



HAL
open science

Extraction de palettes de couleurs pour l'aide à la conception cartographique

Sidonie Christophe, Julien Perret, Charlotte Hoarau

► **To cite this version:**

Sidonie Christophe, Julien Perret, Charlotte Hoarau. Extraction de palettes de couleurs pour l'aide à la conception cartographique. *Revue des Sciences et Technologies de l'Information - Série TSI: Technique et Science Informatiques*, 2013, 32 (3-4), pp.401-430. 10.3166/tsi.32.401-430 . hal-02280931

HAL Id: hal-02280931

<https://hal.science/hal-02280931v1>

Submitted on 16 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Titre → TNR 18
Gras, bordure ¼ à
20pts au-dessus
18 après s'il y a un
sous-titre,
24 après s'il n'y a
pas de sous-titre

Extraction de palettes de couleurs pour l'aide à la conception cartographique

Sidonie Christophe¹, Julien Perret¹, Charlotte Hoarau¹

*1. Laboratoire COGIT – Institut National de l'Information Géographique et Forestière (IGN), Université Paris-Est
73 avenue de Paris, F-94165 Saint Mandé, France
{sidonie.christophe; julien.perret; charlotte.hoarau}@ign.fr*

RESUME. La démocratisation des outils de conception cartographique nécessite que ceux-ci intègrent et apportent une expertise cartographique aux utilisateurs. En particulier, l'étape du choix des couleurs, enjeu important pour la lisibilité de la carte, est un problème complexe requérant des connaissances variées sur la couleur. Nous disposons d'un outil COLorLEGend permettant à un utilisateur de concevoir des cartes possibles à partir de peintures célèbres : en entrée, des palettes de couleurs extraites manuellement des peintures sont proposées à l'utilisateur afin qu'il exprime ses préférences de couleurs. Le travail décrit ici consiste à évaluer des méthodes d'extraction automatique de palettes de couleurs depuis ces peintures pouvant être intégrées à la chaîne de traitement de COLorLEGend. La question principale est de déterminer automatiquement les couleurs représentatives de la peinture qui pourraient être utilisées ensuite dans des cartes. Plusieurs contributions sont apportées, d'une part par l'enrichissement de méthodes de classification de couleurs, et d'autre part par la proposition d'une distance entre palettes de couleurs permettant d'effectuer une évaluation quantitative des résultats des algorithmes proposés.

ABSTRACT. The democratization of cartographic design tools suggests that efficient and accessible tools for cartographic design support are needed. Specifically, the color choice step is an important issue for the legibility of the map being designed. Research studies have been conducted to formalize and exploit various knowledge on color. COLorLEGend is a research tool allowing the user to design maps using famous paintings: color palettes manually extracted from the paintings are proposed to the user as inputs in order for her/him to express her/his preferences. The work described in this article consists in evaluating automatic palette extraction methods that could be integrated into the COLorLEGend processing chain. The main question addressed is the automatic determination of the representative colors of a painting that could be used for cartographic design. Several contributions are brought forward, first by the enrichment of color clustering methods, second by the proposal of a distance between color palettes that allows for the quantitative evaluation of the results of the proposed algorithms.

MOTS-CLES : aide à la conception cartographique, style cartographique, choix des couleurs, classification/quantification des couleurs, extraction de palettes, peintures.

KEYWORDS: cartographic design aid, cartographic style, color choice, color classification/quantization, palette extraction, masters' paintings.

1. Introduction

L'offre des outils cartographiques en ligne (CloudMade¹, Google Maps², Géoportail³, etc.) permet à tout utilisateur de se construire facilement et rapidement une carte : chacun peut s'improviser « cartographe ». Ces outils proposent des fonctionnalités plus ou moins avancées de symbolisation : modification des symboles et des couleurs par des palettes prédéfinies et réduites dans Google Maps, création et modification exhaustive des styles dans CloudMade. Quel que soit le degré de personnalisation du rendu cartographique offert, ces outils n'intègrent cependant ni expertise cartographique ni interactions avec l'utilisateur permettant de le guider dans ses choix sémiologiques ou d'évaluer sa carte résultante. En conséquence, les cartes construites en ligne ne sont pas toujours adaptées aux besoins, goûts et intentions de leurs auteurs, les rendant plus ou moins faciles à lire ou compréhensibles. L'appropriation de la notion d'efficacité de la carte par les utilisateurs est une question de recherche majeure en cartographie : tout utilisateur, qu'il soit expert ou novice en cartographie, doit pouvoir construire une carte personnalisée transmettant un message cartographique construit en fonction de ses besoins et de ses goûts.

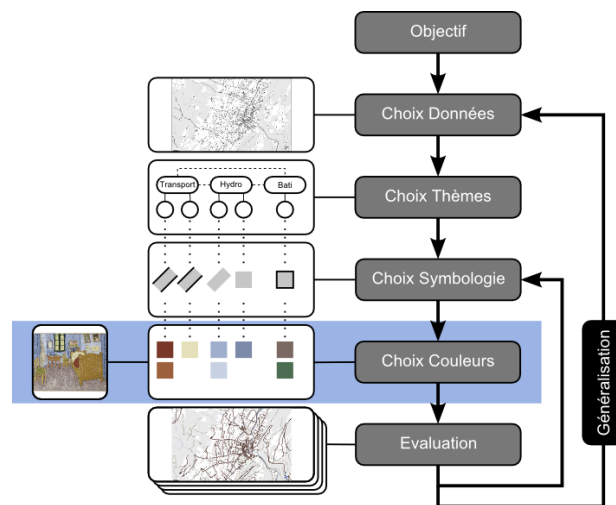


Figure 1. Processus simplifié de conception cartographique du point de vue utilisateur et positionnement de notre problème de choix des couleurs à résoudre (rectangle bleu)

1. <http://maps.cloudmade.com/editor>

2. <http://maps.google.com/maps>

3. <http://geoportail.fr>

La Figure 1 présente un schéma simplifié du processus de conception cartographique du point de vue utilisateur : nous nous intéressons en particulier à l'aide au choix des couleurs dans ce processus. Ceci nécessite que les connaissances sur le choix des couleurs en cartographie soient formalisées.

En cartographie, la théorie nous propose des recommandations sur l'usage de la couleur (MacEachren, 1995 ; Robinson, 1995 ; Denègre, 2005, etc.) liées aux connaissances sur les aspects perceptifs et cognitifs de la couleur en général (Sutter, 1858 ; Pastoureau, 1992 ; Mollard-Desfour, 1998 ; Roque, 2009, etc.) et en sémiologie graphique (Bertin, 1967). La couleur est aujourd'hui la variable visuelle la plus utilisée en cartographie numérique, mais également la plus difficile à manipuler. Des outils d'aide à la conception cartographique intègrent des connaissances sur les couleurs et les contrastes de couleurs à respecter en cartographie (Brewer, 1992 ; Chesneau, 2006 ; Buard *et al.*, 2009). Un des auteurs de cet article s'est intéressé aux usages de la couleur en cartographie afin d'aider les utilisateurs à faire des choix de couleurs lors de la conception d'une carte personnalisée. Dans l'application de conception cartographique COLorLEGend dédiée, deux types de sources d'inspiration sont ainsi proposées à l'utilisateur afin qu'il puisse raisonner par analogie : des échantillons de cartes, proposant des usages conventionnels, traditionnels et plutôt classiques des couleurs en cartographie, et des peintures, proposant des usages artistiques plutôt originaux de couleurs elles-mêmes originales pour la cartographie. L'idée de s'inspirer de sources variées permet de s'adapter aux besoins, goûts et démarches créatives des utilisateurs et ainsi d'améliorer le degré de personnalisation des cartes produites (Christophe, 2011). D'autres travaux sont en cours pour utiliser d'autres sources d'inspiration comme des images ou des styles prédéfinis -- style « pop-art » (Kasbarian *et al.*, 2012), styles européens (Christophe, 2012) -- ou pour proposer des représentations cartographiques plus réalistes (Hoarau, 2012).

Dans (Christophe, 2011), les couleurs issues de peintures ont été extraites manuellement, en considérant les couleurs saillantes d'une peinture. Dans cet article, nous tentons de résoudre la question d'une extraction automatique de palettes de couleurs dans ces peintures, pouvant être ensuite proposées à un utilisateur pour l'aider à symboliser ses données géographiques. Dans une première partie, nous décrivons brièvement les connaissances sur la couleur en cartographie ainsi que les usages cartographiques des couleurs. Dans une deuxième partie, nous présentons notre approche qui consiste à utiliser des couleurs originales pour la cartographie, provenant d'images spécifiques que sont des peintures. Dans une troisième partie, nous confrontons différentes méthodes d'extraction automatique de palettes. Finalement, nous présentons nos résultats et nous discutons des différentes méthodes proposées.

2. Du bon usage de la couleur en cartographie

La représentation cartographique consiste, entre autres, à faire des choix de signes graphiques pour représenter des informations géographiques. La difficulté de ce problème réside dans la maîtrise de la relation entre un signe graphique et un sens, ainsi que dans la façon de créer de telles relations pour construire efficacement des messages visuels liés à l'espace géographique. A ce titre, la variable visuelle couleur est particulièrement difficile à manipuler et est souvent mal utilisée : inadaptation entre couleur et thème cartographié, mauvaise association de couleurs, choix d'une couleur trop symbolique utilisée à contre-sens, etc. (Bord, 2008). Cette utilisation inappropriée de la couleur perturbe la lisibilité ainsi que l'efficacité de la carte.

La teinte et la valeur⁴ sont deux des sept variables visuelles définies par Jacques Bertin, dont il a décrit les propriétés perceptives dans son ouvrage 'Sémiologie Graphique' (Bertin, 1967) servant de référence fondamentale pour la théorie cartographique (MacEachren, 1995 ; Robinson, 1995, etc.) : la teinte permet de différencier des entités géographiques et d'associer les entités géographiques d'un même thème, la valeur permet de mettre en évidence un ordre entre les entités représentées. La couleur dans ses trois dimensions (teinte, valeur, saturation) joue donc un rôle important pour représenter les relations sémantiques entre les thèmes géographiques représentés et à l'intérieur de ceux-ci. En manipulant ces trois dimensions de la couleur et les contrastes de couleurs (les sept contrastes⁵ d'Itten (1967), le contraste simultané de Chevreul (1839)), il est possible de mettre en évidence des informations de natures différentes : densités, oppositions entre thèmes, thème dominant, etc. (Mersey, 1990 ; Brewer, 1992 ; Klippel *et al.*, 2009). Les aspects perceptifs et cognitifs de la couleur (Sutter, 1858 ; Pastoureau, 1992 ; Mollard-Desfour, 1998 ; Roque, 2009 ; etc.) sont également largement repris en cartographie. La couleur a un fort pouvoir évocateur (le bleu fait penser à l'eau, le vert fait penser à la végétation, etc.) qui est utilisé par les conventions d'utilisation des couleurs en cartographie (Robinson, 1995) ; elle a aussi un fort pouvoir symbolique (le rouge fait penser à quelque chose de dangereux) souvent utilisé en cartographie pour mettre en évidence une information ou un phénomène prédominant (Klippel *et al.*, 2009). Par ailleurs, l'harmonie dans la carte, recommandée dans les ouvrages de cartographie, commence seulement à être formalisée à partir de connaissances issues des arts visuels (Kandinsky, 1989 ; Sawahata, 2001 ; Sutton *et al.*, 2004 ; Gage, 2006 ; etc.) et considérée comme un facteur majeur d'amélioration de la lisibilité de la carte (Brewer, 1994 ; 2003 ; Christophe *et al.*, 2011). Pour finir, la perception et l'interprétation des couleurs dépendent des capacités perceptives (Dhee, 2011) et cognitives de celui qui les regarde, ainsi que de sa culture (le vert est par exemple utilisé pour représenter les édifices religieux dans de nombreux pays musulmans).

4. Valeur : progression continue du blanc au noir en passant par la saturation complète d'une teinte

5. Couleur en soi, clair - obscur, complémentaires, chaud - froid, quantité, qualité, simultané (Itten, 1967)

Il est important de souligner que la théorie cartographique ne précise pas comment toutes ces recommandations doivent être appliquées en pratique afin de construire une carte : il n'y a pas de méthode générique pour choisir, associer et appliquer correctement des variables visuelles sur n'importe quel jeu de données géographiques, pour transmettre un message particulier. Ces recommandations n'ont effectivement été que peu ou pas du tout expérimentées et commencent seulement à l'être (Fabrikant *et al.*, 2010). Si dans certains domaines spécifiques, des recommandations pour la représentation graphique apparaissent (CNIG, 2011), c'est en général pour viser un consensus et une représentation standardisée (ce qui peut potentiellement conduire à une meilleure lisibilité et efficacité de la carte) et non pour viser à améliorer la créativité du processus de conception cartographique (au contraire).

Des travaux du COGIT (Buard *et al.*, 2009 ; Christophe, 2011) ont considéré les règles suivantes comme étant des règles majeures pour l'utilisation des couleurs en cartographie et ce sont ces règles que nous utilisons dans notre travail :

- représenter les relations sémantiques (association, différence, ordre) entre les entités géographiques par des relations visuelles spécifiques entre les couleurs ;
- représenter les entités géographiques (ex. : mer, forêt, etc.) de façon conventionnelle ;
- utiliser un contraste fort entre le fond cartographique et les entités géographiques en superposition à mettre en évidence.

Il faut également souligner que même si les aspects esthétiques et artistiques de la cartographie sont systématiquement évoqués dans la théorie, il n'existe pas de recommandations pratiques pour mieux utiliser la couleur et les autres signes graphiques de façon esthétique et artistique pendant le processus de conception cartographique.

3. Explorer des couleurs originales en cartographie

En tant que chercheurs en cartographie, nous souhaitons améliorer les connaissances sur le processus de conception cartographique, et nos outils automatiques d'aide à la conception cartographique. Le choix des couleurs pendant l'étape de symbolisation est un processus complexe et nous devons fournir aux utilisateurs les moyens de personnaliser ces choix afin de s'adapter à leurs besoins et goûts. Nous ne sommes pas dans le contexte de la production de cartes papier en série, dont les choix de légende et de couleurs ont été spécifiquement étudiés et affinés pendant des décennies pour répondre à un besoin et un contexte d'utilisation précis (ex. : la carte de randonnée au 1/25000).

3.1. Limites imposées par les outils et les habitudes

Pendant la thèse de (Christophe, 2009), deux constats se sont imposés dans le domaine de la conception de cartes personnalisées et soulèvent les deux questions suivantes :

– les mêmes couleurs sont souvent utilisées : d’une part, les outils proposent souvent des palettes avec peu de couleurs et des choix peu variés pour une teinte, d’autre part, les utilisateurs ont l’habitude des codes couleurs des cartes papier nationales ou des outils comme Google Maps ou OpenStreetMap⁶ et peinent à s’en détacher. Les utilisateurs sont donc restreints par les outils et par leurs habitudes, freinant toute démarche créative inhérente au processus cartographique, et risquant ainsi de ne pas réussir à construire un message cartographique adapté : est-il possible de trouver des couleurs différentes de « ce qui est connu » et d’améliorer ainsi la qualité des cartes personnalisées ?

– des usages traditionnels des couleurs sont mis en pratique en cartographie (conventions, contraste du fond) : serait-il possible d’explorer d’autres usages plus originaux des couleurs qui seraient adaptés à la conception cartographique et qui amélioreraient la qualité des cartes personnalisées ?

Nous présentons ici des travaux existants sur l’utilisation de couleurs originales en cartographie, comment nous en sommes venus à utiliser des peintures comme sources d’inspiration, ainsi que des travaux sur l’extraction de palettes de couleurs depuis des peintures.

3.2. Travaux existants sur l’utilisation de couleurs originales en cartographie

L’originalité a été peu traitée jusqu’à maintenant en tant que telle par les cartographes ou spécialistes de l’information géographique. Elle est effectivement traitée indirectement par exemple par l’analyse de la diversité des styles cartographiques européens (Kent *et al.*, 2009), soulignant une certaine variabilité et donc originalité entre les choix de représentations cartographiques dans les différents pays : la notion d’originalité est donc aussi dépendante des usages et pratiques cartographiques développés par les agences de cartographie nationales. Elle est depuis peu directement abordée par exemple dans l’évaluation de différents types de cartes pour évaluer leur niveau de « décorativité », i.e. l’ajout d’éléments graphiques 2D ou 3D, « expressivité », i.e. l’accentuation visuelle d’un élément (grossissement, assombrissement, surbrillance, etc.) ou « originalité », i.e. la proposition de solutions de visualisation inhabituelles observables dans la carte entière (Beconyte, 2011). Beconyte (2011) définit alors l’originalité comme le degré de déviation des choix de visualisation standards pour un type particulier de carte. Des propositions originales de rendu sont de plus en plus proposées par les outils en ligne, par exemple le rendu « aquarelle » sur des données OpenStreetMap⁷.

Christophe (2009) propose d’aider les utilisateurs à faire des choix de couleurs originaux : il s’agit à la fois d’utiliser des couleurs et des associations de couleurs inhabituelles en cartographie, en utilisant des exemples de cartes existantes (il faut noter que la légende topographique norvégienne peut paraître très originale pour un français) et des peintures célèbres, utilisant des associations originales de couleurs

⁶ <http://openstreetmap.fr/>

⁷ <http://maps.stamen.com/#watercolor/12/37.7706/-122.3782>

originales pour la cartographie. Le choix a été fait de se tourner vers les peintres et les théoriciens de l'art afin de bénéficier d'une part de connaissances différentes de celles issues du domaine cartographique, d'autre part de pratiques purement artistiques de la couleur et, au final, de choix artistiques forts en termes d'impacts perceptif et cognitif. L'application COLorLEGend développée par Christophe (2009) propose d'aider un utilisateur à faire des choix de couleurs pendant le processus de conception cartographique : l'utilisateur fournit son jeu de données et le système l'aide à sélectionner des couleurs pour symboliser ses données. COLorLEGend propose à l'utilisateur d'exprimer ses préférences de couleurs, à partir d'un type de source d'inspiration (des échantillons de carte ou des peintures). Le système confronte ensuite ces préférences de couleurs avec des contraintes cartographiques et artistiques stockées : un algorithme de résolution d'un problème à satisfaction de contraintes applique les couleurs préférées aux thèmes géographiques de la légende, construisant ainsi un ensemble de cartes variées possibles, les contraintes courantes pouvant être ou non respectées. L'utilisateur peut naviguer dans ces cartes résultantes, choisir la carte la plus satisfaisante pour lui et faire quelques retouches pour finaliser sa carte personnalisée. La Figure 2 présente deux cartes parmi d'autres pouvant être obtenues avec les couleurs de la palette du tableau de van Gogh⁸ présenté, selon que des règles artistiques « à la van Gogh » (carte a) ou des règles cartographiques (conventions, contrastes) (carte b) sont respectées.

Dans ce travail, les extractions de palettes de couleurs ont été réalisées manuellement (cf. §4.1). Le problème qui nous intéresse ici est le suivant : comment ces peintures peuvent-elles être exploitées de façon automatique afin d'être utilisées en entrée d'un processus d'aide à la conception de cartes ?

8. Vincent van Gogh, La Chambre à Arles, 1888, Musée van Gogh, Amsterdam, Pays-Bas.



Figure 2. Applications artistique « à la van Gogh » (a) et traditionnelle (b) des couleurs d'un tableau de van Gogh

3.3. Travaux existants en extraction de palettes

Dans ce paragraphe, nous nous intéressons aux travaux existants en matière d'extraction de palettes : des outils en ligne et des travaux de recherche sont présentés. Nous devons préciser que notre objectif n'est pas d'obtenir la palette du « vrai » tableau, ce qui est impossible : nous nous situons en effet dans un contexte applicatif où c'est l'utilisateur qui va apporter une image, quelle qu'elle soit, qu'il souhaitera utiliser comme source d'inspiration pour construire sa carte personnalisée. Les reproductions des peintures utilisées peuvent donc être légèrement différentes des vraies peintures⁹.

Outils en ligne

Nous avons testé deux outils – Color Palette Generator¹⁰ et Pic2Color¹¹ – parmi un très grand nombre d'outils en ligne permettant de récupérer automatiquement une gamme colorée dans une image. Le principe est de charger une image et l'outil construit automatiquement une palette de couleurs. Nous soulignons qu'aucune

9. Des questions peuvent alors se poser sur la qualité requise pour ces images qui sont un autre problème à résoudre.

10. <http://www.degraeve.com/color-palette/>

11. <http://www.pic2color.com>

explication sur la façon dont la sélection de couleurs est effectuée n'est donnée. Quelle que soit l'image fournie, Color Palette Generator donne cinq couleurs ternes et cinq couleurs pures, Pic2Color donne huit couleurs. La Figure 3 montre l'utilisation respective de Color Palette Generator et de Pic2Color sur un tableau d'André Derain¹², la Figure 4 sur le tableau de van Gogh de la Figure 2.

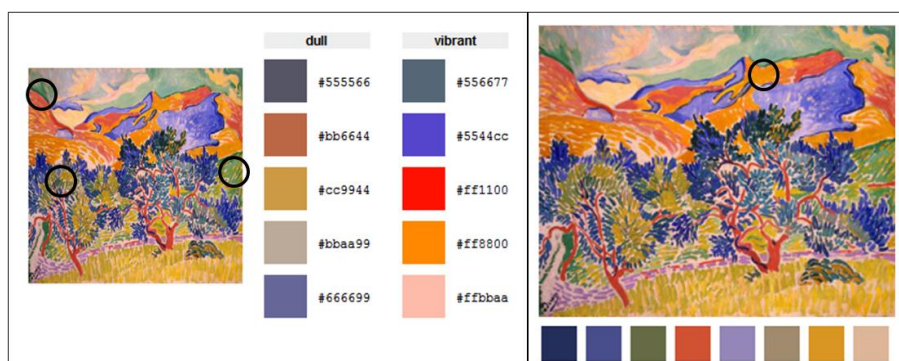


Figure 3. Extraction automatique d'une palette à partir d'un tableau de Derain (à gauche : Color Palette Generator ; à droite : Pic2Color). NB : les cercles noirs marquent des couleurs non détectées.

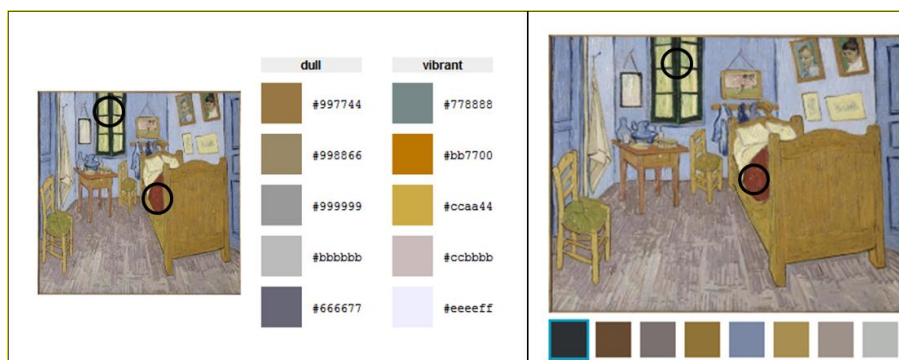


Figure 4. Extraction automatique d'une palette à partir d'un tableau de van Gogh (à gauche : Color Palette Generator ; à droite : Pic2Color). NB : les cercles noirs marquent des couleurs non détectées.

12. André Derain, Montagnes à Collioure, 1905, John Hay Whitney Collection, National Gallery of Art Washington, USA.

Pour les deux tableaux de Derain et de van Gogh testés, même si de nombreuses nuances ressortent bien, il manque des couleurs que nous jugeons comme représentatives du tableau, mises en évidence par des cercles noirs dans les figures :

– pour le Derain : il manque les verts avec le *Color Palette Generator*, et l’orange avec le *Pic2Color* (Figure 3) ;

– pour le van Gogh : il manque le rouge bordeaux avec les deux outils (Figure 4).

Dans notre contexte, nous sommes intéressés par les couleurs identifiées par ces outils, mais également celles qui, même si elles sont peu représentées en nombre de pixels, sont visuellement saillantes. Le rouge bordeaux du tableau de van Gogh est à ce titre caractéristique : la tâche de rouge en petite quantité est centrale et saute aux yeux au milieu des couleurs froides du tableau. Ainsi, la quantité d’information n’est pas forcément reliée à la quantité de couleur présente dans le tableau. Certaines couleurs peuvent être utilisées pour une région colorée, petite mais visuellement saillante, et donc apporter une quantité d’information suffisamment importante pour que nous les considérons comme représentatives du tableau selon notre point de vue applicatif de conception cartographique. Elles pourraient par exemple être utilisées pour mettre une information importante en valeur dans la carte.

Ces outils ne fournissent pas les critères utilisés pour sélectionner les couleurs dans une image rendant difficile leur réutilisation dans le cadre de l’automatisation du processus de conception cartographique. Ces outils qui ne sont ni paramétrables ni utilisables depuis une autre application (sous la forme d’un service ou d’une bibliothèque par exemple) donnent des résultats de couleurs insatisfaisants pour nos besoins.

Travaux de recherche

Il existe un grand nombre de travaux sur les couleurs en peinture sur des sujets variés en informatique graphique : comparaison de peintures et de styles de peintures à partir de l’étude des lignes, lumières, textures et couleurs (Lombardi, 2005), validation de l’utilisation de contrastes des complémentaires chez van Gogh (Berezhnoy, 2004), détection des coups de pinceaux dans les tableaux de van Gogh (Li *et al.*, 2012), entre autres. Barla (2006) définit une palette de couleurs comme un ensemble de classes de couleurs appelées prototypes dans un espace de couleurs donné. Cette palette décrit la distribution des couleurs de régions d’intérêt dans une image selon un paramètre de tolérance. Sa méthode consiste à segmenter l’image en régions d’intérêt, à regrouper les régions voisines de couleurs similaires et à en extraire des prototypes de couleurs à l’aide d’une Analyse en Composantes Principales. Cependant, l’auteur relève qu’il obtient un très grand nombre de prototypes, pouvant être redondants¹³ ou pouvant avoir des variations de couleurs intra-prototype importantes¹⁴.

13. Des régions de couleurs similaires distantes dans l’image ou des couleurs sombres ou désaturées donnent des prototypes redondants.

14. A l’observation, on peut nommer des couleurs différentes dans un prototype alors qu’on ne devrait pouvoir en nommer qu’une seule.

Notre objectif applicatif final étant de pouvoir proposer, au travers d'applications (en ligne ou hors ligne), d'utiliser des peintures pour symboliser des données géographiques nous impose une exigence : pouvoir présenter aux utilisateurs une palette, extraite d'une peinture, d'une dizaine de couleurs différentes.

Les travaux présentés ne répondent pas à nos besoins. Nous allons donc proposer, adapter et tester différentes méthodes afin de réaliser une extraction automatique de couleurs dans des peintures.

4. Proposition de méthodes d'extraction de palettes dans des peintures

Nous souhaitons extraire de l'image d'une peinture des couleurs significatives sous la forme d'une palette de couleurs. Une méthode manuelle et des méthodes automatiques sont présentées dans cette partie. Enfin, nous présentons une méthode de comparaison de palettes. Celle-ci nous permettra de mener une démarche similaire à une classification supervisée d'algorithmes afin de déterminer les méthodes et espaces de couleurs les plus adaptés à notre besoin.

4.1. Méthode manuelle

Dans le travail de (Christophe, 2009), cinq toiles de maîtres ont été analysées et exploitées par le modèle d'aide au choix des couleurs COLorLEGend : pour chacune, des couleurs significatives ont été extraites : le choix des palettes a été guidé par l'importance de certaines couleurs significatives dans le tableau, et l'idée qu'il fallait en sélectionner une dizaine¹⁵. La Figure 5 présente ces cinq peintures et leur palette respective.

¹⁵ Ceci explique que, pour chaque palette, on n'ait pas nécessairement le même nombre de couleurs.

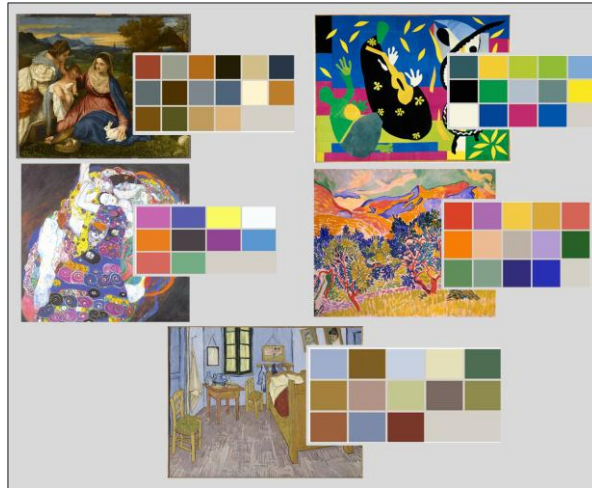


Figure 5. Peintures et palettes associées (Christophe, 2009)¹⁶

Les *couleurs significatives* sont des couleurs du tableau qui ont une importance significative dans l'effet visuel du tableau. Une couleur significative n'est donc pas nécessairement une couleur qui a une grande emprise dans le tableau (donc représentée par un grand nombre de pixels), mais bien une couleur qui participe activement à l'impression visuelle qu'une personne peut avoir de ce tableau : une couleur peu représentée mais très brillante (le jaune de Klimt), deux couleurs dans un contraste précis (le bleu-orange du ciel et le rouge-bleu de la robe dans le Titien), une couleur distribuée de façon très particulière (les collages jaunes ou les aplats bleus et roses de Matisse), des couleurs très différentes (les couleurs du Klimt), des couleurs donnant une sensation de température (le côté froid du tableau de van Gogh, le côté fauve du tableau de Derain).

Nous choisissons ces palettes comme références pour la proposition de méthodes automatiques d'extraction de palettes : les palettes construites automatiquement doivent être similaires à ces palettes de référence.

16. « De gauche à droite et de haut en bas : Le Titien, La Vierge au lapin, ap. 1530, Musée du Louvre, Paris, France ; Henri Matisse, La Tristesse du Roi, 1952, Centre Georges Pompidou, Paris, France. Succession H. Matisse ; Gustav Klimt, Die JungFrau, 1913, Narodni Gallery, Prague, Rép. Tchèque ; André Derain, Montagnes à Collioure, 1905, John Hay Whitney Collection, National Gallery of Art, Washington, USA ; Vincent van Gogh, La Chambre à Arles, 1888, Musée van Gogh, Amsterdam, Pays-Bas. »

4.2. Méthodes d'extraction automatique de palettes

Il s'agit maintenant de proposer et de tester des méthodes automatiques et de confronter leurs résultats avec ceux obtenus lors de l'extraction manuelle. Dans cette partie nous décrivons les différentes méthodes proposées et nous nous concentrons sur le fait de vérifier que les palettes construites sont bien représentatives du tableau initial¹⁷.

Notre objectif est de classer et réduire le nombre de couleurs dans la peinture analysée afin d'en extraire un sous-ensemble représentatif. Afin de résoudre ce problème, nous avons choisi d'utiliser des méthodes de classification non supervisée (clustering). En effet, d'autres méthodes de classification (supervisée) auraient pu être utilisées. Néanmoins, ces méthodes introduiraient un *a priori* en modélisant, dans une certaine mesure, les connaissances des experts, ce qui n'est pas notre objectif. Par ailleurs, un ensemble de méthodes existe dans le domaine de la compression d'image (avec perte) afin de réduire le nombre de couleurs d'une image (et donc sa taille en mémoire). Ces méthodes sont généralement appelées des méthodes de quantification de couleurs (color quantization) : leur objectif est la compression, c'est-à-dire la réduction du nombre de couleurs de l'image tout en conservant sa qualité visuelle (après l'étape de réaffectation ou remapping). Notre objectif est donc différent : en effet, nous ne cherchons pas à reconstruire une nouvelle image avec ces couleurs, mais bien à réutiliser ces couleurs sur un jeu de données géographiques. C'est pourquoi, nous devons trouver une méthode la plus judicieuse possible pour extraire une palette adaptée aux besoins cartographiques. La qualité des palettes extraites est donc, pour nous, à évaluer vis-à-vis de leur utilisation pour la conception de cartes, et non vis-à-vis de l'erreur entre l'image initiale et l'image utilisant la palette extraite. Dans un second temps, les palettes peuvent aussi être évaluées à l'aide des cartes qu'elles ont permis de générer. Ce processus est illustré en figure 1 par la boucle entre la phase d'évaluation de la carte et la phase de choix de la symbologie. Néanmoins, nous n'aborderons pas ces aspects dans cet article pour nous focaliser sur l'extraction de couleurs.

Parmi les méthodes de quantification existantes, trois méthodes ont été testées : les *K-means*, *octree*, classification hiérarchique ascendante. L'approche proposée est illustrée par la figure 6. Tout d'abord, les couleurs de l'image en RVB sont converties dans l'espace de couleur CS (les espaces de couleurs utilisés sont décrits au §4.2.3), considéré comme un espace de travail dans lequel les couleurs sont décrites par des coordonnées et qui est muni d'une métrique (distance entre couleurs D). Elles sont ensuite utilisées par un algorithme de clustering (*K-means*, *octree* ou *CAH*) puis converties en RVB à nouveau. Tous les algorithmes proposés prennent par ailleurs en entrée une distance entre couleurs D et le nombre de couleurs de la palette finale K .

¹⁷ Plus tard, nous testerons si les palettes sont bien adaptées pour ce que l'on veut en faire (il y a du bleu, il y a assez de couleurs, il y a assez de couleurs différentes, etc.)

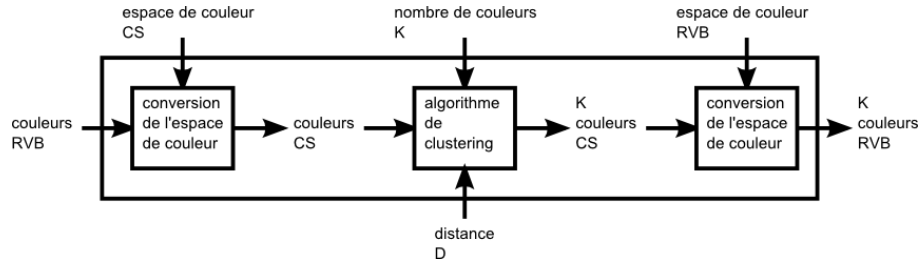


Figure 6. Schéma global de l'approche proposée.

4.2.1. Méthode des *K-means* (Mac Queen, 1967)

La méthode des *K-means* est un outil de classification classique qui permet de répartir un ensemble de données en k classes (cluster) en minimisant un critère : la somme de carrés des distances entre les données associées à une classe et le centroïde de cette classe. L'algorithme des *K-means* permet ainsi de découper l'espace des couleurs de l'image donnée (les données) en K sous-espaces de couleurs (les classes) représentées par leur couleur moyenne (le centroïde de la classe). L'algorithme consiste, à partir de K classes et de leurs centroïdes initiaux, à identifier l'ensemble des données associées à cette classe (les données plus près du centroïde de cette classe que de tous les autres) et à réévaluer ensuite la position des centroïdes. Cette étape est exécutée de façon itérative jusqu'à convergence de l'algorithme (les classes auxquelles sont affectées les données n'ont pas changées après réévaluation des centroïdes). Deux paramètres sont importants pour cet algorithme : l'espace de travail et la distance utilisée pour choisir le centroïde le plus proche de chaque couleur (et donc la classe qui lui est associée).

Afin d'obtenir une solution stable, la méthode des *K-means* (dont les résultats dépendent de l'initialisation, c'est-à-dire des K premiers centroïdes choisis de façon plus ou moins arbitraire) a été lancée plusieurs fois. La meilleure solution vis-à-vis de la méthode des *K-Means* a été conservée ; c'est à dire la solution minimisant la somme des carrés des distances entre chaque couleur d'un cluster et son centroïde.

4.2.2. Méthode de quantification par *Octree* (Gervautz, 1988)

La méthode de l'*octree* est une méthode de partitionnement de l'espace (des couleurs) hiérarchique où chaque nœud de l'arbre peut avoir jusqu'à 8 descendants. On parle de feuille lorsqu'un nœud n'a pas de descendant. L'algorithme consiste à insérer les couleurs à l'arbre en s'appuyant sur la représentation binaire de la couleur pour déterminer sa position dans l'arbre. En effet, l'espace des couleurs est divisé en huit partitions à chaque niveau N de l'arbre. Chaque partition est ainsi découpée en 2 pour chaque composante de la couleur (nous n'utilisons que des espaces de couleur à 3 composantes). Pour ce faire, la couleur étant composée des 3 composantes A , B et C (en pratique RVB, XYZ, etc.), et chaque composante étant représentée à l'aide de 8 bits ($A_0A_1A_2A_3A_4A_5A_6A_7$, etc.), l'index où insérer une nouvelle couleur utilise le N ème bit de chaque composante (A_N , B_N , C_N)

(Bloomberg, 2008). Chaque couleur est initialement insérée au niveau le plus bas (7 pour une image 24 bits) et à chaque insertion, l'arbre est réduit (ou élagué) à chaque fois que le nombre de feuilles de l'arbre dépasse le nombre de couleur recherché. La procédure de réduction consiste à choisir le nœud de plus haut niveau qui vérifie certaines conditions, à en supprimer les feuilles et à marquer ce nœud comme une feuille. On garantit ainsi de ne jamais avoir dans l'arbre plus de feuilles que nécessaire. La figure 7 montre l'octree obtenu sur le tableau de Derain et dont le nombre de couleurs a été réduit de 99227 à 64.



Figure 7. Octree à 64 couleurs (en CIELab sur la peinture de Derain).

Outre l'espace de couleur utilisé, la réduction de l'arbre constitue une étape influente sur les résultats de la quantification par octree. La question à se poser est alors : quel nœud réduire ? Nous avons testé les choix suivants :

- afin de minimiser l'inertie intra-nœuds, on peut choisir le nœud dont l'écartype des distances entre feuilles est le plus faible ;
- on peut, comme pour les K-means, utiliser la somme des carrés des distances intra-nœud.
- ou alors, on peut choisir de réduire le plus petit cluster, c'est-à-dire le cluster dont le rayon (distance maximale entre la moyenne des feuilles et l'ensemble des feuilles) est minimal.

Nous avons choisi d'utiliser la deuxième solution car en éliminant le nœud minimisant la somme des carrés des distances intra-nœud, l'erreur globale est minimisée par la même occasion. En effet, l'erreur faite en conservant le centroïde de ce nœud à la place de toutes les couleurs qui le constituaient est celle qui aura le moins d'impact global.

4.2.3. Méthode de classification hiérarchique ascendante (Ward, 1963)

La méthode présentée ici est une approche ascendante consistant à regrouper progressivement des clusters ne contenant au départ qu'une seule couleur, jusqu'à l'obtention d'une hiérarchie complète de clusters. Les clusters créés portent l'indice

d'agrégation calculé par le critère de liaison. Une fois cette hiérarchie créée, le nombre de couleurs peut être réduit en élaguant l'arbre, par exemple en choisissant l'indice d'agrégation permettant d'obtenir le nombre de couleur souhaité ou le nombre le plus proche possible. En effet, plusieurs clusters pouvant avoir le même indice d'agrégation, il est parfois impossible d'obtenir le nombre de clusters K sans faire de choix arbitraire. Trois facteurs influencent le résultat : l'espace de représentation des couleurs noté CS pour *ColorSpace* dans la figure 8 (RVB, CIE Lab, CIE XYZ), la métrique noté D dans la figure 8 (euclidienne, Manhattan, etc.) et le critère de liaison noté L dans la figure 8. Le critère de liaison est une mesure de la distance ou de la dissimilarité entre clusters. Il existe de nombreux critères de liaison dans la littérature. Le saut minimum (Sibson, 1973) est le minimum des distances entre individus des deux clusters, le saut maximum (Defays, 1977) est le maximum de ces mêmes distances et le lien moyen est la moyenne de celles-ci (Sokal *et al.* 1958, Murtagh, 1984). Enfin, (Szekely *et al.*, 2005) généralisent le critère de Ward, qui représente le minimum de variance inter-cluster (Ward, 1963) en proposant une méthode combinant les distances inter et intra cluster. De nombreuses autres méthodes ont par ailleurs été proposées (Murtagh *et al.*, 2012), mais nous choisissons d'utiliser le critère de Ward afin d'utiliser un critère de liaison similaire aux critères utilisés dans les deux autres méthodes utilisées dans ce travail.

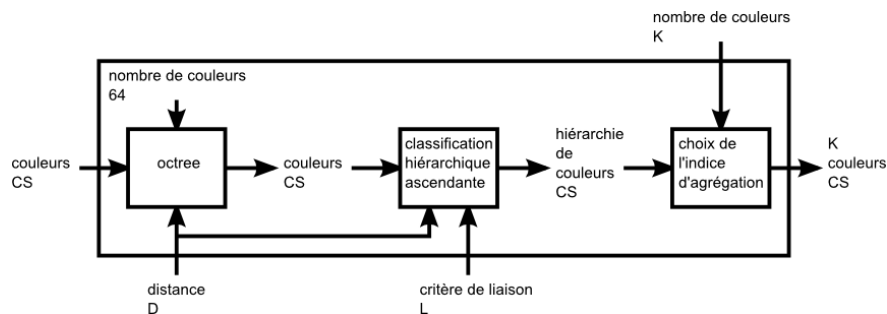


Figure 8. Schéma du fonctionnement de notre méthode de classification hiérarchique ascendante

Pour simplifier les calculs, nous utilisons tout d'abord un octree réduisant le nombre de couleurs à 64 puis nous utilisons la classification hiérarchique ascendante afin de trouver les K couleurs recherchées (voir figure 8). Les Figures 9 et 10 montrent les résultats de la méthode hiérarchique respectivement en 16 et 64 couleurs. Nous obtenons des palettes de couleurs proches des couleurs caractéristiques recherchées.

La palette à 16 couleurs présentée figure 9 peut être proposée telle quelle à l'utilisateur lui permettant d'exprimer ses préférences directement sur les couleurs. Néanmoins, certaines couleurs peuvent paraître très proches visuellement pour un œil non exercé (Ce sont les quantités de couleurs utilisées dans la carte qui feront toute la différence).

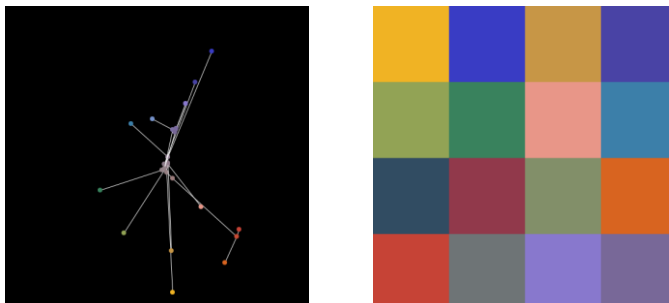


Figure 9. Visualisation des 16 clusters et couleurs de la palette de Derain (RVB)

La palette à 64 couleurs présentée figure 10 est intéressante car elle contient des dégradés en valeur des couleurs caractéristiques permettant d’avoir des couleurs supplémentaires si les 16 couleurs proposées précédemment sont insuffisantes pour produire des cartes satisfaisant l’utilisateur (nombre de thèmes, contrastes à mettre en œuvre suivant relations sémantiques dans la carte).

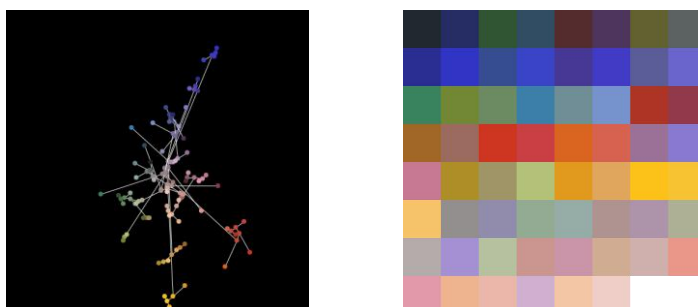


Figure 10. Visualisation des 64 clusters et couleurs de la palette de Derain (RVB)

4.2.4. Améliorations apportées aux méthodes de quantification sur les espaces de couleurs

Traditionnellement, les algorithmes d’indexation et de quantification d’images manipulent des images numériques encodées en couleurs RVB. Afin de mieux représenter l’ambiance colorée des toiles étudiées, nous avons manipulé différents espaces de couleurs dans l’implémentation des différentes méthodes de quantification décrites ci-dessus :

- l’espace RVB suit la décomposition trichromatique de la couleur selon les trois faisceaux lumineux primaires que sont le Rouge, le Vert et le Bleu (Sève 2009).

Cette décomposition se retrouve par exemple dans les bâtonnets de la rétine de l'œil humain ou encore dans les canaux lumineux des écrans modernes de télévision ou d'ordinateur.

- l'espace CIEXYZ est un des premiers systèmes de repérage spatial des couleurs dans un espace à trois dimensions orthogonales et uniformes. Il est utilisé dans l'article surtout comme un espace de passage, de conversion.

- l'espace chromatique uniforme CIELab, mis en place par la CIE en 1976, est fidèle à la perception humaine des couleurs. C'est à dire que les distances euclidiennes entre couleurs dans cet espace ont une signification proche de la perception de l'œil humain (CIE 1976, Sève 2009).

- l'espace YCbCr est utilisé dans le domaine de la vidéo et souvent dans les méthodes de compression par quantification au format jpeg. Il renseigne sur la luminosité en Y et sur la chromaticité bleue en Cb et rouge en Cr. De plus, les différentes sources lumineuses (rouge, verte et bleue) sont pondérées par rapport à la perception de l'œil.

Une des améliorations notables due au choix de l'espace de couleurs est la qualité des distances entre couleurs. En évoluant dans un espace de couleur uniforme comme CIELab, on manipule des distances entre couleurs qui sont significatives pour la perception humaine. Cela permet de mieux regrouper des couleurs similaires lors des répartitions des méthodes ci-dessus :

- en cluster dans la méthode des K-Means,
- en nœud dans la méthode des octrees,
- en classe dans la méthode de classification ascendante hiérarchique.

D'autre part, dans la première étape de l'algorithme de quantification par octree proposé par (Gervautz, 1988) les couleurs de l'image en entrée sont distribuées dans l'octree selon les valeurs des bits de poids fort de leurs coordonnées dans le système RVB. Notre proposition consiste à faire évoluer les couleurs dans l'espace chromatique uniforme CIELab.

Pour cela, nous avons d'abord normalisé les coordonnées CIELab des couleurs du tableau afin qu'elles soient encodées sur 8 bits comme les coordonnées RVB, entre 0 et 255. Puis nous avons utilisé les bits de poids fort de ces coordonnées normalisées pour ordonner les couleurs dans l'octree. Ainsi, le système CIELab représentant en premier lieu la luminosité par la coordonnée L, les couleurs se trouvent organisées d'abord par luminosité puis par chromaticité, ce qui favorise la conservation dans l'octree de couleurs claires et lumineuses. Nous voulions pouvoir conserver ces couleurs claires et lumineuses car elles ne sont pas moins représentatives que d'autres. En outre, elles sont parfois visuellement saillantes et participent à l'effet visuel général du tableau. Mais elles sont plus sensibles que les couleurs sombres lors des étapes de moyennage de couleurs. L'utilisation particulière des coordonnées CIELab décrite ci-dessus a pour but de les préserver en les séparant a priori des couleurs sombres dans l'espace des couleurs.

4.3. Méthode de comparaison de palettes

Afin de comparer les palettes résultantes des méthodes de quantification d'image, nous avons mis au point une mesure de distance entre deux palettes en s'inspirant des distances d'édition de chaîne de caractères. Notre but final étant d'évaluer les méthodes de quantification décrites au paragraphe précédent, nous considérons les palettes extraites manuellement comme une référence, un objectif à atteindre. Les couleurs des palettes résultant des méthodes de quantification sont donc comparées à celles des palettes de référence décrites au paragraphe 4.1.

La première étape consiste à aligner les couleurs de la palette à évaluer avec celles de la palette de référence. Pour cela, pour chaque couleur de la palette à étudier, nous déterminons la couleur la plus proche dans la palette de référence (dans l'espace de couleurs CIELab). Ces deux couleurs seront associées si elles sont assez proches. Un exemple d'alignement entre la palette extraite manuellement et la palette générée par l'outil en ligne pic2Color à partir du tableau de Derain, est donné en Figure 11.

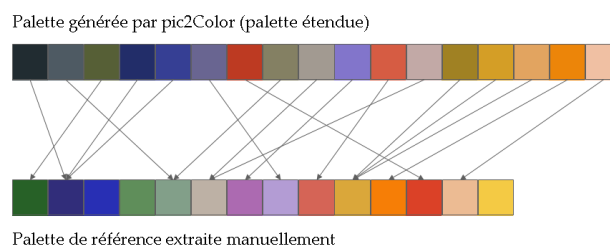


Figure 11. Alignement de deux palettes de couleurs

Pour toutes les couleurs de la palette de référence auxquelles a été associée une unique couleur (entourées en rose sur la figure 12), un coût de transformation de couleur est calculé par la distance euclidienne entre ces deux couleurs. A titre d'exemple, sept couleurs de la palette de la référence ont été associées à une seule couleur de la palette générée par l'outil pic2Color comme le montre la figure 12, ce qui donne un coût de transformation de 1754.

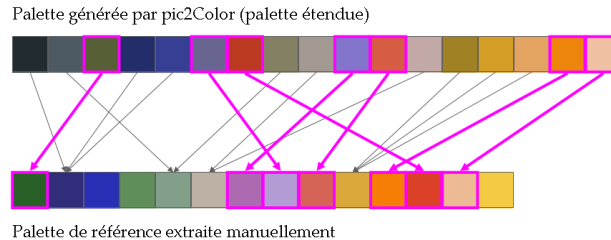


Figure 12. Une unique couleur générée est associée à une couleur de référence : couleurs pour lesquelles on calcule un coût de transformation (entourées en rose)

Pour chaque couleur de la palette de référence à laquelle ont été associées plusieurs couleurs (entourées en bleu foncé sur la figure 13), un coût de fusion de couleurs est calculé par la distance minimale entre les couleurs associées et la couleur correspondante dans la palette de référence (cf. figure 14).

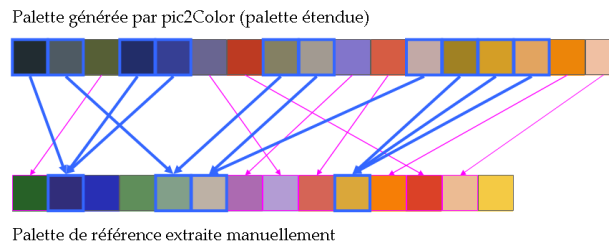


Figure 13. Plusieurs couleurs générées sont associées à une couleur de référence : couleurs (entourées en bleu foncé)

A titre d'exemple, trois couleurs de la palette générée par l'outil pic2Color ont été associées au premier bleu de la palette de référence (elles sont entourées en bleu foncé dans la figure 13). Pour ces trois couleurs, le coût de fusion correspond à la distance entre le bleu de la palette de référence et la couleur la plus proche parmi ces trois couleurs de la palette générée par l'outil pic2Color (entourée en bleu foncé dans la figure 14), les deux autres couleurs (entourées en cyan dans la figure 14) ne sont pas prises en compte dans le coût de fusion. Au total, des associations multiples ont eu lieu pour quatre couleurs de la palette de référence comme le montre la figure 14 ce qui donne un coût de fusion de 421.

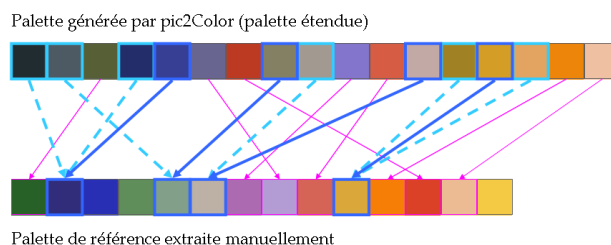


Figure 14. Couleurs pour lesquelles un coût de fusion est calculé (entourées en bleu foncé) et couleurs ne présentant pas une distance minimale à une couleur de la palette de référence donc non prises en compte pour le calcul du coût de fusion (entourées en bleu clair).

Un coût d'ajout a été calculé pour les couleurs de la palette de référence qui n'ont été associées à aucune des couleurs de la palette à étudier. Elles sont entourées en vert sur la figure 15. Le coût d'ajout d'une couleur est égal à la distance euclidienne entre la couleur (de la palette de référence) à ajouter et la couleur la plus proche dans la palette à évaluer. Afin de privilégier certaines couleurs essentielles de la palette de référence, ce coût a été pondéré par cinq¹⁸ pour les couleurs qui avaient été identifiées comme essentielles¹⁹. À titre d'exemple, quatre couleurs de la palette générée par l'outil pic2Color ont été associées à plusieurs couleurs de la palette de référence comme le montre la figure 15. Les couleurs essentielles identifiées dans la palette de référence ont été signalées par des étoiles. Pour ce couple de palettes, le coût d'ajout est de 1597.

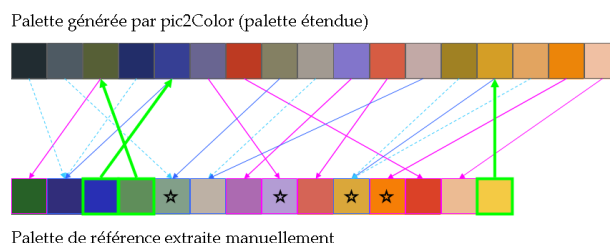


Figure 15. Couleurs de la palette de référence manquantes dans la palette à étudier (entourées en vert)

18. Cette pondération est le résultat de tests empiriques. L'objectif était de privilégier fortement les méthodes conservant les couleurs essentielles.

19. Les couleurs essentielles ont été identifiées par un expert cartographe. Ce sont soit des couleurs saillantes qui participent fortement à l'impact visuel du tableau soit des couleurs qui avaient généralement été perdues avec les outils d'extraction en ligne et qui sont donc potentiellement difficiles à extraire.

Les coûts de transformation, fusion et ajout sont ensuite ajoutés pour évaluer la distance existant entre la palette à étudier et la palette de référence (cf. Equation (1))²⁰. Un poids plus important (double) a été affecté au coût d'ajout, l'absence d'une couleur de la palette de référence dans la palette à étudier étant considéré comme critique. Selon (1), pour notre exemple, cela donne une distance de 5369 :

$$\text{Distance (palette1, palette2)} = \text{coût_transformation} + \text{coût_fusion} + 2 \times \text{coût_ajout} \quad [1].$$

5. Résultats et évaluations

Afin d'évaluer les méthodes d'extraction automatique de palettes de couleurs, une gamme de palettes comprenant entre 10 et 19 couleurs a été générée pour chacune des cinq toiles décrites au paragraphe 4.1, pour chacune des trois méthodes décrites au paragraphe 4.2 et pour chacun des quatre espaces de couleurs décrits au paragraphe 4.2.4., soit un total de $10 \times 5 \times 3 \times 4 = 600$ palettes.

5.1. Comparaison des palettes

5.1.1. Méthode de comparaison automatique

Une première façon de comparer les palettes générées automatiquement est d'observer les palettes issues d'un tableau avec les différentes méthodes, dans les différents espaces de couleurs, pour un même nombre de couleurs initial. Un tel jeu de palettes est présenté en figure 16 pour le tableau de van Gogh, ordonné par la distance entre palettes présentée au paragraphe 4.3, de la meilleure à la plus mauvaise.

20. Un coût de suppression pourrait être ajouté pour tenir compte des couleurs qui n'ont pas été associées à une couleur de la palette de référence. Cependant, notre objectif premier étant avant tout d'obtenir les couleurs essentielles, nous n'avons pas pénalisé les palettes qui présentaient d'autres couleurs, d'autant plus qu'elles pourraient avoir un intérêt pour une application de conception cartographique. En conséquence, nous n'avons pas défini de seuil de distance en dessous duquel deux couleurs seraient considérées comme proches : les couleurs sont donc toujours associées à une couleur de la palette de référence. Cela pourrait être modifié selon d'autres objectifs de comparaison de palettes.

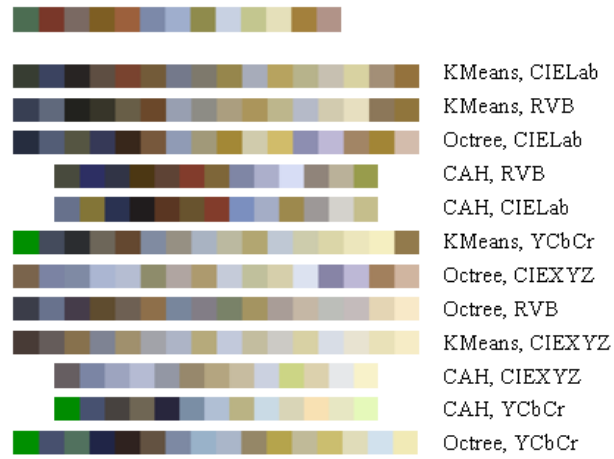


Figure 16. Palette de référence et Palettes de 16 couleurs générées par les différentes méthodes d'extraction automatique, dans les différents espaces de couleurs

La figure 17 montre quant à elle l'évolution des distances entre ces palettes et la palette de référence.

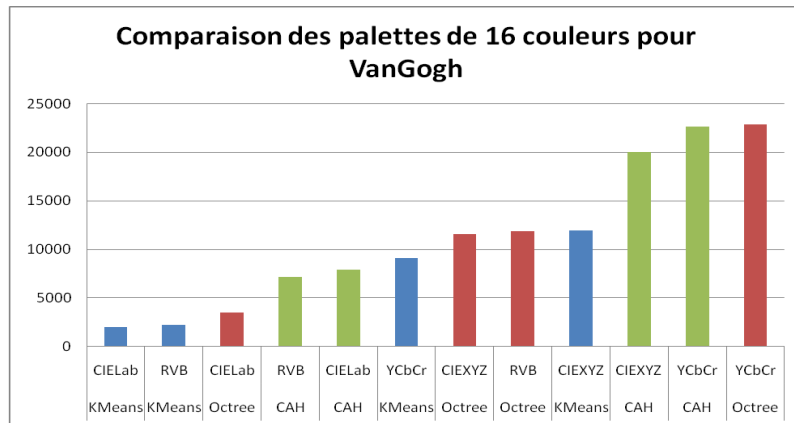


Figure 17. Classement des palettes de couleurs par similarité à la palette de référence

Les méthodes d'extraction utilisées dans cet article (cf. §4) comportent toutes des étapes de moyennage des couleurs, c'est pourquoi certaines palettes présentent des couleurs inexistantes initialement dans les peintures. De plus, on remarque la présence de trois verts particulièrement vifs que l'on ne retrouve pas dans le tableau de van Gogh. Ces verts ont été extraits dans l'espace de couleur YCbCr dont la répartition des couleurs offre une grande part de l'espace aux verts. L'extraction de couleurs légèrement étrangères au tableau peut donc être due soit à la méthode d'extraction soit à l'espace de travail.

Ces résultats nous montrent que les couleurs majoritaires du tableau se retrouvent généralement dans les palettes résultantes. Cependant, certaines de nos couleurs significatives (le rouge de l'édredon ou le vert foncé de la fenêtre) apparaissent moins souvent. Nous considérons que les méthodes qui les conservent sont les meilleures.

En revanche, on remarque que les palettes issues de la méthode des K-Means sont un peu plus ternes que celles issues de la CAH (la quatrième et la cinquième). Cela est une conséquence de l'utilisation du centroïde intrinsèque à la méthode des K-Means. Cela n'est cependant pas rédhibitoire, les palettes issues du K-Means restent chromatiquement plus proches de celle de l'expert. D'autre part, dans les palettes contenant un grand nombre de couleurs, certaines couleurs sont inexistantes dans les peintures. Ces couleurs modifient l'impression globale de la palette, l'éloignant ainsi de l'ambiance colorée de la peinture initiale. Il serait donc peut-être pertinent de modifier le coût de suppression dans la méthode de distance entre palettes afin de pénaliser ces couleurs.

5.1.2. Réaffectation des couleurs

Afin d'évaluer nos distances entre palettes, nous les réaffectons une à une dans l'image initiale de la peinture et dans une carte précédemment réalisée avec la palette de référence. La Figure 18 présente la réaffectation des palettes ayant obtenu respectivement la plus petite distance, la distance moyenne et la plus grande distance à la palette extraite de la peinture de van Gogh, dans l'ensemble des palettes entre 10 et 19 couleurs.

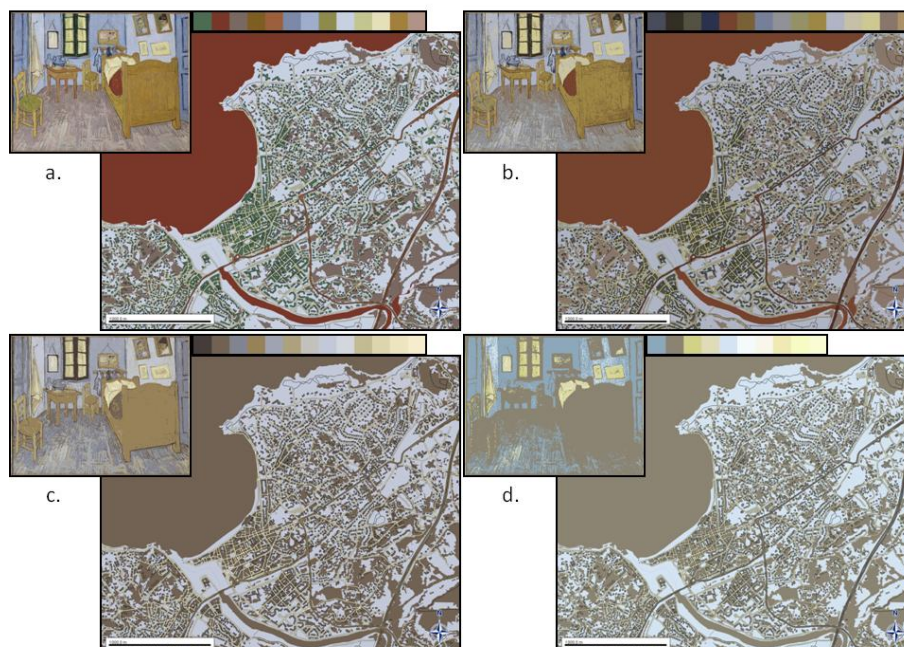


Figure 18. Application de la palette de références (a) et application des couleurs des palettes extrêmes : b) la meilleure, c) la moyenne et d) la pire

La meilleure réaffectation dans le tableau reproduit bien l'impression initiale de la peinture. Les deux autres palettes sont de plus en plus ternes et s'éloignent de la peinture initiale. En revanche, les contrastes clair-obscur sont bien conservés, même dans la plus mauvaise palette, ce qui se retrouve également dans les réaffectations similaires sur la peinture du Titien. En ce qui concerne la réaffectation dans les cartes, le résultat est sensiblement identique. La carte obtenue avec la meilleure palette est cartographiquement acceptable dans le sens où elle présente six thèmes symbolisés différemment et suffisamment contrastés pour être lisibles. Les deux autres cartes ne sont plus lisibles : les objets restent parfois identifiables grâce aux contrastes clair-obscur mais le message cartographique est perdu par manque d'organisation et de hiérarchisation.

5.2. Evaluation des méthodes proposées

Les figures 19 et 20 illustrent l'évaluation quantitative des différentes méthodes proposées dans cet article appliquée aux toiles de Klimt et Matisse respectivement. On remarque, dans un premier temps, une relative fluctuation de la qualité des palettes générées par le même algorithme en fonction du nombre de couleurs

demandé. Ce résultat, qui peut paraître contre-intuitif, est en fait expliqué par la nature des méthodes utilisées. En effet, pour la méthode des K-means par exemple, l'ajout d'une classe peut complètement modifier la répartition dans l'espace (colorimétrique) des classes et, potentiellement, diminuer la qualité globale de la palette (même si la somme des distances intra-classe reste faible).

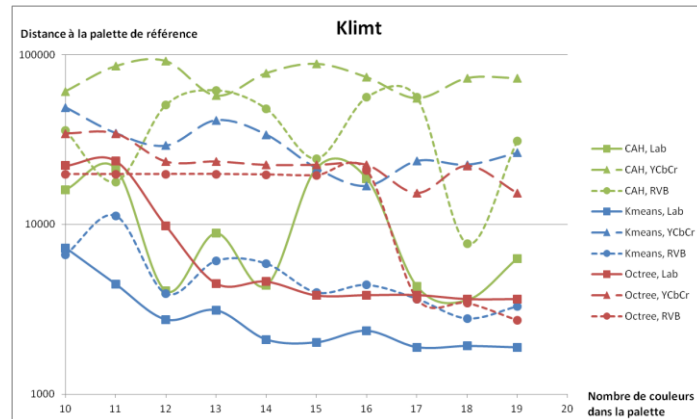


Figure 19. Classement des palettes générées automatiquement pour Klimt

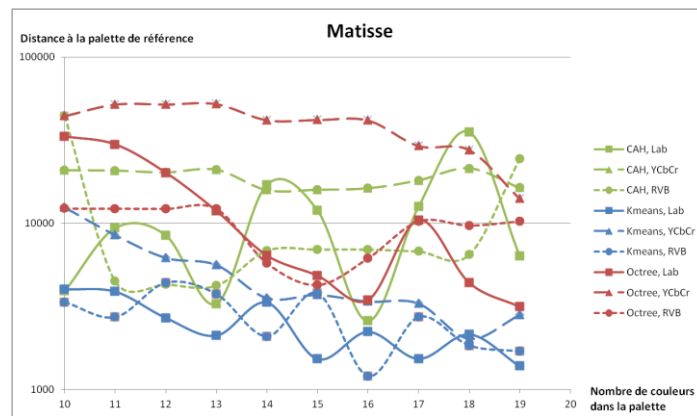


Figure 20. Classement des palettes générées automatiquement pour Matisse

Par ailleurs, on constate que les résultats des différents algorithmes sont le plus souvent meilleurs dans l'espace CIELab pour une méthode donnée. En effet, la distance euclidienne dans cet espace colorimétrique, étant mieux corrélée à la perception de l'œil, permet aux algorithmes de mieux classer les couleurs des toiles et de se rapprocher des résultats proposés par les experts. Pour finir, nous constatons que la méthode des K-Means semble fournir les meilleurs résultats dans la plupart des cas.

6. Conclusions et Perspectives

L'objectif de cet article est de proposer, tester et évaluer des méthodes automatiques d'extraction de palettes dans des peintures. Nous avons cherché à obtenir, pour une peinture donnée, une palette réduite de couleurs significatives pouvant être ensuite présentée à un utilisateur pour qu'il exprime des préférences de couleurs. Nous avons comparé les différentes palettes obtenues automatiquement à la palette de référence obtenue manuellement par un expert, en analysant systématiquement les distances entre les palettes, dans différents espaces de couleurs. Concernant la comparaison des méthodes, nous concluons que la méthode des K-Means est la plus efficace. Concernant la comparaison des espaces de couleurs, les résultats nous confirment qu'il est plus judicieux d'effectuer ces travaux dans l'espace de couleur CIELab. De nombreuses améliorations sont possibles concernant les méthodes proposées. En effet, de nombreuses autres méthodes de classification pourraient être utilisées pour résoudre ce problème (Kaufman *et al.*, 1990) et nous avons choisi dans cet article de nous concentrer sur trois méthodes simples et pertinentes. En particulier, la méthode des K-medoids (Theodoridis *et al.*, 2006) nous paraît pertinente puisque, partant du même principe général que les K-means, elle possède l'avantage de ne pas moyenner les couleurs (ce qui est un inconvénient que nous avons constaté dans la plupart des algorithmes) en utilisant le médonoïde, c'est-à-dire un individu appartenant aux données, comme représentant de la classe au lieu de son centroïde. De nombreuses méthodes ont par ailleurs été proposées pour l'utilisation de médonoïdes dans les méthodes de clustering telles que PAM (Kaufman *et al.*, 1990) et PAMSIL (Van der Laan, 2003) ou encore (Park *et al.* 2009). Enfin, la méthode DBKMedoids (Pratap *et al.*, 2011) combine les propriétés de la méthode K-medoids avec celles de la méthode DBSCAN (Ester *et al.*, 1996) basée sur la densité et permettant de créer des clusters de forme quelconque, alors que les méthodes K-means et K-medoids partitionnent l'espace en cellules de Voronoï.

Si nous comparons les palettes obtenues avec les outils sur Internet, présentées en figures 3 et 4, nous voyons qu'avec les méthodes automatiques, nous obtenons plus de couleurs dans le tableau : à la fois des couleurs différentes et des nuances autour de certaines couleurs caractéristiques. Le faible nombre de couleurs obtenues avec les outils en ligne est dû à l'absence de paramétrage de ces outils, ce qui est un inconvénient majeur dans notre cas. Avec les méthodes automatiques, nous obtenons également des couleurs saillantes qui ne représentent pas nécessairement un grand nombre de pixels. L'utilisation de ces méthodes doit être poursuivie en parallèle d'une analyse de la caractérisation des couleurs dans une image pouvant venir alimenter automatiquement notre processus d'aide à la conception de légendes : par exemple, il serait utile d'analyser le rendu graphique des traits dans une image choisie par l'utilisateur pour l'aider à représenter graphiquement les réseaux de transport parmi ses données géographiques.

Maintenant que nous avons extrait des palettes de couleurs significatives, sans a priori lié au domaine cartographique, il sera intéressant de les confronter à diverses contraintes que l'utilisateur pourra rencontrer pendant la conception de sa carte :

- des contraintes conventionnelles : la palette doit avoir au moins un bleu et un vert pour représenter le réseau hydrographique et la végétation ;
- des contraintes sur la variété des couleurs : la palette doit contenir suffisamment de couleurs différentes pour représenter les différents thèmes de l'utilisateur ;
- des contraintes sur la représentation des relations sémantiques : la palette doit contenir des dégradés de certaines de ses couleurs pour pouvoir représenter une relation d'ordre par exemple.

La palette extraite devra donc être étendue selon des méthodes restant à proposer afin de répondre à ces exigences.

7. Références

- Barla P. (2006). *Representation and acquisition models for expressive rendering*. Thèse de doctorat, INP Grenoble.
- Beconyte G. (2011). « Cartographic styles: criteria and parameters », *In Proc. 25th International Cartographic Conference (ICC'11)*, 3-5 July 2011, Paris, France. URL : http://icaci.org/files/documents/ICC_proceedings/ICC2011/
- Berezhnoy I., Postma E., van den Herik J. (2004). Digital Analysis of van Gogh's Complementary Colors, *Proceedings of the 16th Belgian-Dutch Conference on Artificial Intelligence (BNAIC'04)*.
- Bertin J. (1967). *Sémiologie graphique: les diagrammes, les réseaux, les cartes*, Paris, Mouton/ Gauthiers-Villars.
- Bloomberg D. S. (2008). *Color quantization using octrees*, Leptonica. <http://www.leptonica.org/papers/colorquant.pdf>
- Bord J.-P. (2008). De la cartographie à la géomatique : une amnésie récurrente, *Le Monde des Cartes, Comité Français de Cartographie*, vol. 197, p. 59-62.
- Brewer C. A. (1992). Review of colour terms and simultaneous contrast research for cartography, *Cartographica*, vol.29, n°3&4, p. 20-30.
- Brewer C. A. (1994). Chapitre 7 - Color use guidelines for mapping and visualization, *Visualization in modern cartography*, A.M. Mac Eachren and D.R.F. Taylor, Elsevier Science, Tarrytown, NY, p. 123-147.
- Brewer C. A. (2003). A transition in improving maps: the ColorBrewer example, *Cartography and Geographic Information Science*, vol. 30, n°2, p. 159-162.
- Buard E., Ruas A. (2009). Processes for improving the colours of topographic maps in the context of maps on-demand, *Proceedings of the 24th International Cartographic Conference*, Santiago du Chili, 14-19 Novembre 2009.
- Chesneau E. (2006). *Modèle d'amélioration automatique des contrastes de couleurs en cartographie*. Thèse de doctorat, Université Paris Est-Marne La Vallée.

- Chevreur E. (1839). *De la loi du contraste simultané des couleurs et de l'assortiment des objets colorés*, Paris, Pitois-Levrault et Cie.
- Christophe S. (2009). *Aide à la conception de légendes personnalisées et originales : proposition d'une méthode coopérative pour le choix des couleurs*. Thèse de doctorat, Université Paris Est.
- Christophe S. (2011). Creative colours specification based on knowledge (ColorLegend System), *The Cartographic Journal*, International Cartographic Conference, Paris 2011-Special Issue May 2011, vol.48 n°2, p. 138–145.
- Christophe S. (2012). Cartographic Styles between traditional and original (towards a cartographic style model), *In Proc. AutoCarto Conference 2012*, Columbus, Ohio, USA, 16-18 Septembre 2012 (*to be published*).
- Christophe S., Zanin C., Roussaffa H. (2011). Colours harmony in cartography, *In Proc. 25th International Cartographic Conference (ICC'11)*, Paris, France, 3rd-8th July 2011. URL : http://icaci.org/files/documents/ICC_proceedings/ICC2011/
- CIE. (1976). *Commission Internationale de l'éclairage*. Rapport technique, Cambridge University Press, Cambridge.
- CNIG. (2011). *Prescriptions nationales pour la dématérialisation des documents d'urbanisme : Annexe E – Recommandation pour la représentation – Sémiologie graphique*. Rapport technique, août 2011.
- Defays D. (1977). An efficient algorithm for a complete link method, *The Computer Journal*, vol. 20, n°4, January 1977, p. 364-366.
- Denègre J. (2005). *Sémiologie et conception cartographique*, Paris, Hermès-Lavoisier.
- Dhee F. (2011). Improving topographic maps for color-blind users, *In Proc. 25th International Cartographic Conference (ICC'11)*, Paris, France, 3rd-8th July 2011. URL : http://icaci.org/files/documents/ICC_proceedings/ICC2011/
- Ester M., Kriegel H.-P., Jörg S., Xu X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise, *In Proc. of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, p. 226-231.
- Fabrikant S. I., Rebich-Hespanha S., Hegarty M. (2010). Cognitively inspired and perceptually salient graphic displays for efficient spatial inference making, *Annals of the Association of American Geographers, Association of American Geographers*, vol.100, n°1, p. 17–29.
- Gage J. (2006). *Colour in Art*, Thames & Hudson World of Art.
- Gervautz M., Purgathofer W. (1988). A simple method for color quantization : octree quantization, *New trends in computer graphics*, Academic press Professional, Inc. San Diego, CA, USA, p. 219-231.
- Hoarau C. (2012). Orthoimage or map visualization in use in the french geoportal API, *In Proc. AutoCarto Conference 2012*, Columbus, Ohio, USA, 16-18 Septembre 2012 (*to be published*).

- Kandinsky V. (1989). *Du spirituel dans l'art, et dans la peinture en particulier*, Paris, Gallimard, coll. « Folio Essais ».
- Kaufman L., Rousseeuw P. J. (1990). *Finding Groups in Data*. Hoboken, NJ, USA: John Wiley & Sons, Inc., Mars 1990.
- Kent A.J., Vujakovic P. (2009). Stylistic Diversity in European State 1:50 000 Topographic Maps, *The Cartographic Journal*, vol.46, n°3, p.179–213.
- Klippel A., Hardisty F., Li R., Weaver C. (2009). Color enhanced star plot glyphs - Can salient shape characteristics be overcome?, *Cartographica*, vol.44, n°3, p. 217-232.
- Li J., Yao L., Hendriks E., Wang J. Z. (2012). Rhythmic brushstrokes distinguish van Gogh from his contemporaries: findings via automated brushstroke, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, n°6, p. 1159-1176.
- Itten J. (1977). *Art de la couleur*, Paris, Dessain et Tolra.
- Lombardi T. (2005). *The Classification of Style in Fine-Art Painting*. ETD Collection for Pace University, Paper AAI3189084, Paper presented at the Pace University CSIS Student/Faculty Research Day, White Plains, New York, May 2005. <http://digitalcommons.pace.edu/dissertations/AAI3189084>.
- MacEachren A. (1995). *How Maps Work: Representation, Visualization and Design*, New York, Guilford Press.
- MacQueen J. B. (1967). Some methods for classification and analysis of multivariate observations, *In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California press, vol. 1, p. 281-297.
- Mersey J. (1990). Colour and thematic map design: The role of colour scheme and map complexity in choropleth map communication, *Cartographica*, vol.27, n°3, p. 129-138.
- Mollard-Desfour A. (1998). *Le dictionnaire des mots et expressions de couleur. Le Bleu*, CNRS Editions.
- Murtagh F. (1984). Complexities of hierarchic clustering algorithms: state of the art, *Computational Statistics Quarterly*, vol. 1, p. 101-113.
- Murtagh F., Contreras P. (2012). Methods of Hierarchical Clustering, *Data Mining and Knowledge Discovery*, Wiley-Interscience, vol. 2, n° 1, p. 86–97.
- Park H.-S., Jun C.-H. (2009). A simple and fast algorithm for k-medoids clustering, *Expert Systems with Applications*, vol. 36, n°2, Mar. 2009, p. 3336-3341.
- Pastoureau M. (1992). *Dictionnaire des couleurs de notre temps. Symboliques et société contemporaines*, Paris, C. Bonneton, 232p.
- Raghuvira Pratap A., Vani K. S., Devi J. R., Rao K. N. (2011). An efficient density based improved k- medoids clustering algorithm, *International Journal of Advanced Computer Sciences and Applications*, vol. 2, n°6, p. 49-54.
- Roque G. (2009). *Art et science de la couleur : Chevreul et les peintres, de Delacroix à l'abstraction*, nouvelle édition revue et augmentée, Paris, Gallimard (Tel), 657 p.
- Sève R. (2009). *Science de la couleur, Aspects physiques et perceptifs*, Marseille, Chalagam.

- Sibson R. (1973). SLINK: An optimally efficient algorithm for the single-link cluster method, *The Computer Journal*, vol. 16, n° 1, Jan. 1973, p. 30-34.
- Sokal R. R., Michener C. D. (1958). A statistical method for evaluating systematic relationships, *University of Kansas Scientific Bulletin*, vol. 28, p. 1409-1438.
- Sutter D. (1858). *Philosophie des beaux-arts appliquée à la peinture*, Paris, Jules Tardieu, 355p.
- Sutton T., Whelan B. (2004). *The complete color harmony: expert color information for professional color results*, Gloucester, MA, Rockport Publishers, 216p.
- Szekely G. J., Rizzo M. L. (2005). Hierarchical clustering via joint Between-Within distances: Extending ward's minimum variance method, *Journal of Classification*, vol. 22, n°2, Sep. 2005, p. 151-183.
- Theodoridis S., Koutroumbas K. (2006). *Pattern recognition*, Orlando, Florida, USA, Academic press inc., 3^e édition, 635 p.
- Van der Laan M., Pollard K., Bryan J. (2003). A new partitioning around medoids algorithm, *Journal of Statistical Computation and Simulation*, vol. 73, n° 8, Aug. 2003, p. 575-584.
- Ward J. H. (1963). Hierarchical grouping to optimize an objective function, *Journal of American Statistical Association*, vol. 58, n°301, p. 236-244.

Article reçu le : 1er septembre 2011

Article accepté le : 21 juin 2012

Sidonie Christophe est Chargée de Recherche du Ministère de l'Ecologie, du Développement Durable et de l'Energie, affectée au laboratoire « Cartographie et Géomatique » COGIT de l'Institut National de l'Information Géographique et Forestière (IGN). Ses travaux portent sur les connaissances et processus en jeu dans la conception cartographique, en particulier l'aide au choix des variables visuelles. Elle est responsable du projet ANR MAPSTYLE portant sur la proposition de nouveaux rendus en cartographie.

Julien Perret est Chargé de Recherche du Ministère de l'Ecologie, du Développement Durable et de l'Energie, affecté au laboratoire COGIT de l'IGN. Ses travaux portent sur l'étude des dynamiques urbaines par la modélisation de systèmes complexes, leur analyse et leur simulation.

Charlotte Hoarau est Ingénieur des Travaux Géographiques et Cartographiques de l'Etat (ITGCE) et doctorante au laboratoire COGIT de l'IGN. Elle s'est intéressée à l'amélioration des couleurs d'une carte topographique, notamment dans le cadre d'une visualisation sur un dispositif mobile. Actuellement, son travail de thèse porte sur la conception de représentations cartographiques intermédiaires entre le réalisme des orthophotographies et l'abstraction des cartes topographiques.