masses. (a) A free pedestrian chain, (c) the associated Lagrange multipliers. (b)A constrained pedestrian chain, (d) the associated Lagrange multipliers.

Numerically, the right-velocities and Lagrange multipliers are given by Eq.19, the values for a dense crowd are given for $k = 0$ where the contact forces become maximal as shown in Fig.1 and Fig.2 . In case of a constrained pedestrians chain of identical masses, the percussion grows with proximity to the obstacle . For a free pedestrian chain, the pressure is maximum in the middle of the chain. We recall that the left-velocity is decreasing from left to right. In the case of an unconstrained pedestrian chain of different mass, we notice the same behavior, except that the Lagrange multipliers curves are not smooth, due to the influence of the masses.

3.2. The dependence mass-Lagrange multipliers

In this subsection we explain analytically the Lagrange multipliers, a computation in the unidimensional case for an unconstrained pedestrian chain with purely inelastic contacts brings us to the following recurrence relation:

$$\begin{cases} \lambda^{1,2} = -m^1(u - u^{1,-}), \\ \lambda^{i,i+1} = \lambda^{i-1,i} - m^i(u - u^{i,-}), \quad 2 \leq i < n_p. \end{cases} \tag{25}$$

and

$$u^{i,+} = u = \frac{\sum_{i=1}^{n_p} m^i u^{i,-}}{\sum_{i=1}^{n_p} m^i}, 1 \leq i \leq n_p, \tag{26}$$

which shows both the dependence between masses, velocities and Lagrange multipliers .

As a validation test of the result of Fig.1 and Fig2, Fig.3 illustrates a comparison between the analytical and numerical solution.

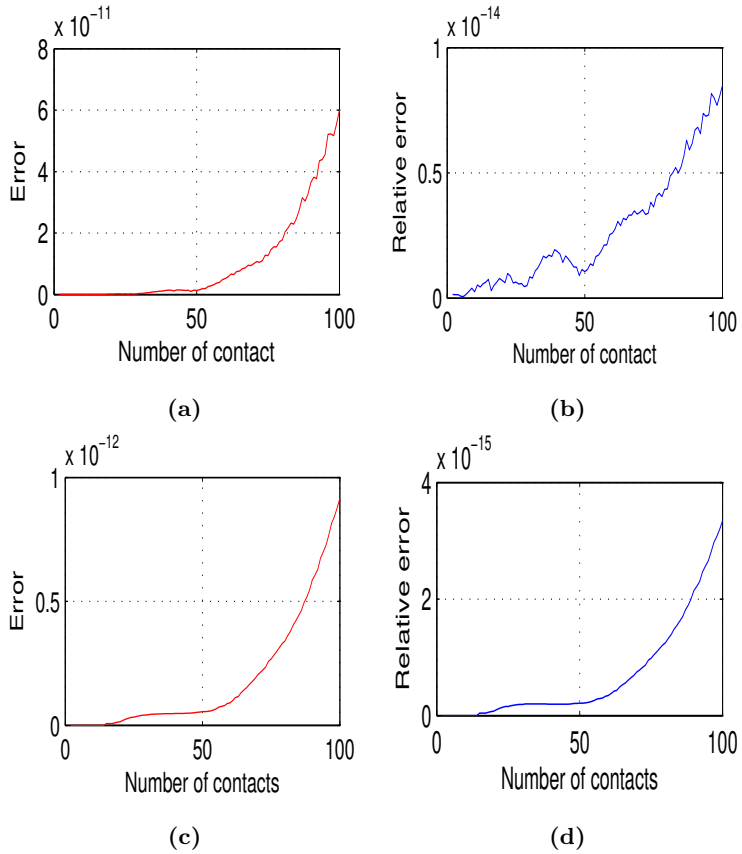


Figure 3: Error between the analytical and numerical solution. (a-b) for Lagrange multipliers, (c-d) for right velocities.

.

3.3. Non-local (at a distance) interactions

The non-interpenetration condition results in non-local interactions. To illustrate this,we consider the one-dimensional problem of simultaneous inelastic collision (the dissipation percussion is zero) of three aligned pedestrians Fig. 4.

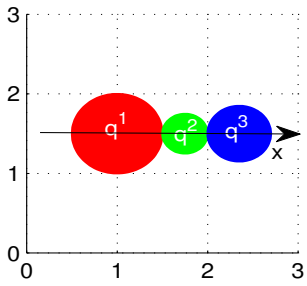


Figure 4: Multiple impact pedestrians chains.

After a simple calculation, the right velocities of the pedestrians are given by:

$$\mathbf{u}^{1,+} = \mathbf{u}^{2,+} = \mathbf{u}^{3,+} = \frac{m^1\mathbf{u}^{1,-} + m^2\mathbf{u}^{2,-} + m^3\mathbf{u}^{3,-}}{m^1 + m^2 + m^3}, \tag{27}$$

and

$$\mathbf{p}^{1,2}_{reac} = -m^1 \frac{m^2(\mathbf{u}^{2,-} - \mathbf{u}^{1,-}) + m^3(\mathbf{u}^{3,-} - \mathbf{u}^{1,-})}{m^1 + m^2 + m^3}, \tag{28}$$

$$\mathbf{p}^{2,3}_{reac} = -m^3 \frac{m^1(\mathbf{u}^{1,-} - \mathbf{u}^{3,-}) + m^2(\mathbf{u}^{2,-} - \mathbf{u}^{3,-})}{m^1 + m^2 + m^3}. \tag{29}$$

From the expressions of $\mathbf{p}^{1,2}_{reac}$ and $\mathbf{p}^{2,3}_{reac}$, it is clearly seen that the condition of non-interpenetration (local constraint on velocity) results in distance interactions in terms of percussion, in fact; the percussion $\mathbf{p}^{1,2}_{reac}$ involves $\mathbf{u}^{3,-}$ besides $\mathbf{u}^{1,-}$ and $\mathbf{u}^{2,-}$, the same for $\mathbf{p}^{2,3}_{reac}$ which involves $\mathbf{u}^{1,-}$ besides $\mathbf{u}^{2,-}$ and $\mathbf{u}^{3,-}$. An explicit at a distance interaction (non-local interaction) can be taken into account by introducing a non-local velocity of deformation $\tilde{D}^{1,3}(\mathbf{u})$ between pedestrians 1 and 3 which are not in contact:

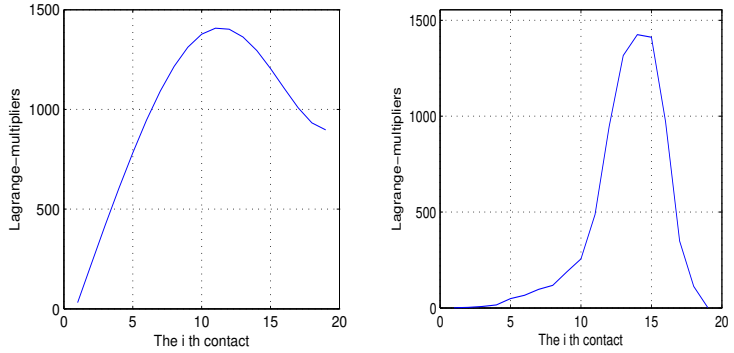
$$\tilde{D}^{1,3}(\mathbf{u}) = 2\left(\mathbf{u}^1 - \mathbf{u}^3\right).\left(\mathbf{q}^1 - \mathbf{q}^3\right), \tag{30}$$

which is the derivative with respect to time of the squared distance between points $\mathbf{q}^1$ and $\mathbf{q}^3$. This deformation can be expressed as a function of the local deformation velocities and can be reduced to the same initial definition.

3.4. Influence and interpretation of the dissipation coefficient k

Fig5 shows Lagrange multipliers as a function of the dissipation coefficient k for a constrained pedestrian chain Fig.5. First 10 left-hand collisions have for $k = 10$ while the following 9 collisions have for $k = 10000$. transmission of contact forces is degraded from left to right instead of being accumulated.

The dissipation coefficient k can identify the situation of the crowd, indeed: for very large k values, contact forces between pedestrians are weak, this corresponds to the case of normal circulation. We obtain the opposite in the case where k tends to zero, ie the contact forces become important, this corresponds to the situation of panic.



(a) **(b)**

Figure 5: Influence of the dissipation parameter k. (a) Pedestrians have the same masses m = 75 kg, (b) pedestrians have different masses of means m = 75 kg and standard deviation 5 kg.

For example, in case of a binary contact of two pedestrians p^i and p^j of mass m^i and m^j and left-velocity $u^{i,-}$ and $u^{j,-}$ respectively. We can identify a critical value k^c of the parameter of dissipation from which the contact force is

canceled. Using Eqs.19, it is easy to establish that:

$$k \leq k^c = \frac{m^i m^j}{m^i + m^j} \Longleftrightarrow \begin{cases} p^{i,j}_{reac} = \frac{m^i m^j - k(m^i+m^j)}{m^i+m^j}(u^{i,-} - u^{j,-}), \\ p^{i,j}_{diss} = k(u^{i,-} - u^{j,-}), \end{cases} \tag{31}$$

and

$$k > k^c = \frac{m^i m^j}{m^i + m^j} \Longleftrightarrow \begin{cases} p^{i,j}_{reac} = 0, \\ p^{i,j}_{diss} = 2\frac{k m^i m^j}{m^i m^j + k(m^i+m^j)}(u^{i,-} - u^{j,-}). \end{cases} \tag{32}$$

One way to illustrate this result is to solve the equation:

$$\mathbf{M}\left(\mathbf{u}^+ - \mathbf{u}^-\right) + k\mathbf{D}^T\mathbf{D}\left(\mathbf{u}^+ + \mathbf{u}^-\right), \tag{33}$$

then trace (Fig.6) the vector $\mathbf{Cu}^+$ (which should be negative, ie,the right-velocity be admissible) according to different values of the parameter k. Left-velocity $\mathbf{u}^-$ is chosen to be inadmissible.

Values of ($k <= k^c$) corresponds to a panic phase (a reaction percussion is necessary to make the configuration admissible), the normal circulation phase corresponding to the values of ($k >= k^c$)(*brownline*). A classical result of the M.Frémond collision model is found: $k = 0$ is a purely inelastic contact and $k >> k^c$ is an elastic contact.

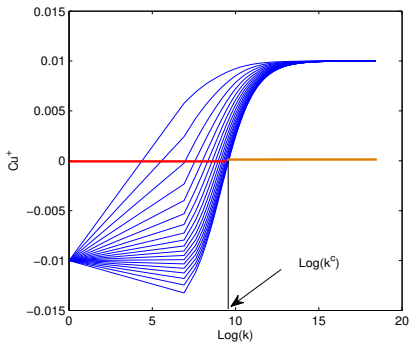


Figure 6: Variation of the vector $\mathbf{Cu}^+$ according to parameter k for the constrained pedestrian chain.

3.5. 2D Simulation: evacuation of a room

Fig.7 shows a simulation of an evacuation of a room with 436 pedestrians of the same mass $m = 75Kg$ with a maximum density of 6 pedestrians per square meter. the contacts are considered purely inelastic (ie k = 0). Right-velocities and the contact forces are given by:

$$\mathbf{u}^+ = \mathbf{u}^- + \mathbf{C}^T\lambda, \tag{34}$$

$$\mathbf{W}\lambda = \mathbf{C}\mathbf{u}^-. \tag{35}$$

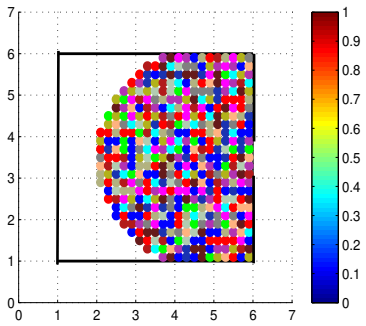


Figure 7: Initial state.

Two configurations are presented: first, the evacuation of a room with an open exit, Fig.8, 9 and 10, then, with an exit reduced to a point: all pedestrians have the attention to go to a single point Fig.11, 12 and 13.

For each configuration we present the results for three types of left-speed (the norm of left-velocity). First for a constant speed, then, the speed is chosen following normal distribution of average $1.34m/s$ and standard deviation $0.26m/s$ [45], and finally for an increasing speed from the exit.

Numerical results are obtained using the GMRES solver on MATLAB.

3.5.1. Open exit case

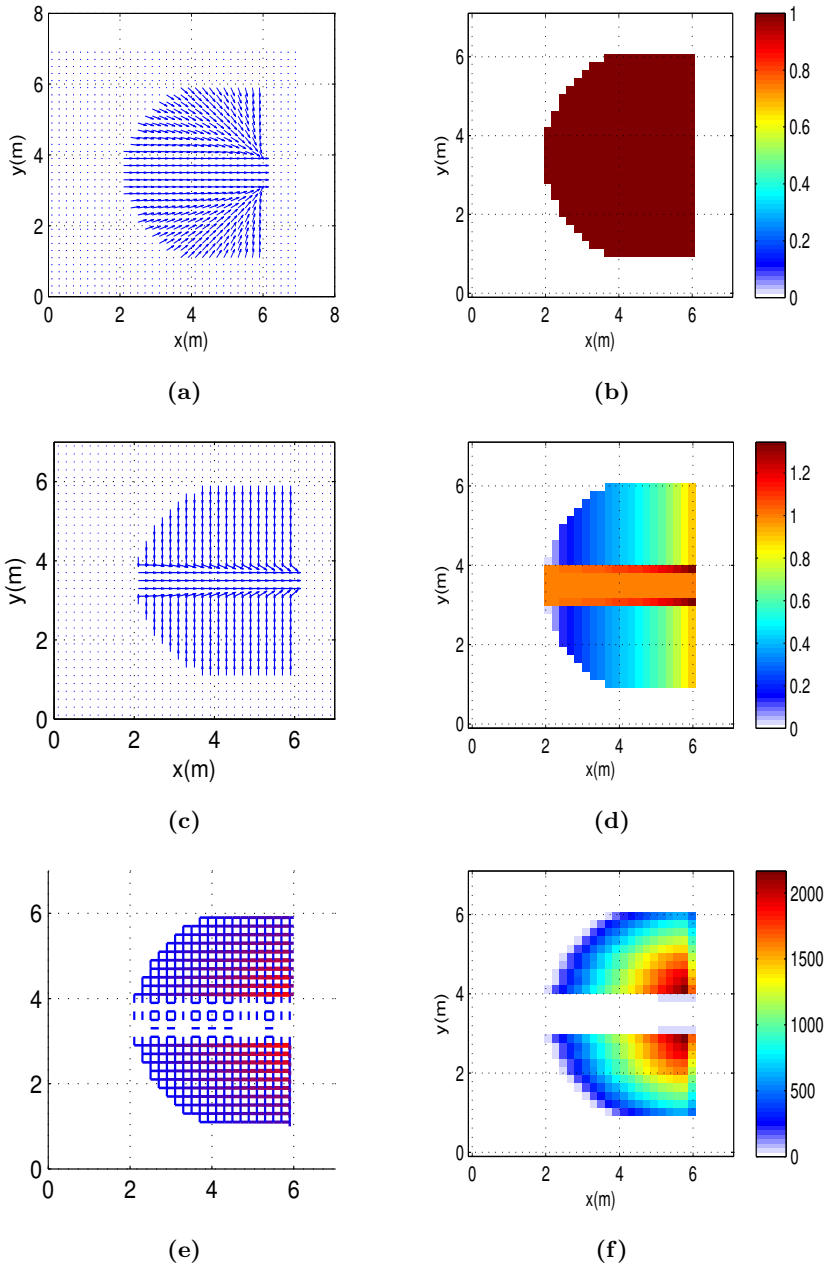


Figure 8: Simulation for a constant left-speed. (a) the desired direction, (b) left-speed, (c)pedestrian directions after contact, (d) right-speed, (e) elementary contact forces, (f) contact force applied to each pedestrian (the accumulated elementary forces)

3.5.2. Exit reduced to a point

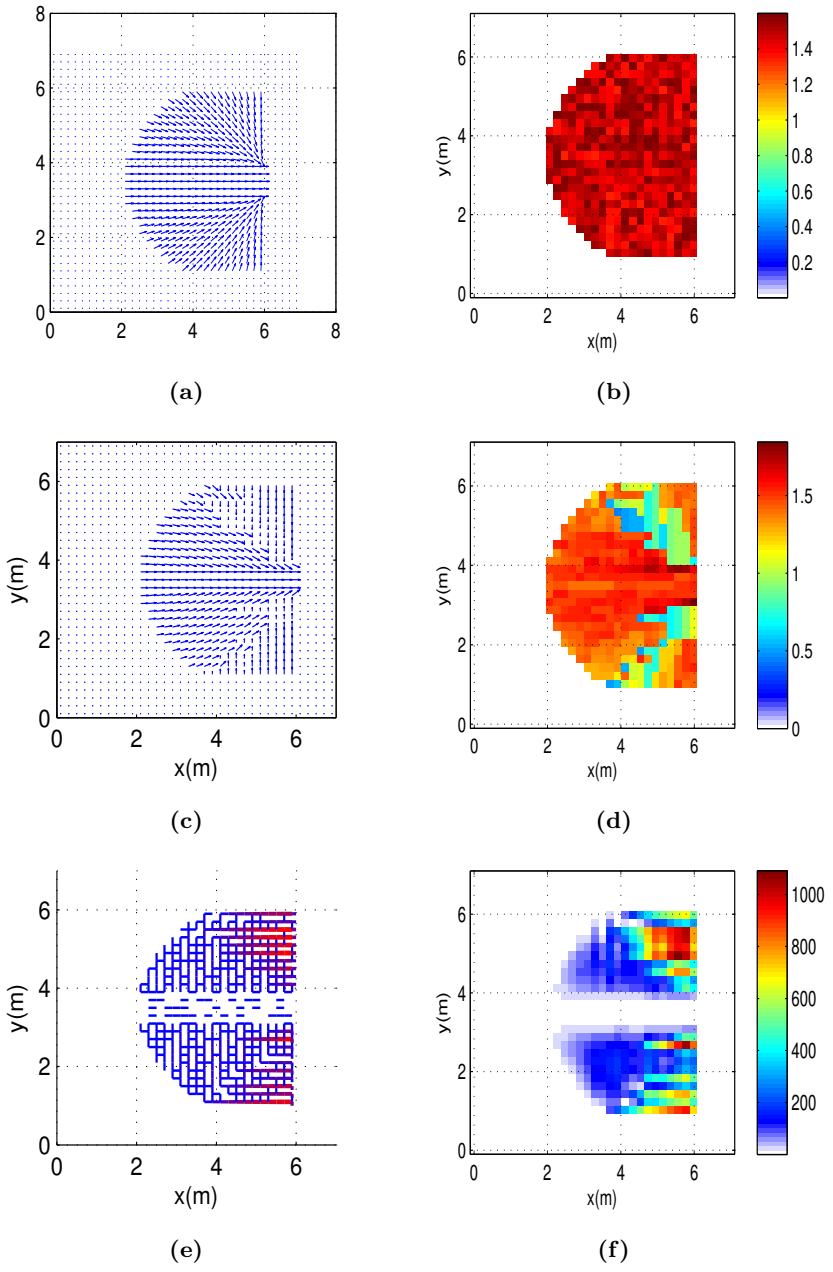


Figure 9: Simulation for a random left-speed. (a) the desired direction , (b) left-speed, (c) pedestrian directions after contact, (d) right-speed, (e) elementary contact forces, (f) contact force applied to each pedestrian (the accumulated elementary forces)

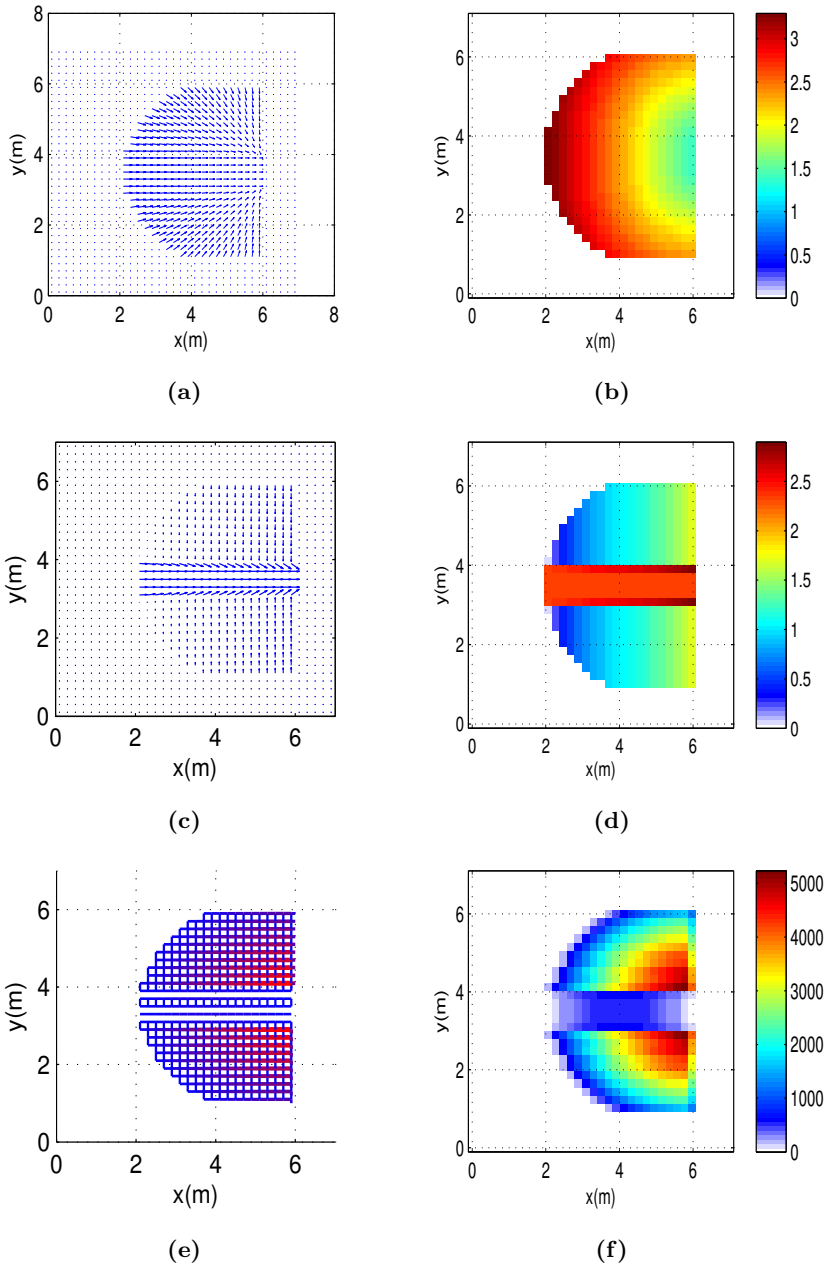


Figure 10: Simulation for a variable left-speed. (a) left-velocity, (b) left-speed, (c) pedestrian directions after contact, (d) right-speed, (e) elementary contact forces, (f) contact force applied to each pedestrian (the accumulated elementary forces)

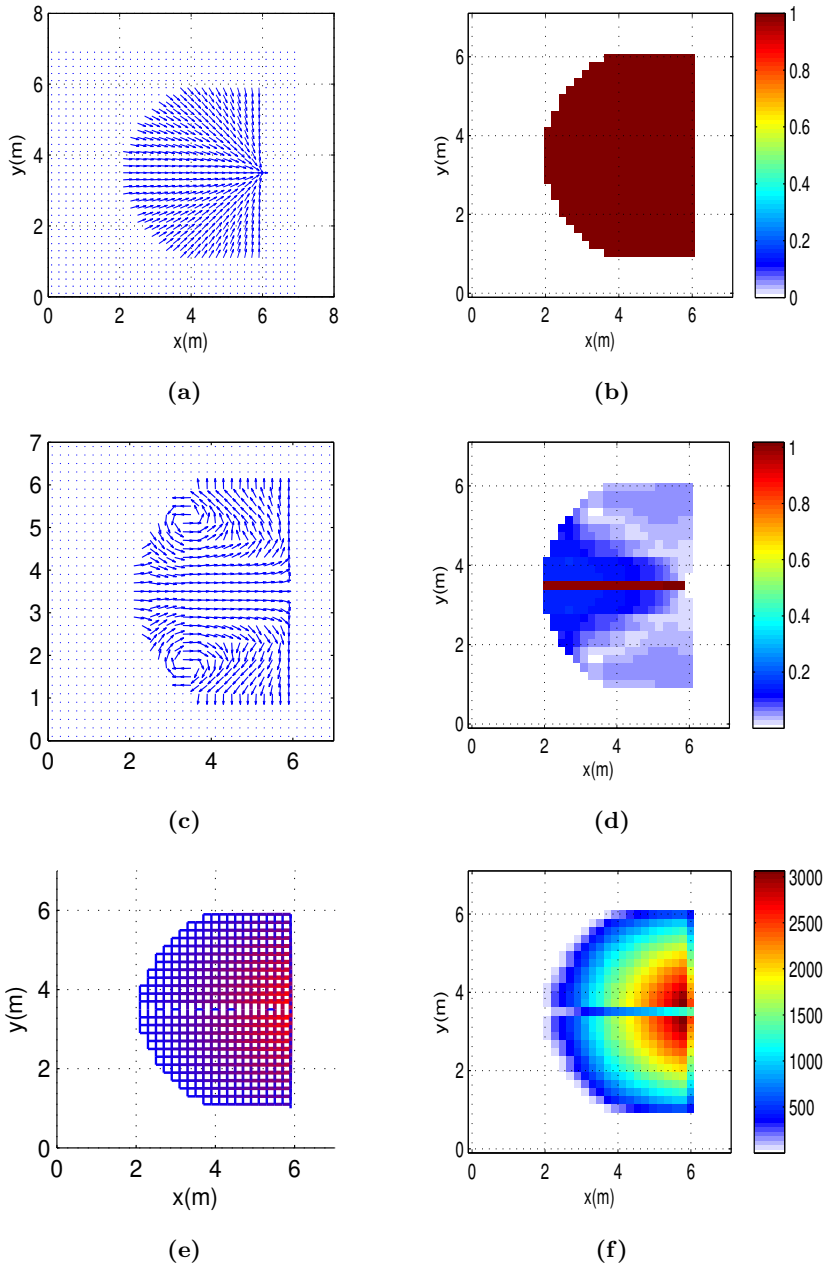


Figure 11: Simulation for a constant left-speed. (a) left-velocity, (b) left-speed, (c) pedestrian directions after contact, (d) right-speed, (e) elementary contact forces, (f) contact force applied to each pedestrian (the accumulated elementary forces)

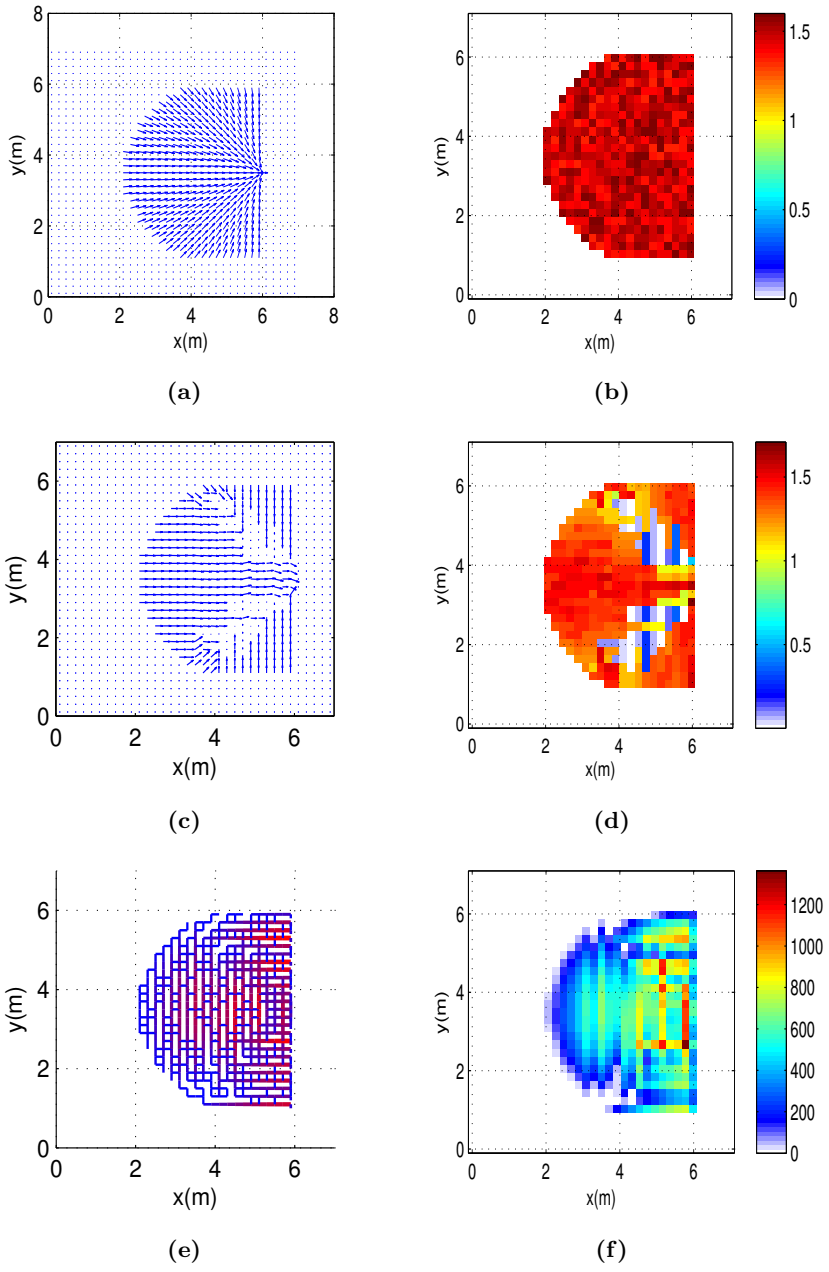


Figure 12: Simulation for a decreasing left-speed. (a) pedestrian directions after contact , (b) left-speed, (c)right-velocity, (d) right-speed, (e) elementary contact forces, (f) contact force applied to each pedestrian (the accumulated elementary forces)

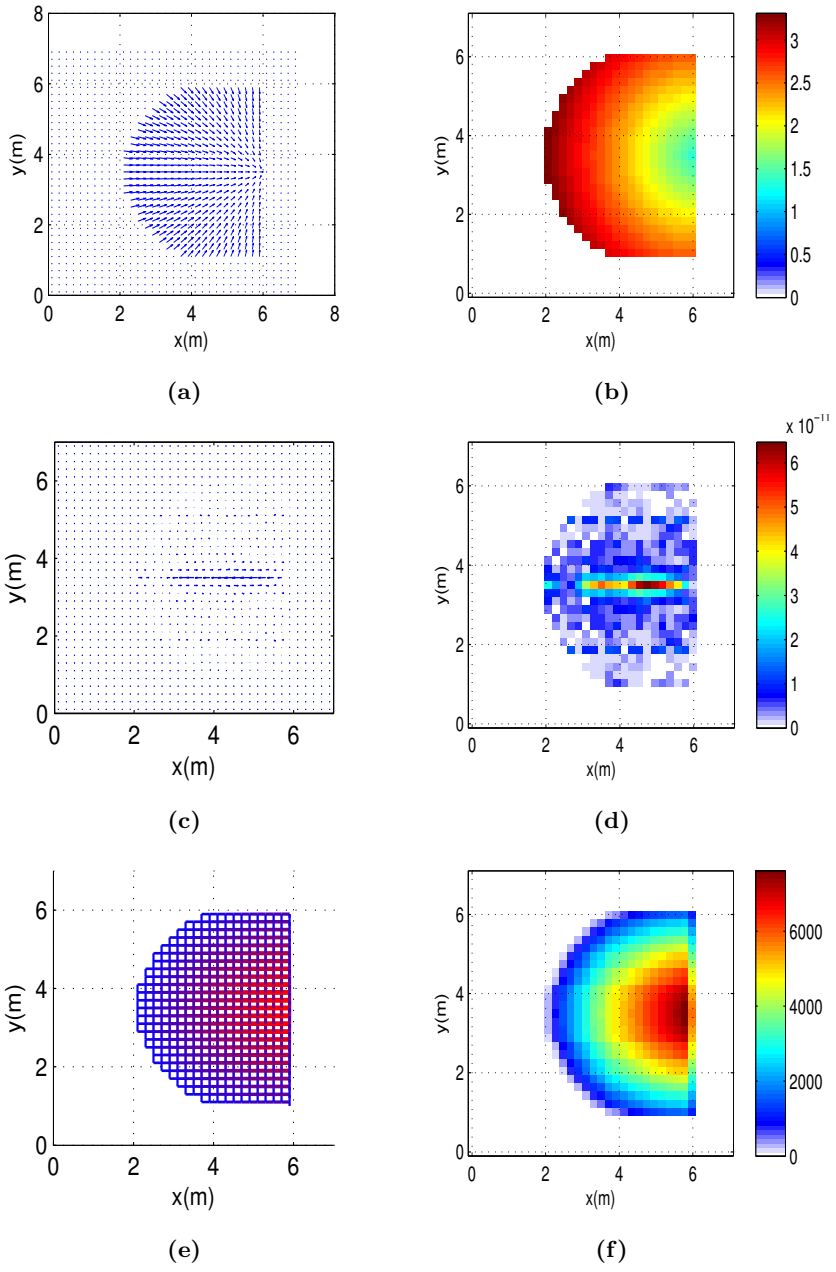


Figure 13: Simulation for a variable left-speed. (a) left-velocity, (b) left-speed, (c)pedestrian directions after contact, (d) right-speed, (e) elementary contact forces, (f) contact force applied to each pedestrian (the accumulated elementary forces)

3.5.3. Comments on the obtained results

1. Open exit case:

   - For a constant left-speed Fig.8: No interaction between the pedestrians in the band corresponding to the exit, they all have the same left-velocity therefore the contact force is zero. The right-speed of pedestrians very close to the exit is greater than their desired speed (left-speed) they are pushed by pedestrians behind them, whereas those who have near the boundary slow down. Contact forces are more important at crossing points.
   - Same conclusion for figures 9 and 10 except that contact forces appear because pedestrians do not have the same left-speed.

2. The exit reduced to a point 11,12 and 13: We notice what is happening in reality, the pedestrians closest to the exit undergo maximum contact force.

As we have mentioned in the one-dimensional case, the parameter k makes it possible to reduce the intensity of the contact force, in other words, this parameter makes it possible to model different situations; indeed: for k small enough (tends to zero), the crowd is completely in panic and contact forces are maximum, when k takes very large values, the crowd behaves in a normal way, and contact forces is reduced or even vanishes.

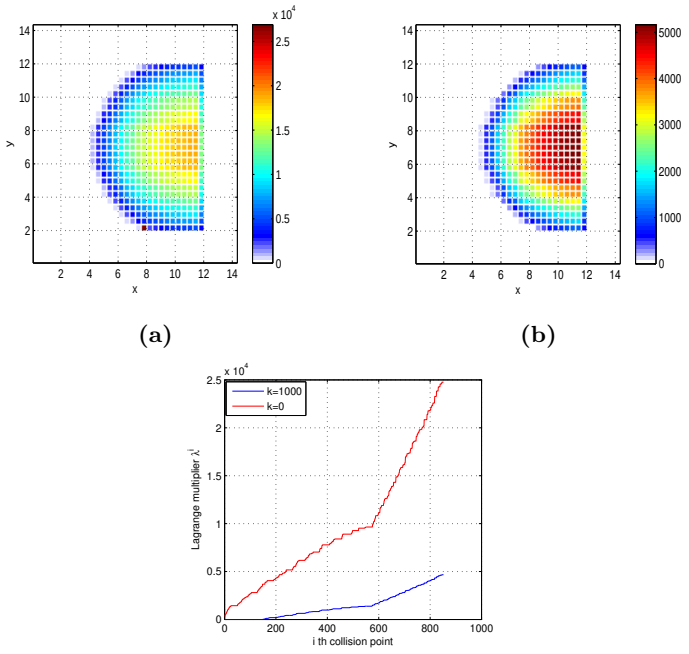


Figure 14: Pressure applied on each pedestrian: the sum of the elementary pressures generated by each contact.(a)$k = 0$, (b) $k = 1000$, (c)one-dimensional illustration.

Managing and treating multiple and simultaneous collisions (including the non-interpenetration condition) in a dense crowd, where panic situations are very common is the strong point of the approach developed above, however, to model a pedestrian as a rigid disk is a restrictive hypothesis. It is this problem that we propose to treat in the following section by identifying the entire crowd as a deformable solid.

4. A macroscopic approach: Crowd is modeled as a deformable solid

The proposed approach can be described as a continuous approach, it is also inspired by the non smooth view of contact dynamics for deformable solids developed by M.Frémond. As we have previously pointed out, we are interested in the problem of contact in a dense crowd, ie calculating the right-velocity and the pressure or the generated contact forces.

Generally, macroscopic models consist of transporting the density (modeling the evolution of the crowd) by an admissible velocity, this last one can depend on several quantities such as the desired direction, overpopulated zones, flow direction, etc . We assume that all these considerations are taken into account by the left-velocity. We take inspiration from the basic approach established by M.Frémond to model the collision of a deformable solid with a rigid plane. This approach makes it possible to estimate the pressure between the pedestrians via voluminal constraints and the pressure undergoes the obstacles is described by the non-interpenetration surface constraints.

4.1. The model: Mathematical and physical framework

Consider a dense moving crowd, whose density is denoted at instant t and at point $\mathbf{x}$ by $\rho(t, \mathbf{x})$, the left-velocity (Velocity before the collision) is $\mathbf{u}^{-}(t_c, \mathbf{x})$, where t_c is an instant of collision. By applying the principle of virtual work to the crowd as a deformable solid we obtain the equations of motion. Here we are

interested only in the collisions equation given by:

$$\rho(t_c, \mathbf{x}) \left(\mathbf{u}^+(t_c, \mathbf{x}) - \mathbf{u}^-(t_c, \mathbf{x})\right) = div\sigma(t_c, \mathbf{x}), \mathbf{x} \in \Omega, \tag{36}$$

$$\sigma(t_c, \mathbf{x})\mathbf{n}(t_c, \mathbf{x}) = -\mathbf{r}_{int}(t_c, \mathbf{x}), \mathbf{x} \in \Gamma_w, \tag{37}$$

$$\sigma(t_c, \mathbf{x})\mathbf{n}(t_c, \mathbf{x}) = \mathbf{r}_{ext}(t_c, \mathbf{x}), \mathbf{x} \in \Gamma_f, \tag{38}$$

$$\sigma(t_c, \mathbf{x})\mathbf{n}(t_c, \mathbf{x}) = 0, \mathbf{x} \in \Gamma_e, \tag{39}$$

where Ω is the area occupied by the crowd, whose border $\partial\Omega$ is defined by: $\partial\Omega = \Gamma_w \cup \Gamma_f \cup \Gamma_e$, with, Γ_w the points of contact with obstacles (walls, ...), Γ_e the exits, and Γ_f the free border that has no contact with any obstacle, see figure Fig.15.

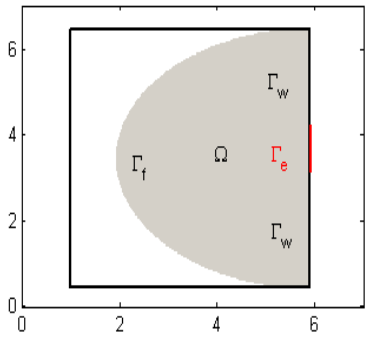


Figure 15: 2D domain of evolution.

A constitutive laws are needed to define the contact percussion $\mathbf{r}_{int}$, and the percussion stresses σ. Constitutive laws are defined by a surface pseudo-potential of dissipation ϕ_s that takes into account the impenetrability condition, and a volume pseudo-potential of dissipation ϕ_v which respect the laws of mechanics.

$$\phi_s\left(D\left(\frac{\mathbf{u}^+(t_c, \mathbf{x}) + \mathbf{u}^-(t_c, \mathbf{x})}{2}\right)\right) = k_s\left(\frac{\mathbf{u}^+(t_c, \mathbf{x}) + \mathbf{u}^-(t_c, \mathbf{x})}{2}\right)^2 ... \\ +I_-\left(\mathbf{u}^+(t_c, \mathbf{x}).\mathbf{n}(t_c, \mathbf{x})\right),$$

$$\phi_v\left(D\left(\frac{\mathbf{u}^+(t_c, \mathbf{x}) + \mathbf{u}^-(t_c, \mathbf{x})}{2}\right)\right) = k_v\left\{D\left(\frac{\mathbf{u}^+(t_c, \mathbf{x}) + \mathbf{u}^-(t_c, \mathbf{x})}{2}\right)\right\}^2.$$

With $D(u) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$.
$\mathbf{r}_{int}$ and σ derives respectively from ϕ_s and ϕ_v:

$$\begin{aligned}\mathbf{r}_{int}(t_c,\mathbf{x}) &= k_s\left(\mathbf{u}^+(t_c,\mathbf{x}) + \mathbf{u}^-(t_c,\mathbf{x})\right)... \\ &+\partial I_-\left(\mathbf{u}^+(t_c,\mathbf{x}).\mathbf{n}(t_c,\mathbf{x})\right).\mathbf{n}(t_c,\mathbf{x}), \\ \sigma(t_c,\mathbf{x}) &= k_v\left(D\left(\mathbf{u}^+(t_c,\mathbf{x}) + \mathbf{u}^-(t_c,\mathbf{x})\right)\right).\end{aligned}$$

Finally, the problem describing the crowd at the instant of collision is:

$$(P)\begin{cases}\rho(t_c,\mathbf{x})\left[\mathbf{u}(t_c,\mathbf{x})\right] = k_v div\left(D\left(\mathbf{u}^+(t_c,\mathbf{x}) + \mathbf{u}^-(t_c,\mathbf{x})\right)\right), \mathbf{x}\in\Omega, \\ \sigma(t_c,\mathbf{x})\mathbf{n}(t_c,\mathbf{x}) + k_s\left(\mathbf{u}^+(t_c,\mathbf{x}) + \mathbf{u}^-(t_c,\mathbf{x})\right)... \\ +\partial I_-\left(\mathbf{u}^+(t_c,\mathbf{x}).\mathbf{n}(t_c,\mathbf{x})\right).\mathbf{n}(t_c,\mathbf{x}) \ni 0, \mathbf{x}\in\Gamma_w, \\ \sigma(t_c,\mathbf{x})\mathbf{n}(t_c,\mathbf{x}) = \mathbf{r}_{ext}(t_c,\mathbf{x}), \mathbf{x}\in\Gamma_f, \\ \sigma(t_c,\mathbf{x})\mathbf{n}(t_c,\mathbf{x}) = 0, \mathbf{x}\in\Gamma_e.\end{cases} . \quad (40)$$

$[\mathbf{u}(t_c,\mathbf{x})] := (\mathbf{u}^+(t_c,\mathbf{x}) - \mathbf{u}^-(t_c,\mathbf{x}))$ the jump of discontinuity of the velocity. For the existence and uniqueness of the solution of the problem (P), it is shown in [17] that at the instant of collision, if the domain Ω is regular, with $mes(\Gamma_w) > 0$, and $\rho(t_c,\mathbf{x}) \geq \rho_0 > 0$, if $\mathbf{u}^-(t_c,.) \in H^1(\Omega)$, $\mathbf{r}_{ext}(t_c,.) \in L^2(\Omega)$ then there is a unique right-velocity $\mathbf{u}^+(t_c,.) \in H^1(\Omega)$ solution of (P).

4.2. Numerical simulations

4.2.1. One-dimensional case : Jostling in a chain of pedestrians lining up.

We take the one-dimensional example of pedestrians chain introduced previously. For a constrained pedestrian chain, ie, a fixed obstacle is placed at the point $x = l$, the problem (P) becomes :

$$(P_{1D})\begin{cases}\rho(t_c,x)\left(\mathbf{u}^+(t_c,x) - \mathbf{u}^-(t_c,x)\right) = k_v\frac{\partial^2(\mathbf{u}^+(t_c,x)+\mathbf{u}^-(t_c,x)}{\partial x^2}, x\in]0,l[, \\ -k_v\frac{\partial(\mathbf{u}^+(t_c,l)+\mathbf{u}^-(t_c,l))}{\partial x} + k_s\left(\mathbf{u}^+(t_c,l) + \mathbf{u}^-(t_c,l)\right) + \partial I_-(\mathbf{u}^+(l)) \ni 0. \\ \frac{\partial(\mathbf{u}^+(t_c,0)+\mathbf{u}^-(t_c,0))}{\partial x} = 0.\end{cases} \quad (41)$$

For a constant left-velocity and a uniform density, and the condition

$$\left\{\tilde{k} \geq \frac{e^{-2l/h}-1}{\left(e^{-2l/h}+1\right)}, \tilde{k} = \frac{k_s}{k_v}, \frac{1}{h^2} = \frac{\rho}{k_v}\right\}, \quad (42)$$

the solution of (P_{1D}) is:

$$u^+(t_c,x) = u^-(t_c,x)(1 - \frac{e^{x/h}+e^{-x/h}}{e^{l/h}+e^{-l/h}}, \text{with } h = \sqrt{\frac{\tilde{k}}{\rho}}, \tag{43}$$

and internal percussion is given by:

$$r_{int}(t_c,x) = u^-(t_c,x)\frac{e^{x/h}-e^{-x/h}}{h\left(e^{l/h}+e^{-l/h}\right)} \tag{44}$$

it becomes maximal as x goes to l. The condition 42 implies that $u^+(l) = 0$, ie, the right pedestrian remains crushed against the obstacle and pedestrians stay glued to each other after the impact. In the case where the condition 42 is not verified, rebounds appear and the solution is changed. Note that for an unconstrained pedestrian chain with a constant left-velocity, the right-velocity after collision is constant and the pressure is zero. However for an unconstrained pedestrian chain with a decreasing left-velocity, we find the same result as for rigid case; the pressure is maximum in the middle of the chain.

By analogy to the constrained pedestrian chain in the second-order microscopic model, we will evaluate the pressure for an uniform density and for a variable density. For this, it suffices to see that the percussion given by the equation Eq.44 is an increasing function with respect to the density ρ, this is the same result as in the microscopic case.

4.2.2. 2D Simulation: Evacuation of a room

Let us consider the two-dimensional problem of the evacuation of a room previously introduced in section 3.5. The left-velocity (desired velocity) is the unit vector directed to the exit, and the density is chosen uniform $6 pedestrian/m^2$. As in the microscopic case we presented two simulations, for an open exit Fig.17 and for an exit reduced to a point Fig.16.

4.3. Exit reduce to a point

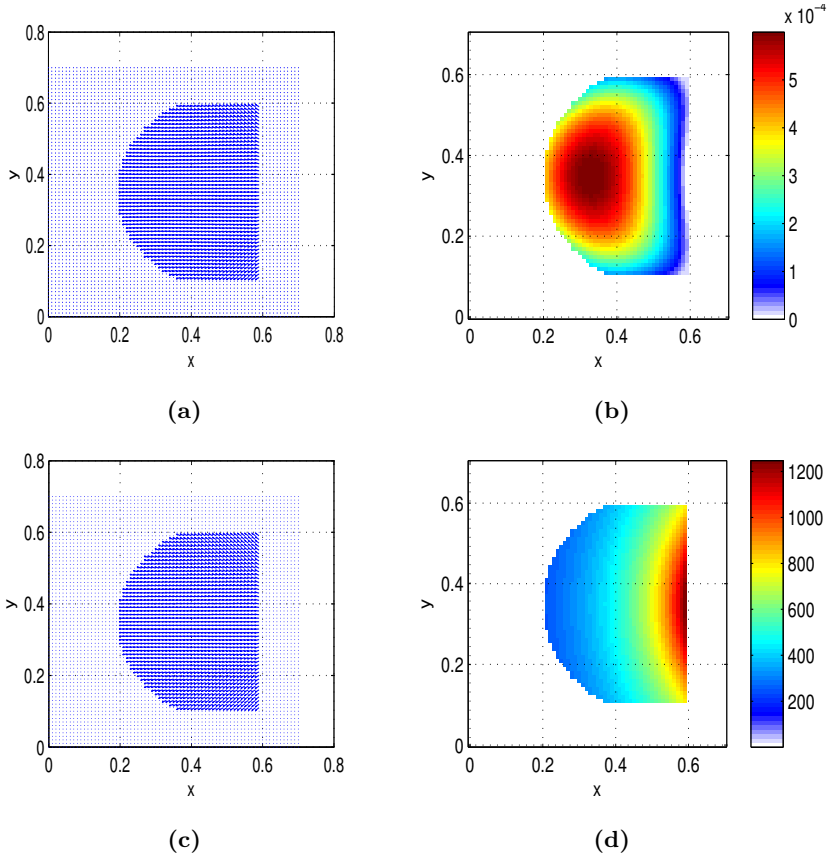


Figure 16: (a) Right-velocity, (b) right-speed, (c) pressure fields,(d) norm of pressure.

Fig.16 shows that the right-velocity is canceled at the boundary of the domain which reflects the condition of non-interpenetration with the wall, and the pressure is maximal around the exit, as for the microscopic case.

4.4. Open exit case

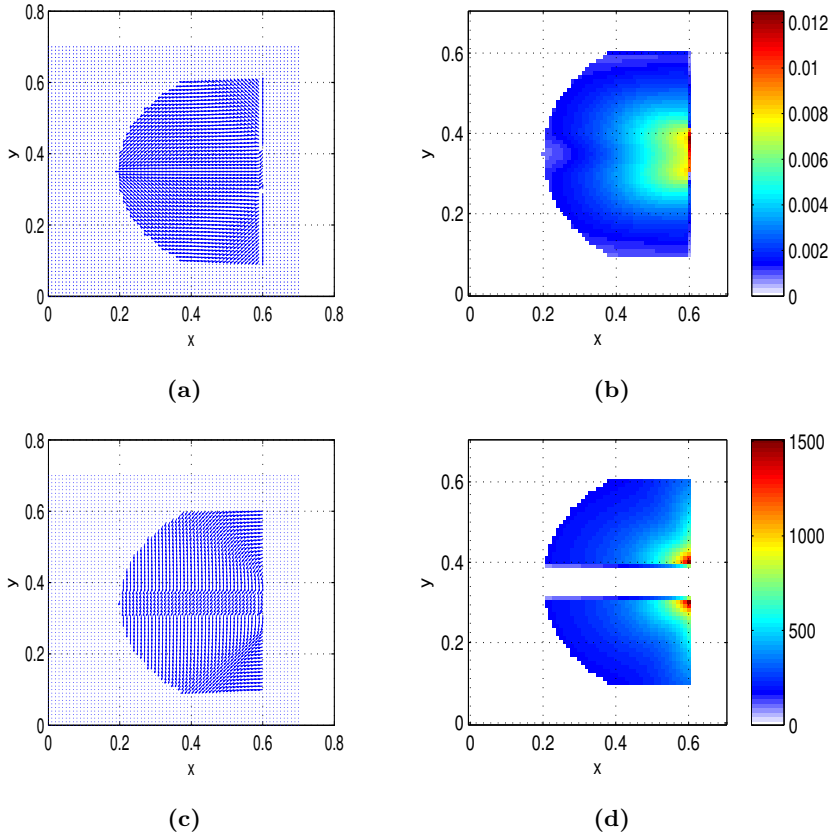


Figure 17: (a) Right-velocity, (b) right-sped, (c) pressure fields,(d) norm of pressure.

Usually we get the same result as in the microscopic case. The difference between the pressure and right-velocity values due to the choice of parameters k_v and k_s.

5. Discussion & perspectives

- **Difference between the two models: Condition of non interpenetration:** In the case of the second-order microscopic model, the velocity is computed under the constraint of non-interpenetration of both pedestrian-pedestrian and pedestrian-obstacle contact, whereas in the case of the continuous model only the non-interpenetration condition of the pedestrian-obstacle contact is taken into account. We take the example of the constrained pedestrian chain (unidimensional case of the macroscopic model) by adding a condition of non-interpenetration. We express the non-interpenetration condition of pedestrians by the non-decrease of the velocity after the contact ie

$$\frac{\partial \mathbf{u}^+(x)}{\partial x} \geq 0, \forall x \in]0, l[\tag{45}$$

So the last problem that considers both pedestrian-pedestrian and pedestrian-obstacle contact types is

$$(\tilde{P}_{1D}) \begin{cases} \rho(t_c, x)\left(\mathbf{u}^+(t_c, x) - \mathbf{u}^-(t_c, x)\right) = k_v \frac{\partial^2(\mathbf{u}^+(t_c,x)+\mathbf{u}^-(t_c,x)}{\partial x^2}, x \in]0, l[, \\ \frac{\partial \mathbf{u}^+(x)}{\partial x} \geq 0, \forall x \in]0, l[\\ -k_v \frac{\partial(\mathbf{u}^+(t_c,l)+\mathbf{u}^-(t_c,l))}{\partial x} + k_s\left(\mathbf{u}^+(t_c, \mathbf{x}) + \mathbf{u}^-(t_c, \mathbf{x})\right) + \partial I_-(\mathbf{u}^+(l)) \ni 0. \\ \frac{\partial(\mathbf{u}^+(t_c,0)+\mathbf{u}^-(t_c,0))}{\partial x} = 0. \end{cases} \tag{46}$$

For a non admissible left-velocity (here a decreasing left-velocity $u^-(t_c, x) = l - x$, the corresponding right-velocity is given by:

$$u^+(t_c, x) = 2h\left(e^{x/h} - e^{l/h}\frac{e^{x/h} + e^{-x/h}}{e^{l/h} + e^{-l/h}}\right) + l - x, \tag{47}$$

which fulfills the condition of non-inter-penetration, the associated internal effort is:

$$r_{int}(t_c, x) = 2\left(e^{x/h} - e^{l/h}\frac{e^{x/h} - e^{-x/h}}{e^{l/h} + e^{-l/h}} - 1\right). \tag{48}$$

which becomes maximal around $x = l$ ie in the vicinity of the obstacle.

- **Notion of deformability of pedestrians: A pedestrian is not a rigid body:** In the first model, pedestrians are assumed rigid, which is a restrictive hypothesis. A more real model would be to introduce the deformability of pedestrians; ie consider each individual as a deformable solid. This question is the subject of our perspectives and the question that arises is around the shape or the geometric representation of the individual. the form of the pedestrian chosen influences the numerical results (for example, in a simulation of evacuation of a room, modeling pedestrians by disks of the same masses can lead to situations of equilibrium and consequently congestion).

- **Convergence of the microscopic model to the macroscopic model:** An important problem that we are trying to address is the question of the convergence (to be defined in what sense) of the microscopic model to the macroscopic model. Numerically, we find that the microscopic model is the discrete representation of the macroscopic model. remains to prove that mathematically.

6. Conclusion

In this paper, we presented two points of view to estimate the pressure generated by a dense crowd in motion. The mathematical and physical framework is inspired by the non smooth approach of contact dynamics developed by M.Frémond and his collaborators for rigid and deformable bodies. A discrete approach is to model the crowd with a rigid disk system where the contact force is the force needed to prevent interpenetration between pedestrians and between pedestrians and obstacles. A continuous approach identifies a crowd with a deformable solid, its main purpose is to estimate correctly the pressure which is defined by volume and surface constraints. Numerical simulations are presented in a one-dimensional case: jostling in a chain of pedestrians lining up, and in a two-dimensional case: evacuation of a room. These simulations show a coherence between the two approaches.

Acknowledgement:

• The authors would like to thank Marie Mathieu-Pruvost and Nicholas Collins-Craft for their help editing the English of this paper.

• This paper was partially supported (financially) by National Center for Scientific and Technical Research (Morocco).

7. References

[1] H. Singh, R. Arter, L. Dodd, P. Langston, E. Lester, J. Drury, Modelling subgroup behaviour in crowd dynamics dem simulation, Applied Mathematical Modelling 33 (12) (2009) 4408 – 4423. doi:https://doi.org/10.1016/j.apm.2009.03.020.

[2] Pcol, Philippe, Pont, Stefano Dal, Erlicher, Silvano, Argoul, Pierre, Modelling crowd-structure interaction, Mécanique and Industries 11 (6) (2010) 495–504. doi:10.1051/meca/2010057.

[3] J. Bodgi, Pedestrian-structure synchronisation : application to swaying footbridges, Ph.D. thesis, Ecole des Ponts ParisTech (2008).

[4] B. Kabalan, P. Argoul, S. Erlicher, Crowd-Structure Interaction in Laterally Vibrating Footbridges: Comparison of Two Fully Coupled Approaches, Springer International Publishing, 2017, pp. 207–223. doi:10.1007/978-3-319-48884-4_11.

[5] B. Kabalan, P. Argoul, G. Cumunel, S. Erlicher, Z. Christoforou, A 2d discrete crowd movement model: pedestrian dynamics -crowd-structure interaction (2015).

[6] P. Pécol, S. Dal Pont, S. Erlicher, P. Argoul, Smooth/non-smooth contact modeling of human crowds movement: numerical aspects and application to emergency evacuations, Annals of Solid and Structural Mechanics 2 (2) (2011) 69–85. doi:10.1007/s12356-011-0019-3.

[7] P. Pécol, P. Argoul, S. Dal Pont, S. Erlicher, The non-smooth view for contact dynamics by michel frémond extended to the modelling of crowd movements, AIMS, Discrete and Continuous Dynamical Systems 1 (2012) 18. doi:hal-00693217.

[8] B. Kabalan, P. Argoul, A. Jebrane, G. Cumunel, S. Erlicher, A crowd movement model for pedestrian flow through bottlenecks, Annals of Solid and Structural Mechanics 8 (1) (2016) 1–15. doi:10.1007/s12356-016-0044-3.

[9] P. Argoul, B. Kabalan, Pedestrian Trajectories and Collisions in Crowd Motion, Springer Berlin Heidelberg, Berlin, Heidelberg, 2017, pp. 79–144. doi:10.1007/978-3-662-52696-5_6.

[10] B. Maury, Handling congestion in crowd motion modeling (2011). doi:10.3934/nhm.2011.6.485.

[11] B. Maury, A. Rouneff, F. Santanbrogio, A macroscopic crowd motion model of gradient flow type, Mathematical Models and Methods in Applied Sciences 20 (10) (2010) 1787–1821. doi:10.1142/S0218202510004799.

[12] B. Maury, J. Venel, A mathematical framework for a crowd motion model, Comptes Rendus Mathematique 346 (23) (2008) 1245 – 1250. doi:https://doi.org/10.1016/j.crma.2008.10.014.

[13] B. Maury, J. Venel, Handling of Contacts in Crowd Motion Simulations, 2009, pp. 171–180. doi:10.1007/978-3-540-77074-9_15.

[14] B. Maury, J. Venel, A discrete contact model for crowd motion, ESAIM: M2AN 45 (1) (2011) 145–168. doi:10.1051/m2an/2010035.

[15] A. Roudneff, Macroscopic modelling of crowd motion, Ph.D. thesis, Université Paris Sud - Paris XI (2011).

[16] S. Faure, B. Maury, Crowd motion from the granular standpoint, Mathematical Models and Methods in Applied Sciences 1. doi:10.1142/S0218202515400035.

[17] M. Frémond, Collisions Engineering: Theory and Applications, Springer Series in Solid and Structural Mechanics, 2017.

[18] M. Frémond, Rigid bodies collisions, Physics Letters A 204 (1) (1995) 33 – 41. doi:http://dx.doi.org/10.1016/0375-9601(95)00418-3.

[19] F. Caselli, M. Frémond, Collision of three balls on a plane, Computational Mechanics 43 (6) (2009) 743–754. doi:10.1007/s00466-008-0342-7.

[20] S. D. Pont, E. Dimnet, A theory for multiple collisions of rigid solids and numerical simulation of granular flow, International Journal of Solids and Structures 43 (20) (2006) 6100 – 6114. doi:http://dx.doi.org/10.1016/j.ijsolstr.2005.10.001.

[21] S. D. Pont, E. Dimnet, Theoretical approach to instantaneous collisions and numerical simulation of granular media using the $a - cd^2$ method, Communications in Applied Mathematics and Computational Science 3 (1) (2008) 1–24.

[22] P. Pecol, P. Argoul, S. Dal Pont, S. Erlicher, A new crowd movement modeling for pedestrians who hold hands, in: XVIIIth Symposium Vibrations, Chocs et Bruit & ASTELAB, EDF Clamart, France, 2012.

[23] M. Moussad, D. Helbing, G. Theraulaz, How simple rules determine pedestrian behavior and crowd disasters, Proceedings of the National Academy of Sciences 108 (17) (2011) 6884–6888. doi:10.1073/pnas.1016507108.

[24] V. T. Helbing D, Farkas I, Simulating dynamical features of escape panic, Nature 108 (17) (2000) 6884–6888. doi:10.1073/pnas.1016507108.

[25] D. Helbing, P. Molnár, Social force model for pedestrian dynamics, Phys. Rev. E 51 (1995) 4282–4286. doi:10.1103/PhysRevE.51.4282.

[26] D. Helbing, I. Farkas, P. Molnàr, T. Vicsek, Simulation of pedestrian crowds in normal and evacuation situations, in: M. Schreckenberg, S. D. Sharma (Eds.), Pedestrian and Evacuation Dynamics, Springer, Berlin, 2002, pp. 21–58.

[27] N. Mahato, A. Klar, S. Tiwari, Particle methods for multi-group pedestrian flow, Applied Mathematical Modelling 53 (2018) 447 – 461. doi:https://doi.org/10.1016/j.apm.2017.08.024.

[28] D. Helbing, A fluid dynamic model for the movement of pedestrians (1992). arXiv:cond-mat/9805213.

[29] Y.-Q. Jiang, R.-Y. Guo, F.-B. Tian, S.-G. Zhou, Macroscopic modeling of pedestrian flow based on a second-order predictive dynamic model, Applied Mathematical Modelling 40 (23) (2016) 9806 – 9820. doi:https://doi.org/10.1016/j.apm.2016.06.041.

[30] M. Twarogowska, P. Goatin, R. Duvigneau, Macroscopic modeling and simulations of room evacuation, Applied Mathematical Modelling 38 (24) (2014) 5781 – 5795. doi:https://doi.org/10.1016/j.apm.2014.03.027.

[31] L. Bruno, A. Tosin, P. Tricerri, F. Venuti, Non-local first-order modelling of crowd dynamics: A multidimensional framework with applications, Applied Mathematical Modelling 35 (1) (2011) 426 – 445. doi:https://doi.org/10.1016/j.apm.2010.07.007.

[32] S.-f. Li, P. Zhang, S. C. Wong, Conservation form of helbing's fluid dynamic traffic flow model, Applied Mathematics and Mechanics 32 (9) (2011) 1109. doi:10.1007/s10483-011-1485-9.

[33] R. L. Hughes, A continuum theory for the flow of pedestrians, Transportation Research Part B: Methodological 36 (6) (2002) 507 – 535. doi:https://doi.org/10.1016/S0191-2615(01)00015-7.

[34] L. Taneja, N. B. Bolia, Network redesign for efficient crowd flow and evacuation, Applied Mathematical Modelling 53 (2018) 251 – 266. doi:https://doi.org/10.1016/j.apm.

2017.08.030.
URL http://www.sciencedirect.com/science/article/pii/S0307904X17305486

[35] N. Bellomo, C. Bianca, V. Coscia, On the modeling of crowd dynamics: An overview and research perspectives, SeMA Journal 54 (1) (2011) 25–46. doi:10.1007/BF03322586.

[36] N. B. G. Ajmone Marsane, L. Gibelli, Stochastic evolutionary differential games toward a systems theory of behavioral social dynamics, Mathematical Models and Methods in Applied Sciences 26 (06) (2016) 1051–1093. doi:10.1142/S0218202516500251.

[37] N. Bellomo, L. Gibelli, Toward a mathematical theory of behavioral-social dynamics for pedestrian crowds, Mathematical Models and Methods in Applied Sciences 25 (13) (2015) 2417–2437. doi:10.1142/S0218202515400138.

[38] A. Elmoussaoui, P. Argoul, M. E. Rhabi, A. Hakim, Discrete kinetic theory for 2d modeling of a moving crowd: Application to the evacuation of a non-connected bounded domain, Computers and Mathematics with Applicationsdoi:https://doi.org/10.1016/j.camwa.2017.10.023.

[39] C. Cholet, Chocs de solides rigides, Ph.D. thesis, Universit Pierre et Marie Curie (Paris) (1998).

[40] C. Cholet, Collisions d'un point et d'un plan, Comptes Rendus de l'Académie des Sciences - Series I - Mathematics 328 (5) (1999) 455 – 458. doi:http://dx.doi.org/10.1016/S0764-4442(99)80189-X.

[41] J. J. Moreau, An introduction to Unilateral Dynamics, Springer Berlin Heidelberg, 2004.

[42] J. J. Moreau, Unilateral Contact and Dry Friction in Finite Freedom Dynamics, Springer Vienna, Vienna, 1988, pp. 1–82.

[43] R. Dzonou, M. D. M. Marques, A sweeping process approach to inelastic contact problems with general inertia operators, European Journal of Mechanics - A/Solids 26 (3) (2007) 474 – 490. doi:http://dx.doi.org/10.1016/j.euromechsol.2006.07.002.

[44] M. Kunze, M. D. M. Marques, An introduction to moreau's sweeping process, in: B. Brogliato (Ed.), Impacts in Mechanical Systems, Springer Berlin Heidelberg, 2000, pp. 1–60.

[45] L. F. HENDERSON, The statistics of crowd fluids, Nature Research journals 229 (1971) 381383. doi:10.1038/229381a0.

Computers and Mathematics with Applications 75 (2018) 1159–1180

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Discrete kinetic theory for 2D modeling of a moving crowd: Application to the evacuation of a non-connected bounded domain

A. Elmoussaoui [a], P. Argoul [b,c], M. El Rhabi [a,d], A. Hakim [a,*]

[a] LAMAI, FST Marrakech, Université Cadi Ayyad, Morocco
[b] Université Paris-Est, LVMT (UMR_T 9403), Ecole des Ponts ParisTech, IFSTTAR, UPEMLV, F-77455 Marne la Vallée, France
[c] Université Paris-Est, MAST, SDOA, IFSTTAR, F-77447 Marne-la-Vallée, France
[d] IMI, Ecole des Ponts ParisTech, Université Paris Est, France

ARTICLE INFO

Article history:
Received 8 February 2017
Received in revised form 22 October 2017
Accepted 28 October 2017
Available online 23 November 2017

Keywords:
Discrete kinetic theory
Complex system
Evacuation
Crowd dynamics
Splitting scheme

ABSTRACT

This paper concerns the mathematical modeling of the motion of a crowd in a non connected bounded domain, based on kinetic and stochastic game theories. The proposed model is a mesoscopic probabilistic approach that retains features obtained from both micro- and macro-scale representations; pedestrian interactions with various obstacles being managed from a probabilistic perspective. A proof of the existence and uniqueness of the proposed mathematical model's solution is given for large times. A numerical resolution scheme based on the splitting method is implemented and then applied to crowd evacuation in a non connected bounded domain with one rectangular obstacle. The evacuation time of the room is then calculated by our technique, according to the dimensions and position of a square-shaped obstacle, and finally compared to the time obtained by a deterministic approach by means of randomly varying some of its parameters.

1. Introduction

The dynamic modeling of crowd motion has recently aroused a great interest in the scientific community and is used in numerous applications, such as engineering and social science [1]. It has become increasingly important to avoid or control panic situations and to ensure the safety of people in congested areas.

Mathematical representations of crowd motion from the microscopic to macroscopic scale have been an active field of study for the last two decades, with a rich scientific literature [2–6]. The aim of this paper is not to present an exhaustive list of references. Only some of the most frequently used will be mentioned: the microscopic approach based on the social forces model, proposed by Helbing [4,5], where the movement of the crowd is characterized by the position and velocity of each individual, and the macroscopic models, given by Hughes [6], that consider the crowd as a fluid.

Recently, an intermediate mesoscopic representation based on the kinetic approach appeared and its application to crowd representation gave promising results for the description of pedestrians' strategy. Very few references of the mesoscopic representation can be found in the literature [7–11].

The kinetic theory is a mathematical description of a volume of material containing a large number of particles interacting with each other, for example, a volume of gas particles [12]. This approach allows us to connect both the macroscopic

* Corresponding author.
E-mail addresses: abdelghani.elmoussaoui@ced.uca.ma (A. Elmoussaoui), pierre.argoul@ifsttar.fr (P. Argoul), mohammed.el-rhabi@enpc.fr (M. El Rhabi), a.hakim@uca.ma (A. Hakim).

https://doi.org/10.1016/j.camwa.2017.10.023
0898-1221/

and microscopic properties. Monte Carlo particle methods have a relevant role in the numerical resolution of kinetic equations [13,14]. Moreover, this theory has been applied in many areas, namely, modeling of vehicular traffic [15,16], and crowd dynamics [7–11,17,18], which is the subject of this study.

The modeling of a crowd by a kinetic approach started with Bellomo's and Bellouquid's work [9], in which the set of main governing equations are introduced. In this approach, the crowd is seen as a complex system in which the interactions between people (particles) are managed from a probabilistic point of view and the microscopic state of each pedestrian (particle) is characterized by his/her position and speed. In addition, the general form of the system is represented by a distribution function in a microscopic state and the dynamic of this distribution function is given by the study of particles' balance in a unit volume element of the phase plane. Then Bellomo et al. [8] developed this approach and handled the movement of a crowd moving in different directions, in an unbounded domain and where the objective of each particle is to reach a fixed target. Afterward, Agnelli et al. [7] studied the case of people moving in a connected bounded domain, without obstacles. Then Bellomo and Gibelli treated the density–velocity diagram in steady flow conditions and studied some collective emerging behaviors that are experimentally observed, namely the self-organized behaviors leading to the creation of lanes in streets and the increasing of evacuation time in stressful conditions [17].

Most of the studies that were previously mentioned, concerned connected areas [7–9,17], while the question of non connected areas is still open. In this paper, the kinetic theory applied to crowd dynamics is extended to its motion in a non-connected bounded domain, with the presence of fixed obstacles. To model interactions, it is assumed that pedestrians can change their direction for various reasons, such as: the wish to reach a target, the avoidance of the edges of the domain and/or fixed obstacles in the field. In a future step, pedestrians will be considered as "active particles" by taking into account their heterogeneity and their capacity to develop a strategy of displacement.

This paper is organized as follows: Section 2 provides the mathematical model for the crowd evacuation in an area including walls and obstacles. Then, probabilistic tools are used to describe pedestrian–pedestrian interactions as well as pedestrians interactions along with the geometry of the area. Section 3 presents a mathematical framework to obtain proofs of the existence and of the uniqueness of the proposed model's solution. Section 4 is devoted to numerical simulations to check the ability of the proposed model to describe the main features of the pedestrian dynamics, particularly the avoidance of fixed obstacles on their walk toward the exit. The influence of the position of a fixed square obstacle in the vicinity of the exit is finally studied with respect to the evacuation time for a group of 50 persons.

2. Position of the problem under study. Mathematical modeling

Let us consider a system composed of N particles (the pedestrians) distributed randomly in a two-dimensional bounded domain $\Omega \subset \mathbb{R}^2$.

This group of N pedestrians present in the room at initial time t_0, wish to evacuate the room by the exit of size S. At initial time $t = t_0$, pedestrians are distributed within a disk $\mathcal{D}_0$ of radius r and of center $M_0(x_0, y_0)$. The initial global density is then: $\rho_0 = \frac{N}{\pi r^2}$ (ped/m^2).

Kinetic type equations derivation requires a detailed analysis of the interactions at a micro-scale, namely at the pedestrian scale related to the statistical representation of the overall system; this requires a suitable probability distribution over the micro-state.

This particle distribution function is given by: $f = f(t, \mathbf{x}, \mathbf{v})$ for all $t \geq t_0, \mathbf{x} \in \Omega, \mathbf{v} \in D_{\mathbf{v}}$, where $D_{\mathbf{v}}$ represents the domain of velocities.

If $f(t, \mathbf{x}, \mathbf{v})$ is locally integrable in $\mathbf{x}$, then $f(t, \mathbf{x}, \mathbf{v})d\mathbf{x}d\mathbf{v}$ represents the number of individuals, located at time t in an infinitesimal rectangle $[x, x+dx] \times [y, y+dy]$ with the velocity belonging to $[v_x, v_x+dv_x] \times [v_y, v_y+dv_y]$, where: $\mathbf{x} = (x, y)$ and $\mathbf{v} = \left(v_x, v_y\right)$.

If $f(t, \mathbf{x}, \mathbf{v})$ is locally integrable in $\mathbf{v}$, the local density (the number of people per square meter) at the point $\mathbf{x}$ and time t can be introduced:

$$\rho\,(t, \mathbf{x}) = \int_{D_{\mathbf{v}}} f(t, \mathbf{x}, \mathbf{v})d\mathbf{v}. \tag{1}$$

At initial time t_0, it can be written that: $\rho\,(t_0, \mathbf{x}) = \rho_0\, \mathbf{1}_{\mathcal{D}_0}(\mathbf{x})$ where $\mathbf{1}_{\mathcal{D}_0}(\mathbf{x})$ is the indicator function of the subset $\mathcal{D}_0$.

The impact of crowd density for a standing crowd and a moving crowd is important to understand for crowd safety. In the UK Guides produced to advise on crowd safety issues (cf. [19]), the safety limit for crowd density is stated as 4 pedestrians per square meter for a moving crowd and 4.7 for standing areas. To be closer to reality, the individual dimensions of pedestrians must be taken into account in the density analysis. For a totally packed metro train (French RATP), the density is between 6 and 8 pedestrians/m^2. In conclusion, a maximum value ρ_{max} for local crowd density, $\rho_{max} \leq 8$ pedestrians/m^2, is introduced, and a maximum number of pedestrians N_{max} is then deduced: $N_{max} = \pi\, r^2\, \rho_{max}$.

In our model, dimensionless quantities are preferred. To do that, from the following reference variables:

- L: a characteristic length of the domain Ω, for example its diagonal when Ω is rectangle shaped,
- V_m: the maximum speed of the pedestrian walking unobstructed in the environment,

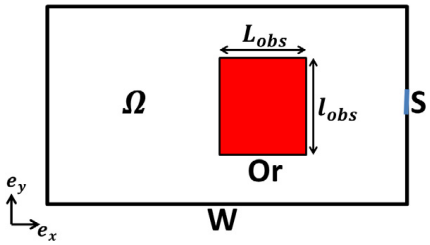


Fig. 1. Illustration of the walking domain. W: wall, Or: rectangular obstacle, S: exit, $\partial\Omega = W \cup S \cup \partial Or$ (with ∂Or is the obstacle's border).

- T_m: a reference value for the evacuation time is given by: $T = L/V_m$,
- ρ_{max}: the maximum local crowd density,

the following unit-less variables are then defined:

- the position variable: $\tilde{\mathbf{x}} = \frac{\mathbf{x}}{L}$.
- the time: $\tilde{t} = \frac{t}{T}$.
- the velocity modulus: $\tilde{v} = \frac{v}{V_m}$,
- The distribution function: $\tilde{f} = \frac{f}{\rho_{max}}$, leading to $\tilde{\rho} = \frac{\rho}{\rho_{max}}$.

In the following, for the sake of simplicity, the tildes are omitted.

2.1. Representation of the pedestrian environment

The bounded domain $\Omega \subset \mathbb{R}^2$, in which pedestrians move, is a room of rectangular shape of length L_x and width L_y. Its external border is made of walls and is noted W.

A single obstacle noted by Or with ∂Or being its boundary is added within Ω. In the following, the obstacle is rectangle-shaped with sides of length L_{obs} and width l_{obs}. A reference frame $(O, \mathbf{e}_x, \mathbf{e}_y)$ is defined and polar coordinates are preferred. For the sake of simplicity, this study is limited to the case of a single obstacle cf. Fig. 1, but the geometry of the domain can be further modified by inserting several obstacles of different shapes.

Two parameters already present in [17] are finally introduced $\alpha \in [0, 1]$ and $\xi \in [0, 1]$, respectively.

α is related to the quality of the domain: when $\alpha = 0$, the domain is of poor quality, that means that pedestrians are forced to stop walking, while when $\alpha = 1$, the quality of the domain is maximum, allowing a pedestrian to walk with the highest speed. ξ characterizes the strength of the pedestrian's preference for areas of low pedestrian density and it is supposed to give indication of the level of pedestrians' anxiety.

2.2. Modeling the velocity vector

The approach developed in the present work refers to the hybrid approach reviewed in [7,8], where the discrete variable for the individual velocity states θ, defined by the angle of the selected velocity $\mathbf{v}$ direction with $\mathbf{e}_x$, is used. The velocity vector can then be expressed as

$$\mathbf{v}[\rho] = v[\rho](t, \mathbf{x})\, cos(\theta)\, \mathbf{e}_x + v[\rho](t, \mathbf{x})\, sin(\theta)\, \mathbf{e}_y,$$

where $v[\rho]$ is the speed; square brackets are used to denote that $\mathbf{v}$ can depend on ρ in a functional way, for instance on ρ and on its gradient $\frac{\partial \rho}{\partial \boldsymbol{x}}$.

The speed variation depends on the interactions between pedestrians. Specifically, pedestrians adjust their velocity modulus according to the level of congestion around them and on the environmental conditions. In this paper, the dependence of the velocity on the local density is motivated by the fundamental diagram developed in [17]. The main idea is that the velocity of the pedestrian decreases monotonically with the density from the maximal value $v[\rho = 0] = 1$ of $\rho = 0$ to $v[\rho = 1] = 0$ where $\rho = 1$, corresponding to the maximal density. Moreover, the maximal velocity observed at very low density increases with the quality of the environmental conditions and/or the pedestrian's anxiety.

From the density speed diagram developed in [17], the following density speed is deduced:

$$v[\rho](t, \mathbf{x}) = \frac{\sigma^3 (1 - \rho(t, \mathbf{x}))^2}{\sigma^2 (1 - \rho(t, \mathbf{x}))^2 + (1 - \sigma)\, \rho^2(t, \mathbf{x})}, \tag{2}$$

where $\sigma = \alpha\, \xi$; the parameters α and ξ, previously introduced, characterize respectively the quality of the domain, and the trend of pedestrians to adapt their walk to their surroundings instead of searching for less crowded areas.

2.3. Mathematical model equations

The number of pedestrians N is limited by the maximum number N_{max} and generally not large enough to justify the continuity assumption of the particle distribution function with respect to the velocity. Indeed, it is assumed that the velocity directions θ_i take discrete values in the following set:

$$S_\theta = \left\{ \theta_i = \frac{i-1}{n} 2\pi : \ i = 1, \ldots, n \right\}.$$

Therefore, due to the deterministic nature of the variable v, the distribution function can be expressed as:

$$f(t, \mathbf{x}, \theta) = \sum_{i=1}^{n} f_i(t, \mathbf{x}) \, \delta(\theta - \theta_i), \tag{3}$$

where δ is the Dirac distribution; $f_i(t, \mathbf{x}) = f(t, \mathbf{x}, \theta_i)$, represents pedestrians, viewed as active particles, moving in direction θ_i, at time t and position $\mathbf{x}$, per unit area.

The local density $\rho(t, \mathbf{x})$ previously defined in Eq. (1), becomes the sum of $f_i(t, \mathbf{x})$ for $1 \leq i \leq n$:

$$\rho(t, \mathbf{x}) = \sum_{i=1}^{n} f_i(t, \mathbf{x}). \tag{4}$$

The derivation of the mathematical structure used in the present paper refers to the theory developed in [8], where the mathematical model is obtained by a balance of particles in a unit volume of the micro state space. Indeed, the motion of a particles group (pedestrians) is governed by the partial derivative equation (PDE) of transport applied to f_i with a second member Γ_i characterizing the different interactions between pedestrians with their environment:

$$\partial_t f_i(t, \mathbf{x}) + \mathbf{v}_i[\rho](t, \mathbf{x}) . \partial_{\mathbf{x}} f_i(t, \mathbf{x}) = \Gamma_i(t, \mathbf{x}) \quad i = 1, \ldots, n, \tag{5}$$

where: $\mathbf{v}_i[\rho](t, \mathbf{x}) = v[\rho](t, \mathbf{x}) \, (cos(\theta_i), sin(\theta_i))^T$ and $\partial_{\mathbf{x}} = \left(\partial_x, \partial_y\right)^T$. This $\Gamma_i(t, \mathbf{x})$ term also allows us to take into account boundary conditions (walls W and the obstacle's borders ∂Or).

2.4. Modeling of interactions

By referring to [7,11,17], modeling the interactions is a decision process in which each particle moves along with the others as well as within the geometry of the domain. The interaction involves three types of particles:

- The test particle with micro-state $(\mathbf{x}, \theta_i)$ and distribution function $f_i(t, \mathbf{x}) = f(t, \mathbf{x}, \theta_i)$.
- The field particle with micro-state $(\mathbf{x}^*, \theta_k)$ and distribution function $f_k(t, \mathbf{x}) = f(t, \mathbf{x}^*, \theta_k)$.
- The candidate particle with micro-state $(\mathbf{x}, \theta_h)$ and distribution function $f_h(t, \mathbf{x}) = f(t, \mathbf{x}, \theta_h)$.

The candidate particle can acquire, in probability, the micro-state of the test particle after interaction with the field particles, while the test particle loses its state in the interaction with the field particles.

Two types of interactions are considered, those between candidate and field particles and those between candidate particle and obstacle (either within the domain or the border walls themselves).

In this way the right hand side of Eq. (5) can be decomposed in two terms as: $\Gamma_i = \Gamma_i^P + \Gamma_i^D$, where Γ_i^P refers to interactions between pedestrians and Γ_i^D to interactions between pedestrians and obstacles (either present within the domain or against the border walls). Both terms are detailed in the following sub-paragraphs.

2.4.1. The Γ_i^p term

The term Γ_i^P of interactions between pedestrians (particles) is defined in a probabilistic sense, since pedestrians will not react in the same way when facing a particular situation. Interactions of test and candidate particle with field particle can be modeled by the following quantities:

- Interaction domain (visibility zone) : it represents the area where the trajectory of each candidate pedestrian can be influenced by those of other field pedestrians, which can be defined as circular sector with radius V symmetric with respect to the velocity direction defined by the visibility angle ϕ (see Fig. 2):

$$V(\mathbf{x}, \mathbf{e}_d, R_V, \phi) = \left\{ \mathbf{y} \in \Omega / \ \ \|\mathbf{x} - \mathbf{y}\|_2 \leq R_V, \ \ \frac{\mathbf{y} - \mathbf{x}}{\|\mathbf{y} - \mathbf{x}\|_2} . \mathbf{e}_d \geq \cos(\phi) \right\}.$$

  This visibility zone is introduced just when a pedestrian located at the point $\mathbf{x}$, interacts with the other pedestrians, but it not introduced in the case of interaction with obstacles and walls.
- The interaction rate $\eta[\rho(t, \mathbf{x})]$ characterizes the contact frequency that a candidate h-pedestrian (or test) in $\mathbf{x}$ develops with a field k-pedestrian in the visibility zone V. The use of the same idea developed in [7] is proposed, by treating the interaction rate with increasing local density $\eta[\rho(t, \mathbf{x})] = \eta_0 \rho(t, \mathbf{x})$, where η_0 is a constant.

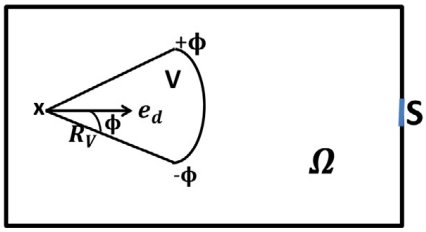


Fig. 2. Vision field V for a pedestrian located at the point $\mathbf{x}$ with unit velocity direction $\mathbf{e}_d$.

- The transition probability density $B_{hk}(i)$, characterizes the fact that the candidate pedestrian changes his/her direction θ_h to the test pedestrian's direction θ_i, due to the interaction with field pedestrians of a direction θ_k. This probability is assumed to be dependent on the density of pedestrians $\rho(t, \mathbf{x})$.

  The probabilities $B_{hk}(i)$ of each i-th pedestrian satisfy the following relationship:

$$\sum_{i=1}^{n} B_{hk}[\rho(\mathbf{x})](i) = 1 \quad for \quad k, h = 1, \ldots, n. \tag{6}$$

  The probability transition definition depends on the desired direction more precisely on the preferred angle of motion θ_d^p, that will be defined in the next paragraph.

  It is referred to [10,17] for the definition of the preferred angle of motion. Indeed, due to the assumption of the deterministic nature of the speed v, interactions between particles are assumed to modify their dynamics by changing the direction of motion. The assumed walking direction modified by two types of stimuli: (1) the tendency to follow the stream and (2) the attempt to avoid overcrowded areas. These are represented by two unit vectors $\mathbf{e}^{(v)}$, $\mathbf{e}^{(s)}$, respectively. It is expected that at high density pedestrians move in the direction $\mathbf{e}^{(v)}$. Conversely at low density they tend to follow the stream in the direction $\mathbf{e}^{(s)}$. Indeed, the desired direction is defined as follows:

$$\mathbf{e}_d^p = \frac{\sigma \mathbf{e}^{(v)} + (1-\sigma)\mathbf{e}^{(s)}}{\|\sigma \mathbf{e}^{(v)} + (1-\sigma)\mathbf{e}^{(s)}\|_2}, \tag{7}$$

where,

- $\mathbf{e}^{(v)} = (cos(\theta_m), sin(\theta_m))$ with $m = \text{argmin}_{j=1,\ldots,n} \partial_j \rho$, ∂_j is the derivative of ρ in the direction θ_j.
- $\mathbf{e}^{(s)} = (cos(\theta_k), sin(\theta_k))$ defines the direction of the particle.
- The preferred angle θ_d^p which allows pedestrians to follow the stream and to avoid overcrowded areas obtained from (7), through the relation: $\mathbf{e}_d^p = \left(cos(\theta_d^p), sin(\theta_d^p)\right)^T$.

It is assumed that each pedestrian can rest in her/his initial state or change his/her direction, in the clockwise direction or in the opposite clockwise direction in the set S_θ. This means that a pedestrian is located either in the states, $h+1$, $h-1$ or remains in state h. Three cases defined and illustrated in Table 1 are considered according to the position of θ_d^p from θ_h.

The interaction term between pedestrians Γ_i^P is defined as the difference between the gain and the loss of pedestrians moving in the direction θ_i due to the interactions with other pedestrians:

$$\Gamma_i^P(t, \mathbf{x}) = \sum_{h=1}^{n}\sum_{k=1}^{n} \int_V \eta[\rho(t, \mathbf{x}^*)] B_{hk}(i)[\rho(t, \mathbf{x}^*)] f_h(t, \mathbf{x}) f_k\left(t, \mathbf{x}^*\right) d\mathbf{x}^* - f_i(t, \mathbf{x}) \sum_{h=1}^{n} \int_V \eta[\rho(t, \mathbf{x}^*)] f_h\left(t, \mathbf{x}^*\right) d\mathbf{x}^*. \tag{8}$$

2.4.2. The Γ_i^d term

This term is introduced to account for boundary conditions, more precisely to model interactions with walls and obstacles. In this sense, we refer to [7]. The assumptions used in the following are summarized below:

- A pedestrian can change his/her direction, due to
  1. his/her willingness to reach the exit.
  2. the presence of walls or fixed obstacles in front of him/her.

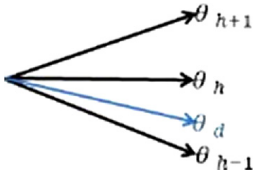
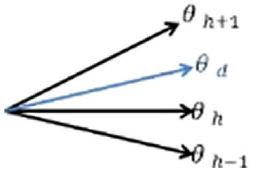
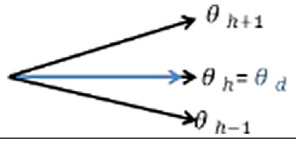
The Γ_i^D term characterizes the interactions between pedestrians and obstacles within the walking area during the evacuation phase. In a similar way as for the Γ_i^P term, the Γ_i^D is modeled by means of the following two interaction terms:

- $P_h(i)$ the probability of the event that the pedestrian changes its direction θ_h into the direction θ_i, due to the presence of walls W, to the obstacle Or and to the exit S, is then introduced. This probability satisfies the following

Table 1
Definition and illustration of the probability term $B_{hk}[\rho](i)$.

Case 1 : $\theta_d^p < \theta_h$	$B_{hk}[\rho](i) = \begin{cases} \sigma\,\rho & \text{if } i = h-1 \\ 1-\sigma\,\rho & \text{if } i = h \end{cases}$
Case 2 : $\theta_h < \theta_d^p$	$B_{hk}[\rho](i) = \begin{cases} 1-\sigma\,\rho & \text{if } i = h \\ \sigma\,\rho & \text{if } i = h+1 \end{cases}$
Case 3 : $\theta_h = \theta_d^p$	$B_{hk}[\rho](i) = \begin{cases} 1 & \text{if } i = h \\ 0 & \text{else} \end{cases}$

Table 2
Definition and illustration of probability's term $P_h(i)$, where $\alpha \in [0, 1]$ sets the quality of the domain.

Case 1 : $\theta_d < \theta_h$	$P_h(i) = \begin{cases} \alpha & \text{if } i = h-1 \\ 1-\alpha & \text{if } i = h \end{cases}$	
Case 2 : $\theta_h < \theta_d$	$P_h(i) = \begin{cases} 1-\alpha & \text{if } i = h \\ \alpha & \text{if } i = h+1 \end{cases}$	
Case 3 : $\theta_h = \theta_d$	$P_h(i) = \begin{cases} 1 & \text{if } i = h \\ 0 & \text{else} \end{cases}$	

relation:

$$\sum_{i=1}^{n} P_h(i) = 1 \quad \textit{for} \quad h = 1, \ldots, n, \tag{9}$$

and its definition depends on the desired direction $\mathbf{e}_d(\mathbf{x})$ for each pedestrian at position $\mathbf{x}$, and more precisely on the angle θ_d that $\mathbf{e}_d(\mathbf{x})$ makes with $\mathbf{e}_\mathbf{x}$.
In the same way, three cases defined and illustrated in Table 2 are considered according to the position of θ_d from θ_h.

- $\mu[\rho(t, \mathbf{x})]$ is an interaction rate, $0 < \mu[\rho(t, \mathbf{x})] \leq 1$, that characterizes the frequency of interactions between the pedestrians and the field. It is a decreasing function of the local density.

The interaction term Γ_i^D between pedestrians and the area characterizes the difference between the gain and the loss of the particles moving in the direction θ_i and is given by:

$$\Gamma_i^D(t, \mathbf{x}) = \mu[\rho(t, \mathbf{x})] \left(\sum_{h=1}^{n} P_h(i) f_h(t, \mathbf{x}) - f_i(t, \mathbf{x}) \right). \tag{10}$$

2.4.3. Determination of the desired direction unit vector $\mathbf{e}_d$

It is then proposed to represent the desired direction unit vector $\mathbf{e}_d$ for each pedestrian by a sum of three vectors: the $\boldsymbol{\gamma}(\mathbf{x})$ vector models the direction which allows the pedestrian to avoid the obstacle, the $\boldsymbol{\tau}(\mathbf{x})$ vector models the pedestrian direction to avoid the walls and finally $\boldsymbol{\nu}(\mathbf{x})$ models the direction which allows the pedestrian to go toward the exit:

$$\mathbf{e}_d(\mathbf{x}) = \frac{\boldsymbol{\gamma}(\mathbf{x}) + \boldsymbol{\tau}(\mathbf{x}) + \boldsymbol{\nu}(\mathbf{x})}{\|\boldsymbol{\gamma}(\mathbf{x}) + \boldsymbol{\tau}(\mathbf{x}) + \boldsymbol{\nu}(\mathbf{x})\|_2}. \tag{11}$$

The directions contribute to the desired direction defined by the previous linear combination defined in Eq. (11). This assumption is a simplification of reality.

To define these three vectors, the domain Ω is decomposed into three zones: the obstacle's influence zone Z, the security zone Z_s to ensure the non collision between pedestrians and the walls, and finally a neutral zone Z_n where pedestrians have only the wish to go toward the exit. These three zones are illustrated in Fig. 3.

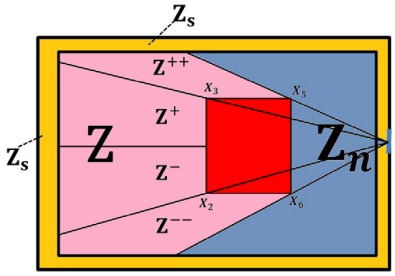


Fig. 3. Representation of the three zones: $Z = Z^+ \cup Z^- \cup Z^{++} \cup Z^{--}$, the obstacle's influence zone for pedestrians, Z_n the neutral zone, and Z_s the security zone.

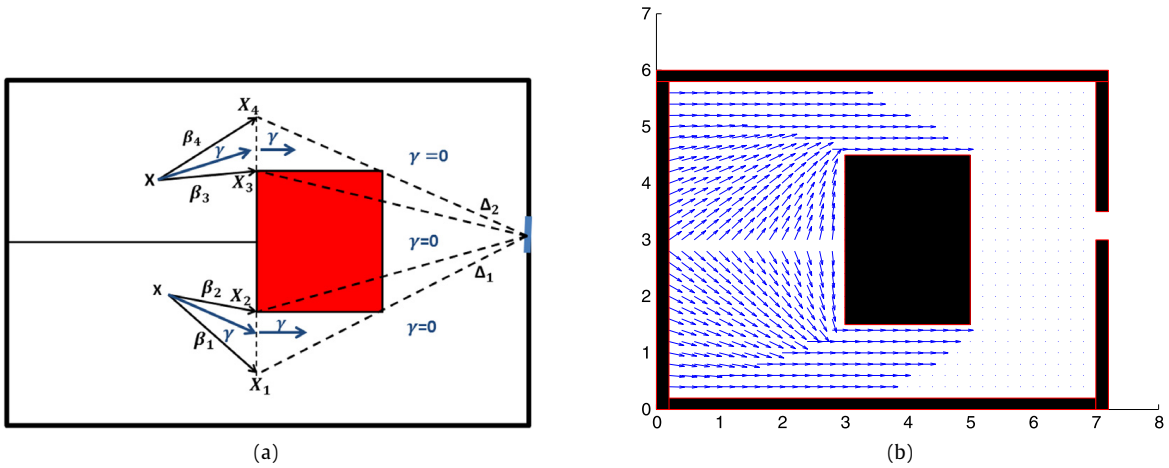


Fig. 4. (a) Illustration of the construction of the $\boldsymbol{\gamma}(\mathbf{x})$ vector which models the direction allowing each pedestrian to avoid the obstacle Or, and (b) the $\boldsymbol{\gamma}$ vector at any point $\mathbf{x}$ of the domain Ω.

The vector field $\boldsymbol{\gamma}(\mathbf{x})$ which models the direction that allows pedestrians to avoid the obstacle, is defined by:

$$\boldsymbol{\gamma}(\mathbf{x}) = \begin{cases} \boldsymbol{\beta}_1(\mathbf{x}) + \boldsymbol{\beta}_2(\mathbf{x}) & \mathbf{x} \in Z^-, \\ \boldsymbol{\beta}_3(\mathbf{x}) + \boldsymbol{\beta}_4(\mathbf{x}) & \mathbf{x} \in Z^+, \\ \mathbf{e}_x & \mathbf{x} \in Z^{++} \cup Z^{--}, \\ 0 & \mathbf{x} \in Z_n \cup Z_s, \end{cases} \tag{12}$$

where:

$$\boldsymbol{\beta}_i(\mathbf{x}) = \frac{\mathbf{x}_i - \mathbf{x}}{\|\mathbf{x}_i - \mathbf{x}\|_2} \quad i = 1, 2, 3, 4.$$

- $\mathbf{x}_2, \mathbf{x}_3$ are two vertices of the obstacle on the opposite side to the exit.
- $\mathbf{x}_1, \mathbf{x}_4$ are respectively the intersection points between the straight line $(\mathbf{x}_2\mathbf{x}_3)$ and two straight lines Δ_1, Δ_2 defining the influence area Z, (see Fig. 4).

$\boldsymbol{\nu}(\mathbf{x})$ is the vector field, which models the direction that allows pedestrians to go toward the exit (see Fig. 5), and is given by the following expression:

$$\boldsymbol{\nu}(\mathbf{x}) = \begin{cases} \mathbf{b}_1(\mathbf{x}) + \mathbf{b}_2(\mathbf{x}) & \textit{if} \quad \mathbf{x} \in Z_n, \\ 0 & \textit{else}, \end{cases} \tag{13}$$

where:

$$\mathbf{b}_i(\mathbf{x}) = \frac{\mathbf{x}_{si} - \mathbf{x}}{\|\mathbf{x}_{si} - \mathbf{x}\|_2} \quad i = 1, 2,$$

$\mathbf{x}_{s1}, \mathbf{x}_{s2}$ are coordinates of the nodes that define the exit.

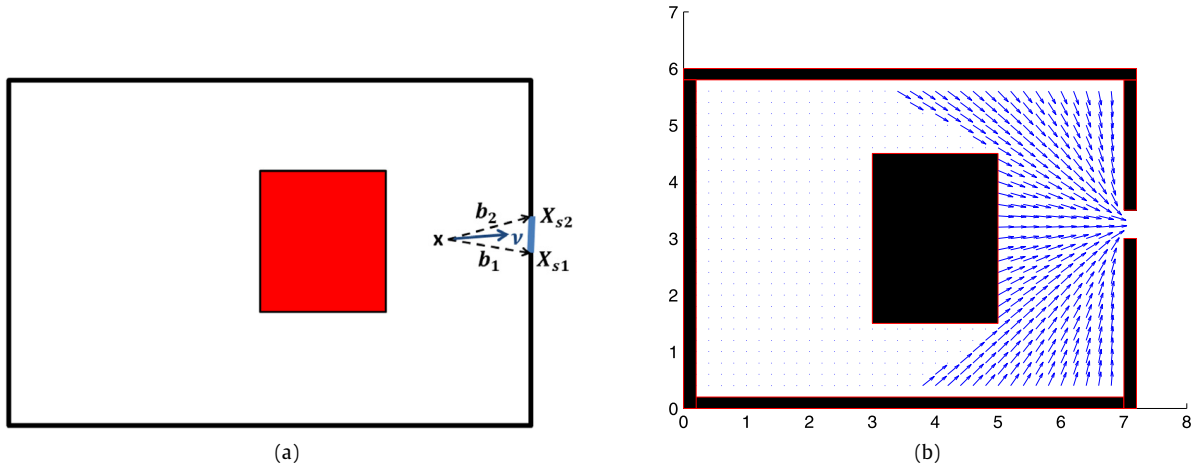


Fig. 5. (a) Illustration of the construction of the $\boldsymbol{\nu}(\mathbf{x})$ vector which models the direction allowing the pedestrian to go to the exit, and (b) the $\boldsymbol{\nu}$ vector in any point $\mathbf{x}$ of the domain Ω.

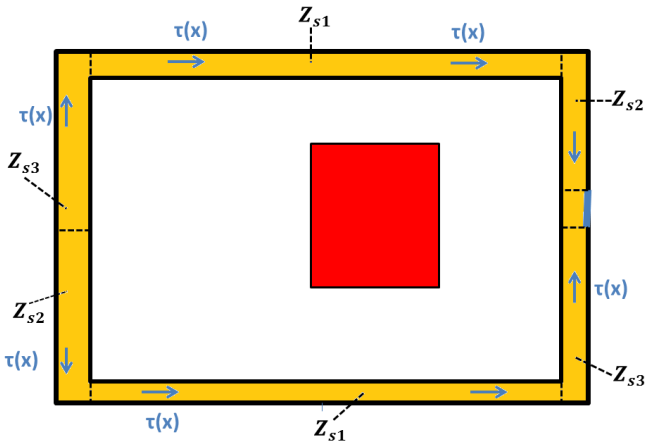


Fig. 6. Illustration of $\boldsymbol{\tau}(\mathbf{x})$ vector which models the direction allowing the pedestrian to avoid the walls.

Finally the field vector $\boldsymbol{\tau}(\mathbf{x})$ which models the direction to avoid collisions between individual pedestrians and the walls is defined as follows:

$$\boldsymbol{\tau}(\mathbf{x}) = \begin{cases} \mathbf{e}_x & \mathbf{x} \in Z_{S1}, \\ -\mathbf{e}_y & \mathbf{x} \in Z_{S2}, \\ \mathbf{e}_y & \mathbf{x} \in Z_{S3}, \\ 0 & else, \end{cases} \tag{14}$$

where the security zone is divided in 3 areas: Z_{S1}, Z_{S2} and Z_{S3} (see Figs. 6 and 7).

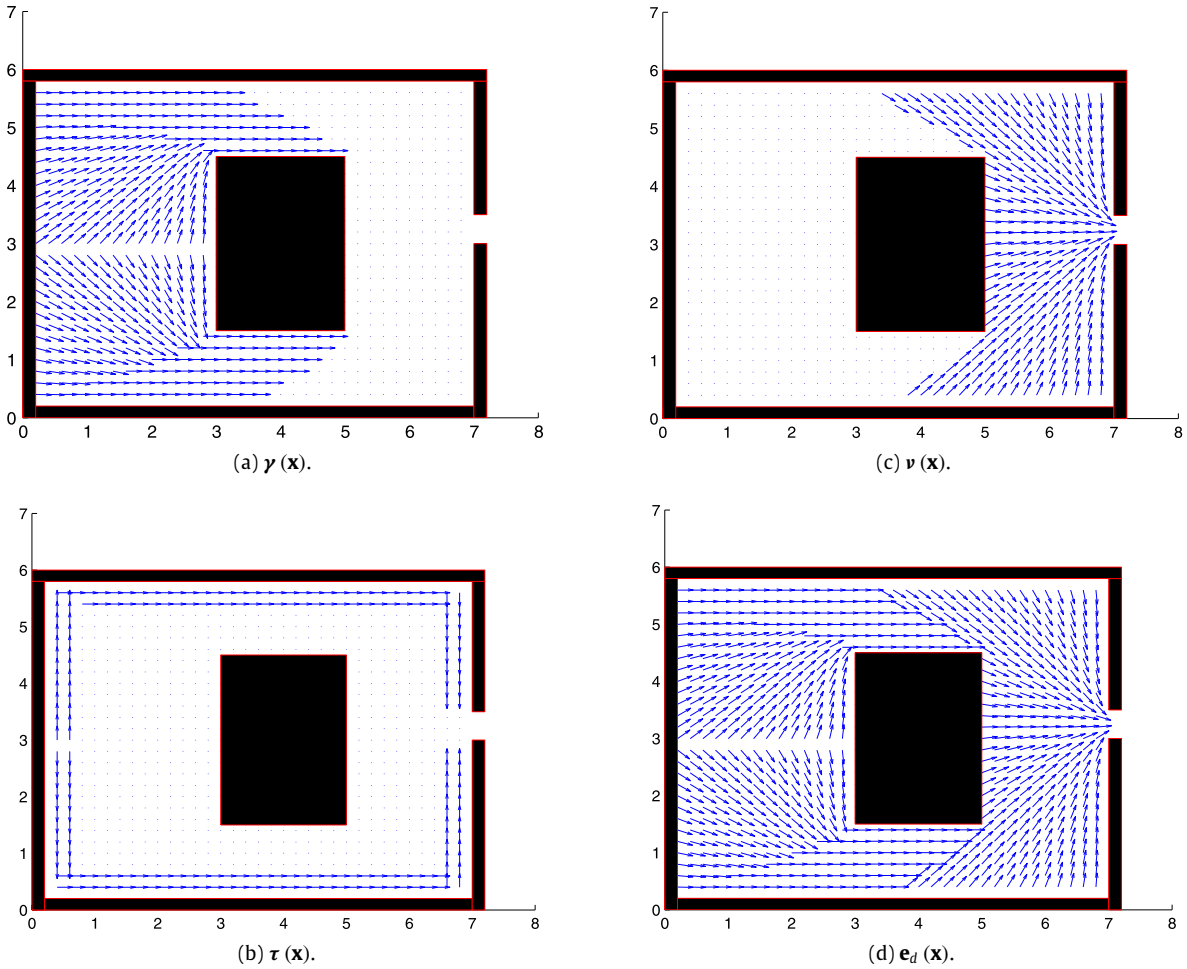
The governing equation

The interaction terms, Γ_i^P defined in Eq. (8) and Γ_i^D defined in Eq. (10) respectively are replaced by their respective expression, and the model partial derivative equation (PDE) (5) becomes:

$$\begin{aligned} \partial_t f_i(t,\mathbf{x}) + \mathbf{v}_i.\partial_{\mathbf{x}} f_i(t,\mathbf{x}) &= \mu[\rho(t,\mathbf{x})] \left(\sum_{h=1}^{n} P_h(i) f_h(t,\mathbf{x}) - f_i(t,\mathbf{x}) \right) \\ &+ \sum_{h=1}^{n}\sum_{k=1}^{n} \int_V \eta[\rho(t,\mathbf{x}^*)] B_{hk}(i)[\rho(t,\mathbf{x}^*)] f_h(t,\mathbf{x}) f_k\left(t,\mathbf{x}^*\right) d\mathbf{x}^* - f_i(t,\mathbf{x}) \sum_{h=1}^{n} \int_V \eta[\rho(t,\mathbf{x}^*)] f_h\left(t,\mathbf{x}^*\right) d\mathbf{x}^*. \end{aligned} \tag{15}$$

The initial conditions expressed as:

$$f_i(t=t_0,\mathbf{x}) = \phi_i(\mathbf{x}) \quad i=1,\ldots,n, \quad \mathbf{x} \in \Omega, \tag{16}$$



(a) $\boldsymbol{\gamma}(\mathbf{x})$. (c) $\boldsymbol{\nu}(\mathbf{x})$.

(b) $\boldsymbol{\tau}(\mathbf{x})$. (d) $\mathbf{e}_d(\mathbf{x})$.

Fig. 7. Illustration of vectors modeling the direction allowing pedestrians to avoid walls, the obstacle and to go toward the exit, (a): $\boldsymbol{\gamma}(\mathbf{x})$, (b): $\boldsymbol{\tau}(\mathbf{x})$, (c): $\boldsymbol{\nu}(\mathbf{x})$ and (d): $\mathbf{e}_d(\mathbf{x})$.

are finally added to the previous PDE (15). After introducing security zones for pedestrians ("pedestrian–wall"), the boundary conditions come from the interactions between pedestrians and the area; therefore, they are included in the management of the interactions.

2.5. Method of characteristics

The method of characteristics is particularly well adapted for solving linear hyperbolic problems and more precisely the transport PDE. Thus we propose to transform the system of Eqs. (15) and (16) from partial differential equations (PDE) to ordinary differential equations (ODE), using the characteristics concept. Indeed, let us consider the characteristic curves associated to the problem defined by Eqs. (15) and (16):

$$\mathbf{X}(t) = (\mathbf{X}_1(t), \mathbf{X}_2(t))^T = (x + t\ v\ cos(\theta_i), y + t\ v\ sin(\theta_i))^T,$$

which is a solution of the following system:

$$\begin{cases} \dfrac{d\mathbf{X}(t)}{dt} = \mathbf{V}_i, \\ \mathbf{X}(0) = (x, y)^T, \end{cases} \tag{17}$$

where: $\mathbf{V}_i = (v\ cos(\theta_i), v\ sin(\theta_i))^T$.

Along these curves, the solution of the system defined by Eqs. (15) and (16) satisfies the following system of ODEs:

$$\begin{cases} \dfrac{d\widehat{f}_i(t, \mathbf{x})}{dt} = \widehat{\Gamma}_i[f, f](t, \mathbf{x}), \\ \widehat{f}_i(0, \mathbf{x}) = \phi_i(\mathbf{x}) \quad \mathbf{x} \in \Omega, \end{cases} \quad for\ \ i = 1, \ldots, n, \tag{18}$$

where: $\widehat{f}_i(t,\mathbf{x}) = f_i(t,\mathbf{X}(t)) = f_i(t, x + v\ cos(\theta_i)t, y + v\ sin(\theta_i)t)$ is the value of f along the characteristics, and $\widehat{\Gamma}_i[f,f](t,\mathbf{x}) = \widehat{\Lambda}_i[f](t,\mathbf{x}) - \mu[\rho(t,\mathbf{x})]\widehat{f}_i(t,\mathbf{x}) + \widehat{\Psi}_i[f,f](t,\mathbf{x}) - \widehat{\Upsilon}_i[f](t,\mathbf{x})\widehat{f}_i(t,\mathbf{x})$, where the operators $\widehat{\Lambda}$, $\widehat{\Psi}$ and $\widehat{\Upsilon}$ are defined for $i = 1,\ldots,n$ by:

$$\widehat{\Lambda}_i[f](t,\mathbf{x}) = \mu[\rho(t,\mathbf{x})]\sum_{h=1}^{n} P_h(i) f_h\left(t, x + v\ \left(cos(\theta_i) - cos(\theta_h)\right)t, y + v\ \left(sin(\theta_i) - sin(\theta_h)\right)t\right),$$

$$\widehat{\Psi}_i[f,f](t,\mathbf{x}) = \sum_{h=1}^{n}\sum_{k=1}^{n}\int_V \eta[\rho(t,\mathbf{x}^*)]B_{hk}(i)[\rho(t,\mathbf{x}^*)] f_h\left(t, x + v\ \left(cos(\theta_i) - cos(\theta_h)\right)t, y + v\ \left(sin(\theta_i) - sin(\theta_h)\right)t\right) \times f_k\left(t, x^* - v\ cos(\theta_k)t, y^* - v\ sin(\theta_k)t\right) d\mathbf{x}^*,$$

and finally,

$$\widehat{\Upsilon}_i[f](t,\mathbf{x}) = \sum_{h=1}^{n}\int_V \eta[\rho(t,\mathbf{x}^*)] f_h\left(t, x^* - v\ cos(\theta_h)t, y^* - v\ sin(\theta_h)t\right) d\mathbf{x}^*.$$

3. Existence and uniqueness of the model solution

This section is devoted to demonstrating proof of the existence and uniqueness of the model solution (18). A first proof of existence and uniqueness was done in [8,15], in a connected domain without obstacles and border walls. Here, an explicit mathematical proof of the model (18) is proposed expressing the crowd's motion in a bounded domain containing obstacles (the same model as in [8], but with a different second member).

To demonstrate the existence and uniqueness of the solution to problems (18), we relied on the proof given in [8]. Indeed it will be proceeded as follows:

(1) Introducing the "mild" solution of the new system of ODE (18).
(2) Using the Banach fixed point theorem, the existence and the uniqueness of a "mild" solution is then proven.

First step:

We introduce the "mild" form of the system (18) obtained by integration along the characteristics, for $i = 1,\ldots,n$,

$$\widehat{f}_i(t,\mathbf{x}) = \phi_i(\mathbf{x}) + \int_0^t \widehat{\Lambda}_i[f](s,\mathbf{x}) - \mu[\rho(s,\mathbf{x})]\widehat{f}_i(s,\mathbf{x}) + \widehat{\Psi}_i[f,f](s,\mathbf{x}) - \widehat{\Upsilon}_i[f](s,\mathbf{x})\widehat{f}_i(s,\mathbf{x})\,ds. \tag{19}$$

Second step, choice of the Banach space:

For a given time t, the functional space defined by:

$$\mathbb{L}^1_n(\Omega) = \left\{\mathbf{f}(t) = (f_1(t),\ldots,f_i(t),\ldots,f_n(t))^T,\ \ \|\mathbf{f}(t)\|_1 = \sum_{i=1}^{n}\int_\Omega |f_i(t,\mathbf{x})|d\mathbf{x} < \infty\right\},$$

is considered.

For a time $T > 0$, let us note the Banach space $\mathbb{X}_T = \mathcal{C}\left([0,T],\mathbb{L}^1_n(\Omega)\right)$, with the norm: $\|\mathbf{f}\|_{\mathbb{X}_T} = sup_{t\in[0,T]}\|\mathbf{f}(t)\|_1$. The following assumptions are then considered :

(A.1): for all positive real $R > 0$ satisfying $0 < \rho < R$, there exist two constants $c_\mu > 0$ and $c_\eta > 0$ such that:

$$\begin{aligned} &0 < \mu(\rho) < c_\mu,\\ &0 < \eta(\rho) < c_\eta. \end{aligned} \tag{20}$$

(A.2): $\mu(\rho)$, $\eta(\rho)$ and $B_{hk}(\rho)$ are Lipschitz functions with respect to the density ρ; namely there are constants L_μ, L_η, L_B such that:

$$\begin{cases} |\mu[\rho_1] - \mu[\rho_2]| \le L_\mu|\rho_1 - \rho_2|, & 0 < \rho_1 < R,\ \ 0 < \rho_2 < R,\\ |\eta[\rho_1] - \eta[\rho_2]| \le L_\eta|\rho_1 - \rho_2|, & 0 < \rho_1 < R,\ \ 0 < \rho_2 < R,\\ |B_{hk}(i)[\rho_1] - B_{hk}(i)[\rho_2]| \le L_B|\rho_1 - \rho_2|, & i,h,k = 1,\ldots,n. \end{cases} \tag{21}$$

The main two theorems for the local and global existence, respectively are given in the following:

Theorem 1 (*Local Existence*). *Let $\boldsymbol{\phi} = (\phi_1,\ldots,\phi_i,\ldots,\phi_n)^T \in \mathbb{L}^\infty_n(\Omega)$ with $\boldsymbol{\phi} \ge 0$, there exists φ^0, a time $T > 0$ and two constants a_0, R, such that if $\|\boldsymbol{\phi}\|_1 \le \varphi^0$, the problem (18) has a unique positive solution $\mathbf{f} = (f_1,\ldots,f_i,\ldots,f_n)^T \in \mathbb{X}_T$, satisfying:*

$$\|\mathbf{f}\|_{\mathbb{X}_T} \le a_0\|\boldsymbol{\phi}\|_1,$$

$$\rho\,(t,\mathbf{x}) \leq R, \quad \forall\ t \ \in [0,T], \quad \forall\ \mathbf{x} \in \Omega.$$

Moreover, if

$$\sum_{i=1}^{n} \|\phi_i\|_\infty \leq 1, \tag{22}$$

we have:

$$\rho\,(t,\mathbf{x}) \leq 1, \quad \forall\ t \ \in [0,T], \quad \forall\ \mathbf{x} \in \Omega. \tag{23}$$

Theorem 2 (*Global Existence*). *Considering the same assumptions as in Theorem 1, there exists* φ^p, $(p = 1, \ldots, m-1, m \in \mathbb{N})$, a_p, $(p = 1, \ldots, m-1, m \in \mathbb{N})$, *such that if* $\|\boldsymbol{\phi}\|_1 \leq \varphi^p$, *the problem (18) has a unique maximum positive solution* $\mathbf{f} \in \mathcal{C}\left([0,(m-1)T], \mathbb{L}^1_n(\Omega)\right)$, *satisfying for any* $p \leq m-1$,

$$sup_t \|\mathbf{f}\,(t+(p-1)T)\,\|_1 \leq a_{p-1}\|\boldsymbol{\phi}\|_1, \quad t \in [0,T],$$

$$\rho\,(t+(p-1)T,\mathbf{x}) \leq R, \quad \forall\ t \in [0,T], \quad \forall\ \mathbf{x} \in \Omega.$$

In addition, if ϕ *satisfies (22), we have:*

$$\rho\,(t+(p-1)T,\mathbf{x}) \leq 1, \quad \forall\ t \ \in [0,T], \quad \forall\ \mathbf{x} \in \Omega.$$

Proof of Theorem 1. The following function:

$$\psi_i(t,\mathbf{x}) = f_i\,(t,\mathbf{x})\,exp\,(\lambda t) \quad for\ \ i = 1, \ldots, n, \quad \lambda > 0,$$

is first introduced. The system (18) is equivalent to the following system:

$$\begin{cases} \dfrac{d\widehat{\psi}_i\,(t,\mathbf{x})}{dt} = \lambda\widehat{\psi}_i\,(t,\mathbf{x}) + \widehat{\Lambda}_i[\psi]\,(t,\mathbf{x}) - \widehat{\psi}_i\mu[\rho(t,\mathbf{x})] + exp\,(-\lambda t)\left[\widehat{\Psi}_i[\psi,\psi]\,(t,\mathbf{x}) - \widehat{\Upsilon}_i[\psi]\,(t,\mathbf{x})\,\widehat{\psi}_i\right], \\ \widehat{\psi}_i\,(0,\mathbf{x}) = \phi_i\,(\mathbf{x}) \quad \mathbf{x} \in \Omega \quad for\ \ i = 1, \ldots, n. \end{cases} \tag{24}$$

For all $t \in [0,T]$, we integrate (24). Then, the following "mild" formulation, is deduced:

$$\begin{aligned} \widehat{\psi}_i\,(t,\mathbf{x}) = \phi_i\,(\mathbf{x}) + \int_0^t \Big[& exp(-\lambda s)\widehat{\Psi}_i[\psi,\psi]\,(s,\mathbf{x}) + \widehat{\Lambda}_i[\psi]\,(s,\mathbf{x}) \\ & + \left(\lambda - \mu[\rho(s,\mathbf{x})] - \widehat{\Upsilon}_i[\psi]\,(s,\mathbf{x})\,exp(-\lambda s)\right)\widehat{\psi}_i(s,\mathbf{x})\Big]\,ds. \end{aligned} \tag{25}$$

Let us consider the operator $\mathbf{A} = \left(\widehat{A(\psi)}_1, \ldots, \widehat{A(\psi)}_i, \ldots, \widehat{A(\psi)}_n\right)^T$, and its i-th component defined by:

$$\begin{aligned} \widehat{A(\psi)}_i\,(t,\mathbf{x}) = \phi_i\,(\mathbf{x}) + \int_0^t \Big[& exp(-\lambda s)\widehat{\Psi}_i[\psi,\psi]\,(s,\mathbf{x}) + \widehat{\Lambda}_i[\psi]\,(s,\mathbf{x}) \\ & + \left(\lambda - \mu[\rho(s,\mathbf{x})] - \widehat{\Upsilon}_i[\psi]\,(s,\mathbf{x})\,exp(-\lambda s)\right)\widehat{\psi}_i\,(s,\mathbf{x})\Big]\,ds. \end{aligned} \tag{26}$$

To show that the system (24) has a solution, it is sufficient to show that the operator $\mathbf{A}$ has a unique fixed point in the Banach space $\mathbb{X}_T$. Indeed, let us introduce the set defined by:

$$\begin{aligned} B_{T,a_0,\lambda,R} = \Bigg\{ & \boldsymbol{\psi} = (\psi_1, \ldots, \psi_i, \ldots, \psi_n)^T \in \mathbb{X}_T : \ \psi_i \geq 0, \ \ \|\boldsymbol{\psi}\|_{\mathbb{X}_T} \leq a_0\|\boldsymbol{\phi}\|_1, \\ & \sum_{i=1}^{n} \psi_i\,(t, x - v\ cos(\theta_i)t, y - v\ sin(\theta_i)t) \leq Rexp(\lambda t), \ \ t \in [0,T], \ \ \mathbf{x} \in \Omega \Bigg\}. \end{aligned}$$

The operator $\mathbf{A}$ has a unique fixed point if the following two properties are satisfied:

- (P.1) Let $\boldsymbol{\psi} \in B_{T,a_0,\lambda,R}$ then, $\mathbf{A}\boldsymbol{\psi} \in B_{T,a_0,\lambda,R}$.
- (P.2) The application $\mathbf{A} : B_{T,a_0,\lambda,R} \to B_{T,a_0,\lambda,R}$ is a contraction.

In what follows, the constants T, a_0, λ, R, must be chosen carefully in order to obtain properties (P.1) and (P.2). The proof of both properties (P.1) and (P.2) is based on the following lemma:

Lemma 1. *Let* $T > 0$, $\lambda > 0$, $\boldsymbol{\psi}^1 = \left(\psi_1^1, \ldots, \psi_i^1, \ldots, \psi_n^1\right)^T \in \mathbb{X}_T$ *and* $\boldsymbol{\psi}^2 = \left(\psi_1^2, \ldots, \psi_i^2, \ldots, \psi_n^2\right)^T \in \mathbb{X}_T$ *such that :*

$$\sum_{i=1}^{n} \psi_i^j\,(t, x - v\,cos(\theta_i)t, y - v\,sin(\theta_i)t) \leq Rexp(\lambda t)\ \ t \in [0, T],\ \ \mathbf{x} \in \Omega\ \ for\ \ j = 1, 2,\ \ \lambda > 0,\ \ R > 0. \tag{27}$$

Then,

1. *There are* $C_1 > 0$, $C_2 > 0$ *such that,*

$$\|\mathbf{A}\boldsymbol{\psi}^1 - \mathbf{A}\boldsymbol{\psi}^2\|_{\mathbb{X}_T} \leq \left[\frac{C_1}{\lambda}\left(\|\boldsymbol{\psi}^1\|_{\mathbb{X}_T} + \|\boldsymbol{\psi}^2\|_{\mathbb{X}_T}\right) + (\lambda + C_2)\,T\right]\|\boldsymbol{\psi}^1 - \boldsymbol{\psi}^2\|_{\mathbb{X}_T}, \tag{28}$$

where:

$$\begin{aligned} C_1 &= 2c_\eta + R\left(nc_\eta L_B + (n+1)L_\eta\right), \\ C_2 &= 2c_\mu + (n+1)RL_\mu. \end{aligned} \tag{29}$$

2. *If* $\psi_i(t, \mathbf{x}) \geq 0$ *and* $\phi_i(\mathbf{x}) \geq 0$ *then, there exists* λ_0 *such that:* $\left(\widehat{A\psi}\right)_i(t, \mathbf{x}) \geq 0$ *for all* $t \in [0, T]$, $\mathbf{x} \in \Omega$ *and* $\lambda \geq \lambda_0$, $i = 1, \ldots, n$, *where:*

$$\begin{aligned} &\lambda_0 = Rc_\eta|V| + c_\mu, \\ &|V| \text{: is the measure of visibility zone } V, \quad |V| = \int_V d\mathbf{x}. \end{aligned} \tag{30}$$

3. *There exist* R_1 *and* T *such that, for all* $R \geq R_1$ *and* $t \leq T$, *we get:*

$$\sum_{i=1}^{n}(A\psi)_i\,(t, x - v\,cos(\theta_i)t, y - v\,sin(\theta_i)t) \leq Rexp(\lambda t)\ \ t \in [0, T_0],\ \ \mathbf{x} \in \Omega. \tag{31}$$

where:

$$R_1 = \sum_{i=1}^{n} \|\phi_i\|_\infty, \tag{32}$$

$$T = \frac{1}{\lambda} ln\left(1 + \frac{\lambda}{nR\left(c_\eta R|V| + c_\mu\right)}\left(R - \sum_{i=1}^{n}\|\phi_i\|_\infty\right)\right) := T_0. \tag{33}$$

4.

$$\|\mathbf{A}\boldsymbol{\psi}\|_{\mathbb{X}_T} \leq \|\phi\|_1 + \frac{C_1}{\lambda}\|\boldsymbol{\psi}\|_{\mathbb{X}_T}^2 + (\lambda + C_2)\,T\|\boldsymbol{\psi}\|_{\mathbb{X}_T}. \tag{34}$$

The proof of this lemma is given in Appendix A.

Let $\boldsymbol{\psi}^1$, $\boldsymbol{\psi}^2 \in B_{T,a_0,\lambda,R}$, and T be defined by:

$$T = \frac{1}{C_2 + \lambda} T_0 \leq T_0, \tag{35}$$

where T_0 is the expression of time defined in Eq. (33), and C_2 is defined in Eq. (29).

According to the inequality (34),

$$\|\mathbf{A}\boldsymbol{\psi}\|_{\mathbb{X}_T} \leq \|\boldsymbol{\phi}\|_1 + \frac{C_1}{\lambda}\|\boldsymbol{\psi}\|_{\mathbb{X}_T}^2 + (\lambda + C_2)\,T\|\boldsymbol{\psi}\|_{\mathbb{X}_T}, \tag{36}$$

since $\boldsymbol{\psi} \in B_{T,a_0,\lambda,R}$; we have $\|\boldsymbol{\psi}\|_{\mathbb{X}_T} \leq a_0\|\boldsymbol{\phi}\|_1$, moreover,

$$\|\mathbf{A}\boldsymbol{\psi}\|_{\mathbb{X}_T} \leq \|\boldsymbol{\phi}\|_1 + \frac{1}{\lambda}C_1 a_0^2\|\boldsymbol{\phi}\|_1^2 + (C_2 + \lambda)\,Ta_0\|\boldsymbol{\phi}\|_1. \tag{37}$$

Then, $\|\mathbf{A}\boldsymbol{\psi}\|_{\mathbb{X}_T} \leq a_0\|\boldsymbol{\phi}\|_1$, if a_0 satisfies the following equation:

$$(E_0) : \|\boldsymbol{\phi}\|_1 + \frac{1}{\lambda}C_1 a_0^2\|\boldsymbol{\phi}\|_1^2 + (C_2 + \lambda)\,Ta_0\|\boldsymbol{\phi}\|_1 = a_0\|\boldsymbol{\phi}\|_1.$$

Indeed, the constant a_0 exists if:

$$\|\boldsymbol{\phi}\|_1 \leq min\left\{\frac{(n-1)^2 c_\mu}{8n^2 c_\eta}, \frac{(n-1)^2 c_\eta |V|}{\left(4n^2\left(nc_\eta L_B + (n+1)L_\eta\right)\right)}\right\} := \varphi^0, \quad \text{(for the proof see Appendix B).}$$

which leads to the existence of a_0 if $\|\boldsymbol{\phi}\|_1$ is "small".

In addition, from Lemma 1 for all $\|\boldsymbol{\phi}\|_1 \leq \varphi^0$, $R \geq R_1$, $\lambda = \lambda_0$, $t \leq T \leq T_0$ and a_0 the smallest solution among the two positive solutions of equation (E_0), where R_1, λ_0, and T, are given by (32), (30), (35), respectively, we get:

$$\textit{if} \quad \boldsymbol{\psi} \in B_{T,a_0,\lambda,R} \quad \textit{then} \quad \mathbf{A}\boldsymbol{\psi} \in B_{T,a_0,\lambda,R} \quad \forall t \in [0,T].$$

This ends the proof of property (P.1).

On the other hand, from Lemma 1, we have:

$$\|\mathbf{A}\boldsymbol{\psi}^1 - \mathbf{A}\boldsymbol{\psi}^2\|_{\mathbb{X}_T} \leq \left(\frac{2C_1}{\lambda} a_0 \|\boldsymbol{\phi}\|_1 + (C_2+\lambda)\,T\right) \|\boldsymbol{\psi}^1 - \boldsymbol{\psi}^2\|_{\mathbb{X}_T}. \tag{38}$$

Let a_0, the small positive solution of equation (E_0) be defined by:

$$a_0 = \lambda \frac{(1-(\lambda+C_2)\,T) - \sqrt{\Delta_0}}{2C_1\|\boldsymbol{\phi}\|_1}, \tag{39}$$

where Δ_0 is the discriminant of equation (E_0)

$$\Delta_0 = ((C_2+\lambda)\,T - 1)^2 - 4\frac{C_1}{\lambda}\|\boldsymbol{\phi}\|_1. \tag{40}$$

Hence,

$$0 \leq \left(\frac{2C_1}{\lambda} a_0 \|\boldsymbol{\phi}\|_1 + (C_2+\lambda)\,T\right) = 1 - \sqrt{\Delta_0} < 1.$$

From there, the operator $\mathbf{A} : B_{T,a_0,\lambda,R} \to B_{T,a_0,\lambda,R}$ is a contraction. This ends the proof of property (P.2).

The fixed point theorem ends the proof of Theorem 1, which refers to local existence.

From the foregoing, there exist φ^0, λ_0, T, a_0 and R, such that the problem (18) has a unique positive solution $\mathbf{f} = (f_1, \ldots, f_i, \ldots, f_n)^T \in \mathbb{X}_T$, satisfying:

$$\rho\,(t, \mathbf{x}) \leq R, \quad \forall\ t \ \in [0,T], \quad \forall\ \mathbf{x} \in \Omega.\ \forall\ R \geq R_1,$$

where

$$R_1 = \sum_{i=1}^{n} \|\phi_i\|_\infty.$$

Moreover if $\boldsymbol{\phi}$ satisfies (22) ($R_1 \leq 1$), then one can choose R such that (23) can be obtained; for example, the R can be chosen as:

$$R = \frac{1 + \sum_{i=1}^{n} \|\phi_i\|_\infty}{2} > R_1. \tag{41}$$

This completes the proof.

Proof of Theorem 2. To prove the existence for large times of the problem's solution (18), it is equivalent to show that the solution obtained by Theorem 1 admits an extension on each interval $[0, pT]$ *for* $p \in \mathbb{N}$.

Let $\boldsymbol{\phi}$ satisfying the conditions of Theorem 1. R, $\lambda = \lambda_0$ and T are given by (41), (30), (35), respectively, and are fixed and depending on initial data $\boldsymbol{\phi}$.

Consequently, there exists φ^1 such that if $\|\boldsymbol{\phi}\|_1 \leq \varphi^1$, there exists a_1 such that the solution to (18) obtained by Theorem 1 can be extended in the interval $[T, 2T]$ and satisfies the estimates

$$\|\boldsymbol{\psi}(t+T)\|_{\mathbb{X}_T} \leq a_1\|\boldsymbol{\phi}\|_1, \quad \forall\ t \ \in [0,T], \tag{42}$$

$$\rho\,(t+T, \mathbf{x}) \leq R, \quad \forall\ t \ \in [0,T], \quad \forall\ \mathbf{x} \in \Omega. \tag{43}$$

Indeed, the problem (24) on $[T, 2T]$ with a given initial condition defined by $\boldsymbol{\psi}(T, \mathbf{x})$ can be solved, more precisely $\forall t \in [0, T]$ we have :

$$\widehat{\psi}_i\,(t+T, \mathbf{x}) = \psi_i\,(T, \mathbf{x}) + \int_0^t \Big[exp(-\lambda(s+T))\widehat{\Psi}_i[\psi, \psi]\,(s+T, \mathbf{x}) + \widehat{\Lambda}_i[\psi]\,(s+T, \mathbf{x}) + \left(\lambda - \mu[\rho(s+T, \mathbf{x})] - \widehat{\Upsilon}_i[\psi]\,(s+T, \mathbf{x})\,exp(-\lambda(s+T))\right)\widehat{\psi}_i\,(s+T, \mathbf{x})\Big] ds, \text{ for } i = 1, \ldots, n.$$

Let us consider the set defined by:

$$B_{T,a_1,\lambda,R} = \Big\{ \boldsymbol{\psi}(t+T) = (\psi_1(t+T), \ldots, \psi_i(t+T), \ldots, \psi_n(t+T))^T \in \mathbb{X}_T : \ \psi_i \geq 0, \ \ \|\boldsymbol{\psi}(t+T)\|_{\mathbb{X}_T} \leq a_1 \|\boldsymbol{\phi}\|_1, \\ \sum_{i=1}^{n} \psi_i \left(t+T, x - v\ cos(\theta_i)(t+T), y - v\ sin(\theta_i)(t+T)\right) \leq Rexp(\lambda(t+T)) \ \ t \in [0,T], \mathbf{x} \in \Omega \Big\}.$$

where R, λ, T, a_0 are fixed and depending on initial data $\boldsymbol{\phi}$.

In the same way as in the case of local existence, it can be shown that, for $\boldsymbol{\psi}(t+T) \geq 0, \ \ t \in [0,T]$, R given by (41), $\lambda = \lambda_0$, and T given by (35), we get:

$$\left(\widehat{A\psi}\right)_i (t+T, \mathbf{x}) \geq 0, \ \ i = 1, \ldots, n, \ \ t \in [0,T], \ \ \mathbf{x} \in \Omega,$$

and

$$\sum_{i=1}^{n} (A\psi)_i \left(t+T, x - v\ cos(\theta_i)(t+T), y - v\ sin(\theta_i)\right) \leq Rexp(\lambda(t+T)) \ \ i = 1, \ldots, n, \ \ t \in [0,T], \ \ \mathbf{x} \in \Omega.$$

In addition, it can be shown that:

$$\|\mathbf{A}\boldsymbol{\psi}(t+T)\|_{\mathbb{X}_T} \leq a_0 \|\boldsymbol{\phi}\|_1 + \frac{1}{\lambda} C_1 a_1^2 \|\boldsymbol{\phi}\|_1^2 + (C_2 + \lambda)\, T a_1 \|\boldsymbol{\phi}\|_1. \tag{44}$$

Then, $\|\mathbf{A}\boldsymbol{\psi}(t+T)\|_{\mathbb{X}_T} \leq a_1 \|\boldsymbol{\phi}\|_1$, if a_1 satisfies the following equation:

$$(E_1): \ a_0 \|\boldsymbol{\phi}\|_1 + \frac{1}{\lambda} C_1 a_1^2 \|\boldsymbol{\phi}\|_1^2 + (C_2 + \lambda)\, T a_1 \|\boldsymbol{\phi}\|_1 = a_1 \|\boldsymbol{\phi}\|_1.$$

Moreover, there exists a solution of equation (E_1) in a_1 if its discriminant $\Delta_1 = ((C_2+\lambda)\,T - 1)^2 - 4\frac{C_1}{\lambda} a_0 \|\boldsymbol{\phi}\|_1$ is positive.

With a choice of $\boldsymbol{\phi}$ such that: $\|\boldsymbol{\phi}\|_1 \leq min\left(\varphi^0, \lambda \frac{((C_2+\lambda)T-1)^2}{4C_1 a_0}\right) := \varphi^1$, we have: $\Delta_1 \geq 0$.

From there, for $\|\boldsymbol{\phi}\|_1 \leq \varphi^1$, $t \in [0,T]$ with T is given by (35), R given by (41), a_1 is a positive solution of equation (E_1), we have:

$$if \ \ \boldsymbol{\psi}(t+T) \in B_{T,a_1,\lambda,R} \ \ then \ \ \mathbf{A}\boldsymbol{\psi}(t+T) \in B_{T,a_1,\lambda,R} \ \ \forall t \in [0,T].$$

In addition, let $\boldsymbol{\psi}_1(t+T)$, $\boldsymbol{\psi}_2(t+T) \ \in B_{T,a_1,\lambda,R}$, then:

$$\|\mathbf{A}\boldsymbol{\psi}_1(t+T) - \mathbf{A}\boldsymbol{\psi}_2(t+T)\|_{\mathbb{X}_T} \leq \left(\frac{2C_1}{\lambda} a_1 \|\boldsymbol{\phi}\|_1 + (C_2+\lambda)\, T\right) \|\boldsymbol{\psi}_1(t+T) - \boldsymbol{\psi}_2(t+T)\|_{\mathbb{X}_T}. \tag{45}$$

Let us consider a_1 the smallest solution among the two positive solutions of the equation (E_1), given by: $a_1 = \lambda \frac{(1-(\lambda+C_2)T) - \sqrt{\Delta_1}}{2C_1 \|\boldsymbol{\phi}\|_1} \geq 0$, since

$$0 \leq \left(\frac{2C_1}{\lambda} a_1 \|\boldsymbol{\phi}\|_1 + (C_2+\lambda)\, T\right) = 1 - \sqrt{\Delta_1} < 1.$$

By applying the fixed point theorem which gives the existence and uniqueness of solution $\boldsymbol{\psi}(t+T) \in \mathbb{X}_T$, more precisely existence of a solution in $[T, 2T]$. This solution is continuous in $[T, 2T]$ and, in particular, it satisfies (42) and (43). Moreover, if $\boldsymbol{\phi}$ satisfies (22) ($\sum_{i=1}^{n} \|\phi_i\|_\infty \leq 1$), for R given by (41), we have,

$$\rho\,(t, \mathbf{x}) \leq 1, \ \ \forall\ t \ \in [T, 2T], \ \ \forall\ \mathbf{x} \in \Omega.$$

The iteration process can be applied to prove the existence for large times. Suppose that the solution exists and is continuous in $[0, (m-1)T]$ and satisfies: $\|\psi(t+(p-1)T)\|_{\mathbb{X}_T} \leq a_{p-1} \|\boldsymbol{\phi}\|_1$, *for* $p = 1, \ldots, m$, $t \in [0,T]$, $\|\psi(pT)\|_{\mathbb{X}_T} \leq a_{p-1} \|\boldsymbol{\phi}\|_1$, *for* $p = 1, \ldots, m$, and $\rho\,(t + (p-1)T, \mathbf{x}) \leq R$, *for* $p = 1, \ldots, m$, $t \ \in [0,T]$ $\mathbf{x} \in \Omega$. where, a_0, Δ_0, R, λ, T are given by (39), (40), (41), (30), (35) respectively.

a_p and Δ_p for $p = 1, \ldots, m-1$ are given by:

$$a_p = \lambda \frac{(1 - (\lambda + C_2)\, T) - \sqrt{\Delta_p}}{2C_1 \|\boldsymbol{\phi}\|_1},$$

$$\Delta_p = ((C_2 + \lambda)\, T - 1)^2 - 4\frac{C_1}{\lambda} a_{p-1} \|\boldsymbol{\phi}\|_1.$$

It can now be proved that we can extend the solution in $[(m-1)T, mT]$, satisfying the following inequality:

$$\|\boldsymbol{\psi}\,(t + (m-1)T)\,\|_1 \leq a_{m-1} \|\boldsymbol{\phi}\|_1, \ \ t \in [0,T], \tag{46}$$

$$\|\boldsymbol{\psi}\,(mT)\,\|_1 \leq a_{m-1} \|\boldsymbol{\phi}\|_1, \tag{47}$$

$$\rho\left(t+(m-1)T,\mathbf{x}\right)\leq R,\quad t\ \in[0,T]\ \ \mathbf{x}\in\Omega. \tag{48}$$

Let $\boldsymbol{\psi}=(\psi_i)_{i=1,\ldots,n}$ be the solution of the following problem, for $i=1,\ldots,n$

$$\widehat{\psi}_i\left(t+(m-1)T,\mathbf{x}\right)=\widehat{\psi}_i\left((m-1)T,\mathbf{x}\right)+\int_{(m-1)T}^{t+(m-1)T}\Big[exp(-\lambda(s))\widehat{\Psi}_i[\psi,\psi]\left(s,\mathbf{x}\right)+\widehat{\Lambda}_i[\psi]\left(s,\mathbf{x}\right)$$
$$+\left(\lambda-\mu[\rho(s,\mathbf{x})]-\widehat{\Upsilon}_i[\psi]\left(s,\mathbf{x}\right)exp(-\lambda(s))\right)\widehat{\psi}_i\left(s,\mathbf{x}\right)\Big]ds.$$

In the same way as in the case of local existence, it can be shown that:

$$\|\mathbf{A}\boldsymbol{\psi}\left(t+(m-1)T\right)\|_{\mathbb{X}_T}\leq a_{m-2}\|\boldsymbol{\phi}\|_1+\frac{1}{\lambda}C_1a_{m-1}^2\|\boldsymbol{\phi}\|_1^2+\left(C_2+\lambda\right)Ta_{m-1}\|\boldsymbol{\phi}\|_1,$$
$$\|\mathbf{A}\boldsymbol{\psi}^1\left(t+(m-1)T\right)-\mathbf{A}\boldsymbol{\psi}^2\left(t+(m-1)T\right)\|_{\mathbb{X}_T}$$
$$\leq\left(\frac{2C_1}{\lambda}a_{m-1}\|\boldsymbol{\phi}\|_1+\left(C_2+\lambda\right)T\right)\|\boldsymbol{\psi}^1\left(t+(m-1)\right)-\boldsymbol{\psi}^2\left(t+(m-1)T\right)\|_{\mathbb{X}_T}.$$

If we choose a_{m-1} so that,

$$a_{m-1}=\lambda\frac{\left(1-\left(\lambda+C_2\right)T\right)-\sqrt{\Delta_{m-1}}}{2C_1\|\boldsymbol{\phi}\|_1}, \tag{49}$$

where,

$$\Delta_{m-1}=\left(\left(C_2+\lambda\right)T-1\right)^2-4\frac{C_1}{\lambda}a_{m-2}\|\boldsymbol{\phi}\|_1. \tag{50}$$

Then a_{m-1} is a solution of

$$(E_{m-1}):\|\boldsymbol{\phi}\|_1+\frac{1}{\lambda}C_1a_{m-1}^2\|\boldsymbol{\phi}\|_1^2+\left(C_2+\lambda\right)Ta_{m-1}\|\boldsymbol{\phi}\|_1=a_{m-1}\|\boldsymbol{\phi}\|_1.$$

Moreover

$$0\leq\left(\frac{2C_1}{\lambda}a_{m-1}\|\boldsymbol{\phi}\|_1+\left(C_2+\lambda\right)T\right)=1-\sqrt{\Delta_{m-1}}<1,$$

which gives the solution in $[(m-1)T,mT]$, satisfying (46)–(48). Moreover, if $\boldsymbol{\phi}$ satisfies (22) ($\sum_{i=1}^n\|\phi_i\|_\infty\leq 1$), for R given by (41), we have:

$$\rho\left(t,\mathbf{x}\right)\leq 1,\quad\forall\ t\ \in[(m-1)T,mT],\quad\forall\ \mathbf{x}\in\Omega.$$

This completes the proof of Theorem 2.

4. Simulations

4.1. Numerical method

The numerical method used in this work is the same numerical scheme described in [7,8], is based on the so called splitting method [20] to solve the proposed mathematical model defined by:

$$\begin{cases}\partial_t f_i\left(t,\mathbf{x}\right)+\mathbf{v}_i.\partial_{\mathbf{x}}f_i\left(t,\mathbf{x}\right)=\Gamma_i[f]\left(t,\mathbf{x}\right)\ \ i=1,\ldots,n,\ \ t\in[0,T],\\ f_i\left(0,\mathbf{x}\right)=\phi_i(\mathbf{x})\ \ \ i=1,\ldots,n,\ \ \mathbf{x}\in\Omega,\end{cases} \tag{51}$$

on an interval of time $[0,T]$. The principle of the method consists in introducing a subdivision $[t_k]_{k=0,\ldots,T}$ of the interval $[0,T]$, and solving on each subdivision $[t_k,t_{k+1}]$ the two following systems:

$$\begin{cases}\partial_t\widehat{f}_i\left(t,\mathbf{x}\right)+\mathbf{v}_i.\partial_{\mathbf{x}}\widehat{f}_i\left(t,\mathbf{x}\right)=0\ \ i=1,\ldots,n,\ \ t\in[t_k,t_{k+\frac{1}{2}}],\\ \widehat{f}_i\left(0,\mathbf{x}\right)=\phi_i(\mathbf{x})\ \ \ i=1,\ldots,n,\ \ \mathbf{x}\in\Omega,\end{cases} \tag{52}$$

where $\phi_i(\mathbf{x})$ is the initial distribution of pedestrians, and

$$\begin{cases}\partial_t f_i\left(t,\mathbf{x}\right)=\Gamma_i[f]\left(t,\mathbf{x}\right)\ \ i=1,\ldots,n,\ \ t\in[t_{k+\frac{1}{2}},t_{k+1}],\\ f_i\left(0,\mathbf{x}\right)=\widehat{f}_i(\mathbf{x})\ \ \ i=1,\ldots,n,\ \ \mathbf{x}\in\Omega,\end{cases} \tag{53}$$

where $\widehat{f}_i(\mathbf{x})$ is the solution of system (52).

For the resolution of the transport equation (52), the upwind scheme of order 1 [21,22] is used, and for the ordinary differential equation (53), Euler's method of order 1 is preferred.

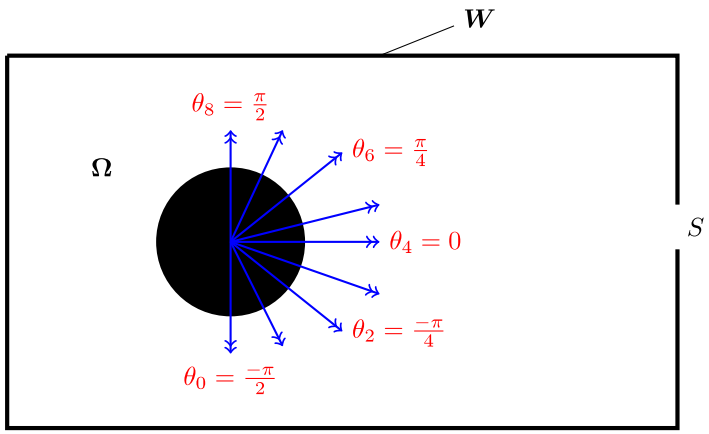


Fig. 8. Illustration of the nine directions defined by: $\theta_k = \frac{k-4}{8}\pi : \ k = 0, \ldots, 8$, that each pedestrian can take.

4.2. Numerical results and application to the evacuation of a room

Several authors [1] reported that an obstacle can either facilitate or obstruct pedestrian evacuation of a room with an exit, depending on its position, size and shape. In particular it has been shown that an obstacle may have a strong influence on pedestrians if it is located close to the exit.

In order to show the performance of the proposed mathematical model, it has been applied to the evacuation of a room with and without an obstacle and our study is restricted to the mathematical model without taking into account the internal interactions between the pedestrians. This choice is motivated by considering a small number of pedestrians ($N \leq 50$ pedestrians). Our goal is to study the influence of an obstacle on evacuation time and describe some characteristics of the dynamics of a crowd in a bounded domain (with walls, size of the exit and dimensions of the obstacle) namely, avoidance and evacuation.

Therefore the mathematical model (51) becomes:

$$\begin{cases} \partial_t f_i(t,\mathbf{x}) + \mathbf{v}_i.\partial_{\mathbf{x}} f_i(t,\mathbf{x}) = \Gamma_i^D[f](t,\mathbf{x}) \quad i = 1,\ldots,n, \quad t \in [0,T], \\ f_i(0,\mathbf{x}) = \phi_i(\mathbf{x}) \quad i = 1,\ldots,n, \quad \mathbf{x} \in \Omega. \end{cases} \tag{54}$$

It is assumed here that $\sigma = 1$ and the velocity has a constant modulus. Moreover pedestrians can take nine directions, and 9 densities: $f_i(t,\mathbf{x}) = f(t,\mathbf{x},\theta_i)$, $i = 1,\ldots,9$ (see Fig. 8), are taken into account.

For all the simulations, the following parameters are chosen:

- dimensions of the domain: length $L_x = 11$ m, and width $L_y = 11$ m.
- size of exit: $S = 1.5$ m.
- quality of domain: $\alpha = 1$.
- 9 different directions of velocity on the set S_θ,

$$S_\theta = \left\{ \theta_i = \frac{i-4}{8}\pi : \ i = 0,\ldots,8 \right\}.$$

- rate of interactions: $\mu = 1 - \rho$, with $0 < \mu < 1$.

The dimensionless quantities are obtained with reference to the following reference variables: $L = 11\sqrt{2}$ m, $V_m = 2$ m s^{-1}, $T_m = 7.78$ s.

A value of $\rho_{max} = 7$ pedestrians/m^2 is chosen to normalize the local density, which is a high value.

The first application deals with the evacuation of $N = 50$ people of a room, first without obstacle and then with an obstacle. This group of 50 pedestrians at initial time is gathered in a circular area of radius $r = 1.91$ m. The maximum number of pedestrians $N_{max} = 7\pi r^2 = 80$. The normalized local density at the initial time (within a circular area) is: $\rho_0 = \frac{50}{\pi r^2 \rho_{max}} = 0.62$. Three cases are considered: first without any obstacle, second with a fixed square-shaped obstacle located between the center of the disk $\mathcal{D}_0$ and the exit and last with a fixed square-shaped obstacle just in front of the exit. For the two last cases, the dimensions of the obstacle are identical : $L_{obs} = l_{obs} = 1.5$ m.

The second application concerns the comparison of the evacuation time obtained by our technique and that obtained by a deterministic discrete method in the simple case of the evacuation of a room without obstacle. The evacuation time is defined as the time when the last pedestrian has left the room. Finally, an obstacle is introduced in the room and the last application examines the effect of the dimensions and location of an obstacle on the time evacuation.

4.2.1. Evacuation of a room without obstacle and with a fixed obstacle

In the first case of the room without obstacle, the coordinates of the center of the disk $\mathcal{D}_0$ at $t = 0$ are $M_0 = (2.3; 3)$; and the results of our model, obtained at four successive times are illustrated in Fig. 9a, b, c and d.

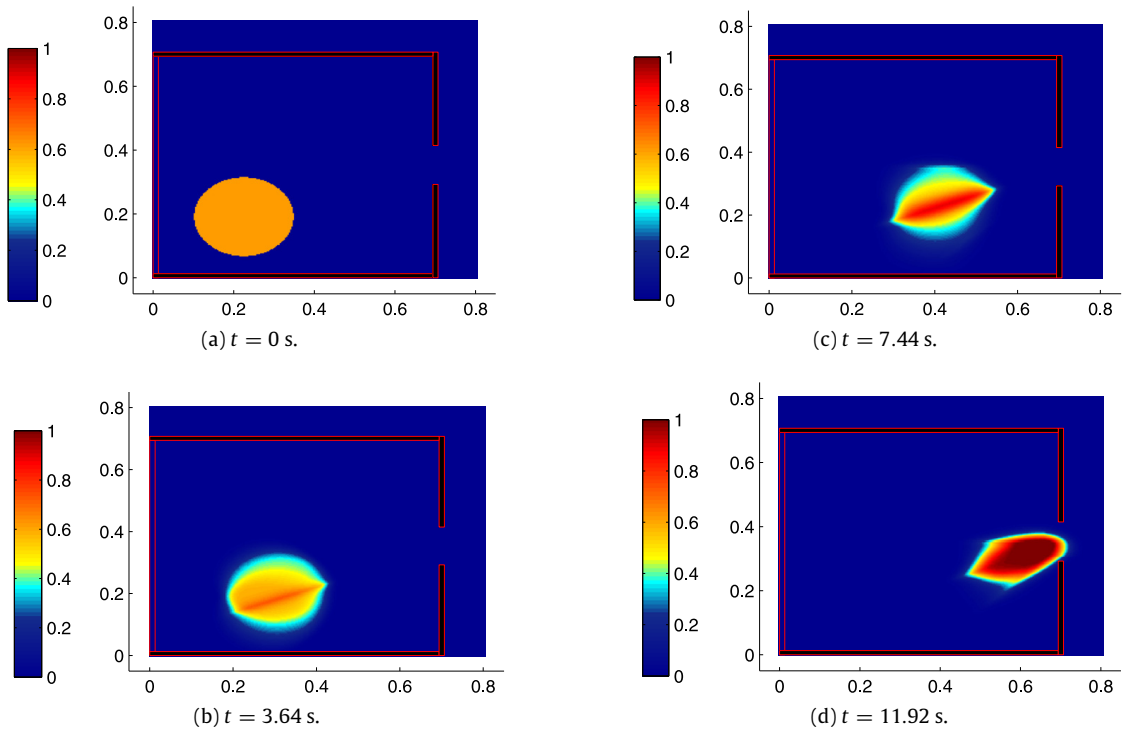


Fig. 9. Time evolution of the local pedestrians density at different times: (a): $t = 0$ s, (b): $t = 3.64$ s, (c): $t = 7.44$ s, and (d): $t = 11.92$ s .

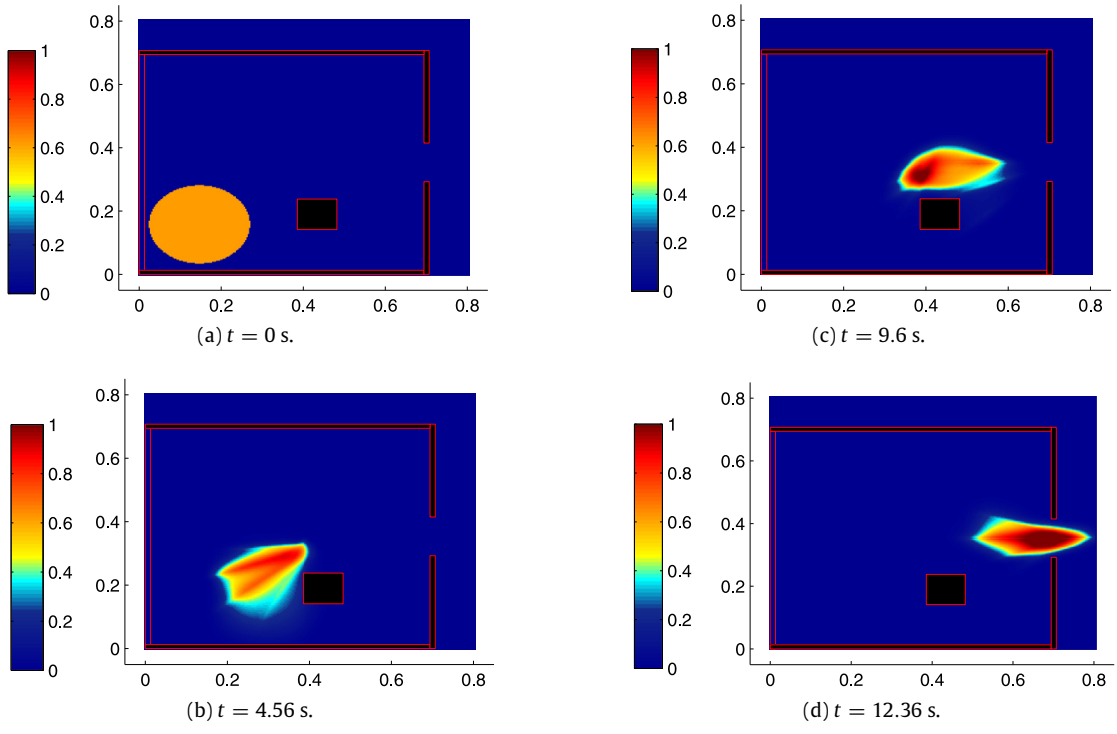


Fig. 10. Time evolution of the local pedestrians density at 4 different times:, (a): $t = 0$ s, (b): $t = 4.56$ s, (c): $t = 9.6$ s, and (d): $t = 12.36$ s .

In the second case, the room contains a fixed square-shaped obstacle (1) whose lower left corner has the coordinates (6.; 2.2), and the center of the initial disk is: $\mathcal{D}_0$ is: $M_0 = (2.3; 2.5)$. The results obtained by our model at four different times are then illustrated in Fig. 10 a, b, c and d.

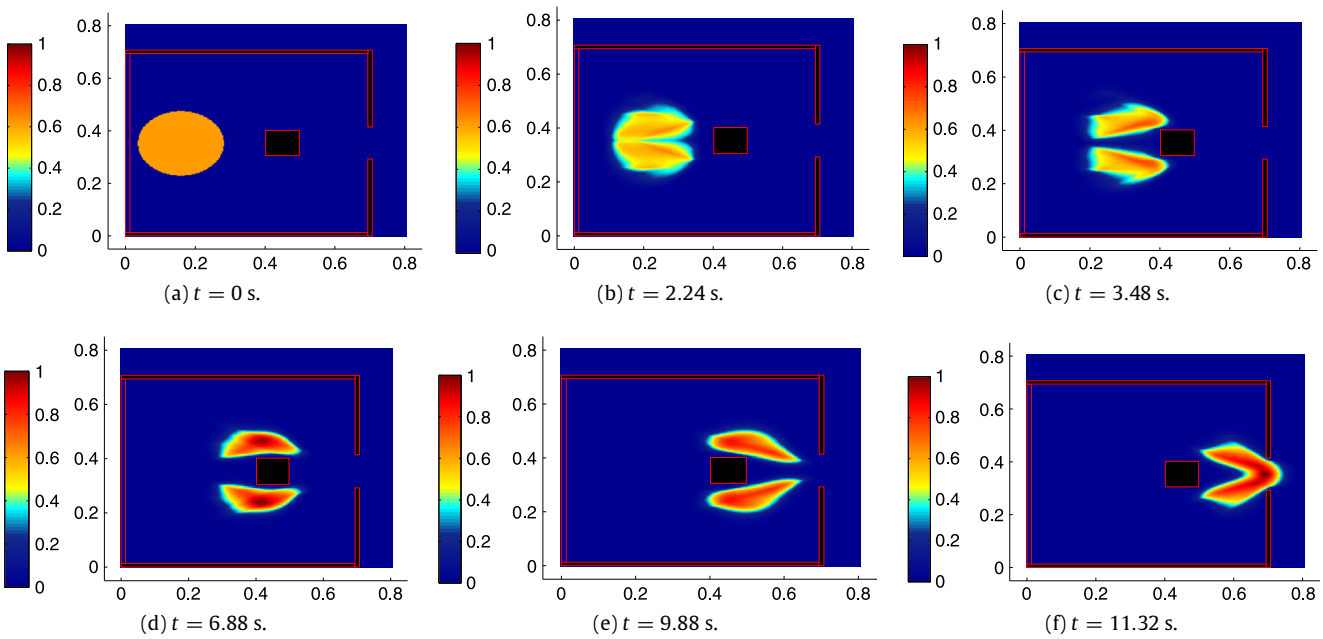


Fig. 11. Time evolution of the local pedestrians density, (a): $t = 0$ s, (b): $t = 2.24$ s, (c): $t = 3.48$ s, (d): $t = 6.88$ s, (e): $t = 9.88$ s and (f): $t = 11.32$ s.

Table 3
Evacuation time (s).

Without obstacle	With obstacle (1)	With obstacle (2)
12.6 s	13.41 s	12.59 s

Table 4
Evacuation time calculated by the kinetic and discrete models for four initial densities.

Initial density (Pedestrian/m^2)	Normalized local density ρ_0	Number of pedestrians N	Kinetic model Evacuation time (s)	Discrete model Mean evacuation time (s)
1.3	0.18	15	11.12	11.10
2.6	0.37	30	11.46	11.51
3.4	0.49	40	11.62	11.67
4.3	0.62	50	11.78	11.80

In the last case, the obstacle (2) whose lower left corner located at the left bottom has the coordinates (6.25; 4.75), is located in front of the exit at a distance of 3.5 m. The center of the initial disk $\mathcal{D}_0$ is: $M_0 = (2.3; 5.5)$. The results obtained from the proposed procedure at six different times are illustrated in Fig. 11 a, b, c, d, e and f.

The group of pedestrians splits into two subgroups when in front of the obstacle to avoid it and with the obstacle bypassed, the group reforms when it leaves the room.

Finally, the evacuation time is collected for the three cases in Table 3. It can be noted that the evacuation time is slightly improved when the obstacle (2) is located in front of the exit.

4.2.2. Comparison of the proposed method with deterministic approach

In order to check the validity of the proposed kinetic model, the evacuation time for a room without obstacle calculated with our kinetic model is compared with a mean evacuation time obtained by a discrete deterministic model: CAPFlow developed by Argoul et al. [3]. In CAPFlow simulations, the pedestrians are randomly distributed in a circular area of radius $r = 1.91$ m. Let the density ρ be in [1.3 4.3] pedestrians/m^2. To examine the sensitivity of the deterministic model to changes in the value of model parameters, the value of some parameters are changed in a random manner (normal law): the radii of the disks representing the pedestrians, the relaxation time after a collision to recover the desired direction etc [1]. 150 simulations with the CAPFlow model are performed. The evacuation time taken by the pedestrians to reach the exit is calculated by both the kinetic and discrete models and the results for four local pedestrians densities are shown in Table 4.

According to the results in the previous table, the evacuation times obtained by the kinetic and discrete models, respectively, for each case are very similar.

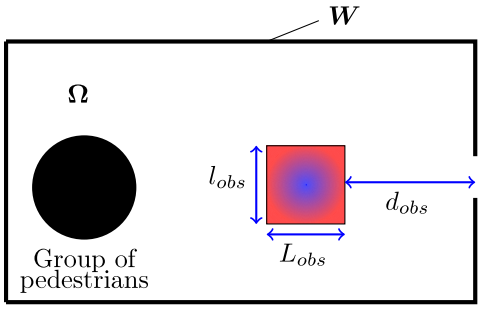


Fig. 12. Illustration of the obstacle characteristics: L_{obs}, l_{obs} are its dimensions, and d_{obs} the distance between the exit and the obstacle.

Table 5
Influence of the distance d_{obs} between the obstacle and the exit and of the dimensions L_{obs} and l_{obs} on the evacuation time in the presence of an obstacle.

d_{obs} (m)	$L_{obs} \times l_{obs}$ (m × m)	Evacuation time (s)
1.5	0.5 × 0.5	12.36
	0.75 × 0.75	12.48
	1 ×1	12.20
	1.25 × 1.25	12.00
	1.5 × 1.5	12.62
1.75	0.5 × 0.5	12.32
	0.75 × 0.75	12.42
	1 ×1	11.86
	1.25 × 1.25	11.42
	1.5 × 1.5	11.48
2	0.5 × 0.5	12.28
	0.75 × 0.75	12.38
	1 ×1	11.78
	1.25 × 1.25	11.38
	1.5 × 1.5	11.26
2.25	0.5 × 0.5	12.24
	0.75 × 0.75	12.32
	1 ×1	12.28
	1.25 × 1.25	11.50
	1.5 × 1.5	11.30
2.5	0.5 × 0.5	12.20
	0.75 × 0.75	12.28
	1 ×1	12.38
	1.25 × 1.25	11.82
	1.5 × 1.5	11.48
2.75	0.5 × 0.5	12.16
	0.75 × 0.75	12.24
	1 ×1	12.32
	1.25 × 1.25	12.42
	1.5 × 1.5	11.90
3	0.5 × 0.5	12.14
	0.75 × 0.75	12.20
	1 ×1	12.28
	1.25 × 1.25	12.38
	1.5 × 1.5	12.50

4.2.3. Influence of the obstacle's position and dimensions on the evacuation time

A fixed obstacle in a room is now located in front of the exit (see Fig. 12). As previously a group of $N = 50$ pedestrians evacuates the room. Our goal is to study the influence of the obstacle on the evacuation time of the group.

The group of $N = 50$ pedestrians is distributed in a circular area of radius $r = 1.91$ m in an empty room (without obstacle). The normed local density at the initial circular area is as previously: $\rho = \frac{50}{\pi r^2 \rho_m} = 0.62$. The evacuation time computed by our technique, to leave the room without any obstacle is: 11.78 s.

Then, the obstacle is located in the room, and the obstacle's dimensions and their distance to the exit are changed. Then, the evacuation time is performed, and its values are collected in Table 5.

According to the results in Table 5, with an accurate choice of obstacle parameters (dimensions, position) the time of the pedestrian evacuation could be improved by putting an obstacle in the room.

The square-shaped obstacle that allows to get a smaller evacuation time is that with dimensions 1.5 m × 1.5 m and located at 1 m from the exit.

Our study confirms that a strategically placed obstacle near an exit can speed evacuations. Several simulations varying its size, shape and orientation have to be performed and tailored to the space and to the width of the exit itself. The proposed technique is interesting for engineers and designers to improve and secure the evacuation of rooms as it executes faster than discrete deterministic model using a random change of some of its parameters.

5. Conclusions

This paper presents the development of a new approach to model crowd dynamics based on the kinetic theory of active particles. The existence and uniqueness of the solution of the proposed mathematical model have been proven for large times thanks to the Banach fixed-point theorem, assuming that the internal interactions between pedestrians are negligible. Numerical simulations using a scheme based on a splitting method have been performed. The proposed model is able to describe some characteristics of the dynamics of a crowd in a bounded domain (with walls, an exit and an obstacle) namely, avoidance and evacuation. Indeed, several cases were studied to show the ability of this model to reproduce pedestrian behavior, and the results were compared to those obtained with a deterministic approach. Finally, the influence of a square-shaped obstacle on the evacuation time is studied. Indeed, with a suitable choice of obstacle parameters (dimensions and position), it has been shown that the time of the pedestrian evacuation may become much shorter than without any obstacle.

Acknowledgments

The authors are grateful to the reviewers for their insightful remarks and corrections. Their feedback has greatly influenced the quality of this paper.

The authors would also like to thank Nicholas Collins-Craft for his help editing the English of this paper.

Appendix A

Proof of Lemma.

1. It is assumed that:

$$|\rho_1(t,x)-\rho_2(t,x)| \leq exp(-\lambda t)\ldots \times \left(\sum_{h=1}^{n}\sum_{k=1}^{n}|\psi_h^1\,(t,x-v\ cos(\theta_h)t,y-v\ sin(\theta_h)t)-\psi_k^2\,(t,x-v\ cos(\theta_k)t,y-v\ sin(\theta_k)t)|\right),$$

from Eq. (27) and the hypothesis (A.1), (A.2), let us note $B_{hk}(i)[\rho(t,\mathbf{x})] \leq 1$, the following estimates are then deduced:

$$\|\widehat{\Psi}[\boldsymbol{\psi}^1,\boldsymbol{\psi}^1](t)-\widehat{\Psi}[\psi^2,\boldsymbol{\psi}^2](t)\|_1 \leq \left(c_\eta+nR\left(c_\eta L_B+L_\eta\right)\right)\left(\|\boldsymbol{\psi}^1(t)\|_1+\|\boldsymbol{\psi}^2(t)\|_1\right)\|\boldsymbol{\psi}^1(t)-\boldsymbol{\psi}^2(t)\|_1, \tag{55}$$

$$\|\widehat{\Upsilon}[\boldsymbol{\psi}^1]\widehat{\boldsymbol{\psi}^1}(t)-\widehat{\Upsilon}[\boldsymbol{\psi}^2]\widehat{\boldsymbol{\psi}^2}(t)\|_1 \leq \left(c_\eta+L_\eta R\right)\left(\|\boldsymbol{\psi}^1(t)\|_1+\|\boldsymbol{\psi}^2(t)\|_1\right)\|\boldsymbol{\psi}^1(t)-\boldsymbol{\psi}^2(t)\|_1, \tag{56}$$

$$\|\widehat{\Lambda}[\boldsymbol{\psi}^1](t)-\widehat{\Lambda}[\psi^2](t)\|_1 \leq \left(c_\mu+nRL_\mu\right)\|\boldsymbol{\psi}^1(t)-\boldsymbol{\psi}^2(t)\|_1. \tag{57}$$

According to estimates (55)–(57), the estimate (28) is obtained with

$$C_1 = 2c_\eta + R\left(nc_\eta L_B+(n+1)L_\eta\right),$$

$$C_2 = 2c_\mu + (n+1)RL_\mu.$$

2. Since $\widehat{\Psi}_i[\boldsymbol{\psi},\boldsymbol{\psi}](t,\mathbf{x}) \geq 0$ and $\widehat{\Lambda}_i[\boldsymbol{\psi}](t,\mathbf{x}) \geq 0$ because $\psi_i(\mathbf{x}) \geq 0$, then $(\widehat{A\psi})_i(t,\mathbf{x}) \geq 0$ if $\lambda-\mu[\rho]-\widehat{\Upsilon}_i[\boldsymbol{\psi}](t,\mathbf{x})exp(-\lambda t) \geq 0$, from (A.1) and Eq. (27), it is deduced that:

$$\Upsilon_i[\boldsymbol{\psi}](t,\mathbf{x}) \leq Rexp(\lambda t)c_\eta|V|,\quad i=1\ldots n,$$

so for that $\widehat{A}(\boldsymbol{\psi})_i(t,\mathbf{x}) \geq 0$, it is enough to choose $\lambda \geq \lambda_0 = Rc_\eta|V|+c_\mu$.

3. Let us note that if $P_h(i) \leq 1$, from (A.1) and Eq. (27), it can deduced that, *for* $i=1,\ldots,n$,

$$\widehat{\Psi}_i[\boldsymbol{\psi},\boldsymbol{\psi}](t,\mathbf{x}) \leq c_\eta|V|R^2exp(2\lambda t),$$
$$\widehat{\Lambda}_i[\boldsymbol{\psi}](t,\mathbf{x}) \leq c_\mu Rexp(\lambda t),$$

hence for a choice of R such that:

$$R \geq \sum_{i=1}^{n}\|\phi_i\|_\infty = R_1,$$

and a time: t such that

$$t \leq T = \frac{1}{\lambda} ln \left(1 + \frac{\lambda}{nR\left(c_\eta R|V| + c_\mu\right)} \left(R - \sum_{i=1}^{n} \|\phi_i\|_\infty \right) \right) := T_0, \tag{58}$$

we obtain the relation (31).

4. Let $\boldsymbol{\psi} \in \mathbb{X}_T$, $T \geq 0$, according to Eq. (9), (6) for $P_h(i)$, $B_{hk}(i)$, the hypothesis (A.1), and using the following variable changes:

$$(z_1, z_2) \leftarrow (x + v\,(cos(\theta_i) - cos(\theta_h))\,t, y + v\,(sin(\theta_i) - sin(\theta_h))\,t)\,,$$
$$(w_1, w_2) \leftarrow (x - v\,cos(\theta_i)t, y - v\,sin(\theta_i)t)\,.$$

We show that:

$$\|\widehat{\boldsymbol{\Psi}}[\boldsymbol{\psi}, \boldsymbol{\psi}](t)\|_1 \leq c_\eta \|\boldsymbol{\psi}(t)\|_1^2, \tag{59}$$
$$\|\widehat{\boldsymbol{\Upsilon}}[\boldsymbol{\psi}]\widehat{\boldsymbol{\psi}}(t)\|_1 \leq c_\eta \|\boldsymbol{\psi}(t)\|_1^2, \tag{60}$$
$$\|\widehat{\Lambda}[\boldsymbol{\psi}](t)\|_1 \leq c_\mu \|\boldsymbol{\psi}(t)\|_1. \tag{61}$$

From the relations (59)–(61) it can be found:

$$\|\mathbf{A}\boldsymbol{\psi}(t)\|_1 \leq \|\boldsymbol{\phi}\|_1 + \frac{2}{\lambda} c_\eta\,(1 - exp(-\lambda t))\,\|\boldsymbol{\psi}\|_{\mathbb{X}_T}^2 + \left(\lambda + 2c_\mu\right) t \|\boldsymbol{\psi}\|_{\mathbb{X}_T}, \tag{62}$$

since $2c_\mu \leq C_1$ and $2c_\eta \leq C_2$, hence the estimate (34) becomes:

$$\|\mathbf{A}\boldsymbol{\psi}\|_{\mathbb{X}_T} \leq \|\boldsymbol{\phi}\|_1 + \frac{C_1}{\lambda} \|\boldsymbol{\psi}\|_{\mathbb{X}_T}^2 + (\lambda + C_2)\,T \|\boldsymbol{\psi}\|_{\mathbb{X}_T}. \tag{63}$$

This ends the proof of the lemma.

Appendix B

The following equation

$$(E) : \|\boldsymbol{\phi}\|_1 + \frac{1}{\lambda} C_1 a_0^2 \|\boldsymbol{\phi}\|_1^2 + (C_2 + \lambda)\,T a_0 \|\boldsymbol{\phi}\|_1 = a_0 \|\boldsymbol{\phi}\|_1$$

has a solution a_0 if,

$$\|\boldsymbol{\phi}\|_1 \leq min \left\{ \frac{(n-1)^2 c_\mu}{8n^2 c_\eta}, \frac{(n-1)^2 c_\eta |V|}{\left(4n^2\left(nc_\eta L_B + (n+1)L_\eta\right)\right)} \right\}.$$

Indeed, the discriminant of the equation (E) is given by,

$$\Delta_0 = ((C_2 + \lambda)\,T - 1)^2 - 4\frac{C_1}{\lambda} \|\boldsymbol{\phi}\|_1.$$

$\Delta_0 \geq 0$ if $((C_2 + \lambda)\,T - 1)^2 \geq 4\frac{C_1}{\lambda} \|\boldsymbol{\phi}\|_1$.

According to Eq. (35) and for $R \geq R_1$, $\lambda = \lambda_0 = c_\eta R|V| + c_\mu$,

$$T = \frac{1}{\lambda + C_2} ln \left(1 + \frac{\lambda}{nR\left(c_\eta R|V| + c_\mu\right)} \left(R - \sum_{i=1}^{n} \|\phi_i\|_\infty \right) \right),$$

by using the following inequality $ln(1 + x) \leq x$, for $x \geq 0$

$$(\lambda + C_2)\,T \leq \frac{1}{n},$$

which implies,

$$(\lambda + C_2)\,T - 1 \leq \frac{1 - n}{n},$$

then,

$$[(\lambda + C_2)\,T - 1]^2 \geq \frac{(1-n)^2}{n^2}.$$

Hence, if we have

$$4\frac{C_1}{\lambda} \|\boldsymbol{\phi}\|_1 \leq \frac{(1-n)^2}{n^2}$$

then we get: $4\frac{C_1}{\lambda}\|\boldsymbol{\phi}\|_1 \leq ((C_2+\lambda)\,T-1)^2$, by taking

$$\|\boldsymbol{\phi}\|_1 \leq \frac{(1-n)^2}{n^2}\frac{\lambda}{4C_1},$$

hence,

$$\|\boldsymbol{\phi}\|_1 \leq \frac{(1-n)^2}{4n^2}\frac{Rc_\eta|V|+c_\mu}{\left(2c_\eta+R\left(nc_\eta L_B+(n+1)L_\eta\right)\right)} := \varphi^*(R).$$

Moreover,

$$\|\boldsymbol{\phi}\|_1 \leq \underset{R\geq 0}{argmin}\left\{\varphi^*(R)\right\}.$$

By studying the following function:

$$\varphi^* \;: R \to \varphi^*(R), \quad with\ R \in [0,+\infty[,$$

we have:

$$\underset{R\geq 0}{argmin}\left\{\varphi^*(R)\right\} = min\left\{\frac{(n-1)^2c_\mu}{8n^2c_\eta}, \frac{(n-1)^2c_\eta|V|}{\left(4n^2\left(nc_\eta L_B+(n+1)L_\eta\right)\right)}\right\}$$

in particular for $R > R_1$, we have:

$$\underset{R>R_1}{argmin}\left\{\varphi^*(R)\right\} = min\left\{\frac{(n-1)^2c_\mu}{8n^2c_\eta}, \frac{(n-1)^2c_\eta|V|}{\left(4n^2\left(nc_\eta L_B+(n+1)L_\eta\right)\right)}\right\} := \varphi^0.$$

References

[1] M. Takashi, T. Akiyasu, I. Mayuko,, S. Kohta, U. Daishin, in: M. Chraibi, M. Boltes, A. Schadschneider, A. Seyfried (Eds.), Effects of an Obstacle Position for Pedestrian Evacuation: SF Model Approach, Traffic and Granular Flow '13, Springer International Publishing, Cham, 2015, pp. 163–170.
[2] E. Cristiani, B. Piccoli, A. Tosin, Multiscale Modeling of Pedestrian Dynamics, Springer International Publishing, Berlin, 2014.
[3] P. Argoul, B. Kabalan, Pedestrian trajectories and collisions in crowd motion, in: Collisions Engineering: Theory and Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, 2017, pp. 79–144.
[4] D. Helbing, P. Molnár, Social force model for pedestrian dynamics, Phys. Rev. E 51 (1995) 4282–4286.
[5] D. Helbing, P. Molnár, I.J. Farkas, K. Bolay, Self-organizing pedestrian movement, Environ. Plan. B: Plan. Des. 28 (2001) 361–383.
[6] R.L. Hughes, A continuum theory for the flow of pedestrians, Transp. Res. 36 (6) (2002) 507–535.
[7] J.P. Agnelli, F. Colasuonno, D. Knopoff, A kinetic theory approach to the dynamics of crowd evacuation from bounded domains, Math. Models Methods Appl. Sci. 25 (01) (2015) 109–129.
[8] N. Bellomo, A. Bellouquid, D. Knopoff, From the microscale to collective crowd dynamics, Multiscale Model. Simul. 11 (3) (2013) 943–963.
[9] N. Bellomo, A. Bellouquid, On the modeling of crowd dynamics: Looking at the beautiful shapes of swarms, Netw. Heterog. Media 6 (3) (2011) 383–399.
[10] N. Bellomo, L. Gibelli, Toward a mathematical theory of behavioral-social dynamics for pedestrian crowds, Math. Models Methods Appl. Sci. 25 (13) (2015) 2417–2437.
[11] N. Bellomo, A. Bellouquid, On multiscale models of pedestrian crowds from mesoscopic to macroscopic, Commun. Math. Sci. 13 (2015) 1649–1664.
[12] N. Bellomo, Modeling complex living systems, in: Modeling Complex Living Systems, Birkhäuser Basel, Birkhäuser Boston, 2008.
[13] V.V. Aristov, Direct Methods for Solving the Boltzmann Equation and Study of Nonequilibrium Flows, Springer Netherlands, Boston, 2001.
[14] P. Barbante, A. Frezzotti, L. Gibelli, A kinetic theory description of liquid menisci at the microscale, Kinet. Relat. Models 8 (2) (2015) 235–254.
[15] N. Bellomo, A. Bellouquid, Global solution to the cauchy problem for discrete velocity models of vehicular traffic, J. Differential Equations 252 (2) (2012) 1350–1368.
[16] E.A., A. Bellouquid, E.D. Angelis, L. Fermo, Towards the modeling of vehicular traffic as a complex system: A kinetic theory approach, Math. Models Methods Appl. Sci. 22 (supp01) (2012) 1140003.
[17] N. Bellomo, L. Gibelli, Behavioral crowds: Modeling and Monte Carlo simulations toward validation, Comput. & Fluids 141 (2016) 13–21. Advances in Fluid-Structure Interaction.
[18] N.B., G. Ajmone Marsane, N. Bellomo, L. Gibelli, Stochastic evolutionary differential games toward a systems theory of behavioral social dynamics, Math. Models Methods Appl. Sci. 26 (06) (2016) 1051–1093.
[19] G. Still, Crowd dynamics, (Ph.D. thesis), University of Warwick, 2000.
[20] J. Geiser, G. Tanoğlu, N. Gücüyenen, Higher order operator splitting methods via zassenhaus product formula: Theory and applications, Comput. Math. Appl. 62 (4) (2011) 1994–2015.
[21] M.Y. Kim, E.-J. Park, An upwind scheme for a nonlinear model in age-structured population dynamics, Comput. Math. Appl. 30 (8) (1995) 5–17.
[22] J.C. Lopez-Marcos, An upwind scheme for a nonlinear hyperbolic integro-differential equation with integral boundary condition, Comput. Math. Appl. 22 (11) (1991) 15–28.