



HAL
open science

Etude de l'apport de la reconstruction des régions occultées du visage pour la reconnaissance des expressions

Delphine Poux, Benjamin Allaert, Nacim Ihaddadene, Ioan Marius Bilasco, Chaabane Djeraba

► To cite this version:

Delphine Poux, Benjamin Allaert, Nacim Ihaddadene, Ioan Marius Bilasco, Chaabane Djeraba. Etude de l'apport de la reconstruction des régions occultées du visage pour la reconnaissance des expressions. CORESA, Nov 2018, Poitiers, France. hal-01993650

HAL Id: hal-01993650

<https://hal.science/hal-01993650v1>

Submitted on 25 Jan 2019

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 :

Etude de l'apport de la reconstruction des régions occultées du visage pour la reconnaissance des expressions

Auteurs :

Delphine Poux [1,2], Benjamin Allaert [1], Nacim Ihaddadene [2], Ioan Marius Bilasco [1], Chaabane Djeraba [1]

[1] Centre de Recherche en Informatique Signal et Automatique de Lille, Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISTAL -, F-59000 Lille, France

[2] ISEN Lille, Yncrea Hauts-de-France, France

Résumé (100-200 mots) :

La reconnaissance des expressions faciales en contexte d'interaction non contrôlée est, encore aujourd'hui, confrontée à un certain nombre de défis qui réduisent considérablement les performances des systèmes. Parmi ces défis, les occultations induites par les variations de pose de la tête ou par des objets externes (accessoires, mains, ...) impliquent une perte d'informations sur le visage. Différentes méthodes proposées dans la littérature tendent à résoudre ces problèmes soit en reconstruisant les régions occultées, soit en exploitant uniquement les informations au sein des régions visibles du visage. Dans cet article, nous étudions l'apport de la reconstruction des régions occultées, par rapport à l'exploitation efficace de l'information contenue dans les régions visibles du visage lors de la reconnaissance des expressions faciales.

Mots-clefs (3 à 5 max.) :

Expressions faciales, Occultations, Mouvement, Reconstruction

Contexte et état de l'art :

Les systèmes de reconnaissance des expressions faciales proposés dans la littérature obtiennent de très bonnes performances dans des conditions contrôlées (pose fixe, visage frontal). Cependant, dans un contexte d'interaction naturelle, où la personne est libre de ses mouvements, les variations de pose ou les gestes occultant le visage engendrent une perte d'information. Pour pallier cette perte d'information, deux familles de méthodes se distinguent : une basée sur la reconstruction de l'information manquante et une basée sur l'exploitation optimale des informations visibles, non-impactées par l'occultation. Les méthodes basées sur la reconstruction amènent le visage dans un contexte idéal d'analyse. Cependant, bien que ces méthodes soient efficaces pour la reconnaissance faciale, elles restent encore mal adaptées pour la reconnaissance des expressions faciales, car elles ne préservent pas efficacement la géométrie du visage [1]. Quant aux méthodes exploitant uniquement les régions visibles, elles garantissent que les données extraites sur le visage ne subissent aucun biais induit par un pré-traitement.

Travail proposé (Décrire l'objectif du travail et indiquer clairement le problème/défi technique étudié) :

Dans cette étude nous comparons l'apport d'une méthode de reconstruction du mouvement par rapport à une méthode de sélection du mouvement dans le cadre de la reconnaissance des expressions en présence d'occultations. Nous exploitons le mouvement facial afin de reconnaître les expressions en nous appuyant sur le descripteur LMP et le modèle de segmentation facial utilisant 25 régions, introduits en [3]. Poux et al. [2] montrent qu'un descripteur de mouvement est pertinent en présence d'occultations du visage car le mouvement extrait à proximité des régions occultées contient des informations importantes induites par la propagation du mouvement.

La méthode de reconstruction proposée estime le mouvement des régions occultées en s'appuyant sur le mouvement de la région symétrique associée. La méthode de sélection suit le processus de sélection de régions significatives introduit dans [2]. Cette méthode consiste à analyser l'importance de chaque région du visage pour une expression et une occultation spécifique. Ceci implique de construire un modèle d'apprentissage par expression et par occultation. Afin de suivre un protocole de validation commun, un modèle d'apprentissage par expression est construit pour les deux méthodes. En nous inspirant de [2], nous mettons en place un protocole expérimental permettant de comparer la méthode de reconstruction et la méthode de sélection. La base de données CK+ [4] est utilisée pour comparer ces deux méthodes. Le jeu de données utilisé contient 374 séquences provenant de 123 participants. Les séquences couvrent six

expressions (colère, dégoût, peur, joie, tristesse et surprise). Les visages sont enregistrés dans de parfaites conditions (pose fixe, visage frontal). D'importantes occultations synthétiques, représentées dans la Figure 1, sont ensuite appliquées sur les visages afin de mieux comparer les performances des deux méthodes d'analyse.

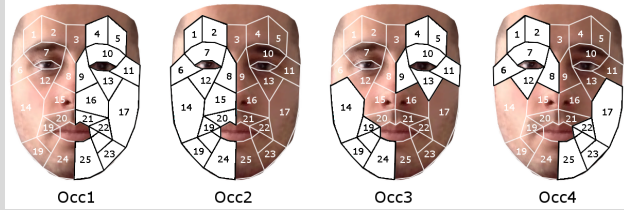


Fig.1 – Représentation des occultations appliquées sur les données de CK+ [Occ 1 et 2: variations de pose ; Occ 3 et 4: objets externes (lunettes, cheveux, main)].

Résultats :

Les performances obtenues pour l'ensemble des occultations sont représentées dans le tableau 1. Le taux de reconnaissance correspond à la moyenne des taux obtenus sur l'ensemble des occultations par expression. Les taux de reconnaissances sont obtenus en utilisant un classifieur SVM et un protocole de validation croisée en 10-fold. Concernant la méthode par reconstruction, le modèle d'apprentissage est construit en s'appuyant sur les informations des 25 régions. Quant à la méthode exploitant les régions visibles, pour chaque expression, seules les 5 meilleures régions pour chaque expression sont utilisées. Ces 5 régions sont définies par [2], où les auteurs montrent qu'en exploitant uniquement 5 régions par expression, les performances obtenues sont similaires aux performances obtenues en exploitant l'ensemble des 25 régions.

Méthodes	Joie	Peur	Surprise	Colère	Dégoût	Tristesse	Moyenne
Reconstruction (25R)	94.25% Std : 2.6	77.46% Std : 6.3	96.13% Std : 1.0	84.81% Std : 2.8	91.75% Std : 2.6	86.78% Std : 3.8	88.53%
Régions visibles (5R)	94.92% Std : 0.1	89.48% Std : 1.0	96.75% Std : 1.0	86.40% Std : 2.0	95.05% Std : 2.0	90.89% Std : 1.8	92.25%

Les résultats montrent qu'en s'appuyant sur un sous-ensemble de 5 régions du visage pour chaque expression, la méthode exploitant uniquement les régions visibles obtient de meilleures performances que celle basée sur la reconstruction. Cependant, celle-ci nécessite de créer un modèle d'apprentissage par expression et par occultation. Quant à la méthode de reconstruction, basée sur les 25 régions du visage, elle a l'avantage de nécessiter un unique modèle d'apprentissage. Il est toutefois important de considérer que les performances de ces méthodes dépendent fortement de leur capacité à reconstruire les régions occultées du visage tout en conservant fidèlement la géométrie faciale.

Conclusion et perspectives :

Dans le cadre de la reconnaissance des expressions faciales, nous avons comparé deux méthodes, une par reconstruction et une par sélection, visant à réduire l'impact de la perte d'information induite par des occultations. Dans ce papier, nous montrons que pour analyser les expressions faciales en présence d'occultations partielles du visage, il est préférable d'exploiter uniquement l'information contenue dans les régions visibles au lieu de reconstruire les régions occultées. Cette solution permet de s'affranchir des problèmes d'estimation des méthodes de reconstruction, qui peuvent avoir de mauvaises répercussions sur la reconnaissance des expressions. De plus, elle permet d'obtenir des performances significatives tout en s'appuyant sur un sous-ensemble de régions, ce qui met en évidence la robustesse de cette méthode en présence d'importantes occultations. Cependant, la construction d'un modèle d'apprentissage par expression et par occultation est nécessaire, contrairement à la méthode basée sur la reconstruction qui nécessite un seul modèle contenant toutes les régions.

Comme perspective, dans un premier temps, nous souhaitons étudier de manière plus fine l'apport de la reconstruction du mouvement au niveau de chaque région. Cela permettra d'avoir une meilleure compréhension du gain apporté par la reconstruction sur les différentes régions du visage en fonction des différentes expressions et des occultations subies. Dans un second temps, nous allons approfondir notre étude en présence d'occultations induites par de réelles variations de pose et d'objets externes.

Références (5 max.):

- [1] Allaert, B., Mennesson, J., Bilasco, I. M. & Djeraba, C. (2018). Impact of the face registration techniques on facial expressions recognition. *Signal Processing: Image Communication*, 61, 44-53.
- [2] Poux, D., Allaert, B., Mennesson, J., Ihaddadene, N., Bilasco, I. M., & Djeraba, C. (2018). Mastering Occlusions by Using Intelligent Facial Frameworks Based on the Propagation of Movement. In *International Conference on Content-Based Multimedia Indexing (CBMI)*.
- [3] Allaert, B., Mennesson, J., Bilasco, I. M. & Djeraba, C. (2018). Advanced local motion patterns for macro and micro facial expression recognition. *arXiv preprint arXiv :1805.01951*.
- [4] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *Computer Vision and Pattern Recognition Workshops (CVPRW), Computer Society Conference on* (pp. 94-101). IEEE.