Signal Processing: Image Communication 61 (2018) 44–53

Contents lists available at ScienceDirect

Signal Processing: Image Communication

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

Impact of the face registration techniques on facial expressions recognition

B. Allaert [a,*], J. Mennesson [b], I.M. Bilasco [a], C. Djeraba [a]

[a] Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRIStAL - Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France
[b] IMT Lille Douai, Univ. Lille, CNRS, UMR 9189 - CRIStAL - Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France

A R T I C L E   I N F O

Keywords:
Facial expressions
Head pose variations
Large displacements

A B S T R A C T

Recent methodologies for facial expression recognition have been proposed and have obtained good results in near-frontal view. However, these situations do not fairly represent in-the-wild challenges, where expressions are natural and the subject is free of its movement. This is reflected in the accuracy drop of facial expression methods obtained on recent databases. Two challenges (head pose variations and large displacements) in facial expression recognition are studied in this paper. Experiments are proposed in order to quantify the impact of free head movements using representative expression recognition approaches (LBP, LBP-TOP, HOOF). We propose an experimental protocol (SNaP-2DFe) that records, under controlled light, facial expressions with two cameras: one attached on the head and one placed in front of the subject. As in both cameras facial expressions are the same, differences in performances measured on each camera show the impact of head pose variations and large displacements on the underlying recognition approach.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Facial expression recognition has attracted great interest over the past decade in various domains. Given the significant role of the face in human communication, several researches have been conducted on facial expression recognition in various contexts.

Several systems evaluate their performances on image collections, where facial expressions are played by actors, in order to obtain exaggerated facial deformations (acted expressions). Several approaches [1–3] obtain very good results in these settings. However, these collections do not fairly represent in-the-wild challenges, where expressions are natural (spontaneous expressions), and problems like head pose variations and large displacements are frequent, as illustrated in Fig. 1. To answer these challenges, recently created collections [4–6] are mainly related to interaction situations where people are free of their movements. They are more challenging due to misalignment in faces, primarily caused by head motions, but also, spontaneous expressions.

State-of-the-art approaches that provide good results in near-frontal view have evolved in order to improve their robustness in the presence of head motions. The most commonly used solution to deal with head motions is to add a pre-processing step generally based on face registration in order to obtain frontal faces [8,9]. However, these methods casually induce texture changes that are not related to the underlying expression.

As in-the-wild settings, expressions are not acted, their intensity is getting smaller, and, hence, the changes induced by the registration interfere with changes induced by the expression itself. Indeed, spontaneous facial expressions are quite different from acted expressions in terms of facial movement amplitudes and/or texture changes. This makes them more difficult to characterize. In this context, systems based on dynamic textures may provide better performance [8,10]. Indeed, they detect subtle changes occurring on the face and do not require large changes in appearance, as texture-based or geometry-based approaches expect. However, these approaches are much more sensitive to varying head motion.

The question about the use and the impact of registration approaches arises especially when facial expression analysis is done in uncontrolled context. The use of registration approaches is increasing, despite a lack of evidence about their effectiveness due to the heterogeneity of the databases.

In this study, we address two challenges : head pose variations and large displacements in facial expressions recognition, denoted HPV and LD, respectively. In Section 2, we discuss the impact of HPV and LD on facial expressions recognition. In Section 3, representative frameworks of automatic facial expression analysis systems are introduced. Representative databases used for facial expressions recognition are reviewed in Section 4. A focus on these two challenges and the performances of

* Corresponding author.
E-mail addresses: benjamin.allaert@ed.univ-lille1.fr (B. Allaert), jose.mennesson@imt-lille-douai.fr (J. Mennesson), marius.bilasco@univ-lille1.fr (I.M. Bilasco), chabane.djeraba@univ-lille1.fr (C. Djeraba).

https://doi.org/10.1016/j.image.2017.11.002
Received 1 December 2016; Received in revised form 10 November 2017; Accepted 11 November 2017
Available online 21 November 2017
0923-5965/© 2017 Elsevier B.V. All rights reserved.

Fig. 1. Faces captured in-the-wild, from GENKI-4K database [7].

several approaches are compared. A common experimental framework using a newly created data collection covering simultaneously free (camera in front of the subject) and constrained (camera attached to the head) facial expressions is proposed in Section 5. A series of experiments are presented in Section 6, in order to quantify the performance degradation induced by HPV and LD considering representative state-of-the-art approaches. In Section 7, we summarize the limits of existing methods and data collections, as well as the benefits brought by the proposed experimental framework.

2. Large displacements (LD) and head pose variations (HPV)

In interaction situations, facial expression analysis has to deal with HPV and LD challenges. LDs involve translation, cinematic blur and scale changes, whereas, HPVs involve 3D-rotations (in-plane and out-of-plane). A first encountered issue with HPV is that most of the state-of-the-art approaches which give the best results in expression recognition are not invariant under 3D geometric transformations, thus computed features for the same face and the same expression vary depending on LD and HPV. For example, it is obvious that histogram-like [8,11] or dynamic texture features computed from equal-sized facial grids are not invariant under translations, rotations and scale changes. Fig. 2 shows an overview of a generic workflow often used in facial features extraction. Faces are divided into a regular grid of $m \times n$ local regions from which features can be extracted. Finally, features are concatenated into one-row vector which depicts the facial expression. HPV induces misalignment of the face (no correspondence of major facial components in each block, across the same facial image from a different point of view) and may results in mismatching between extracted features.

In order to obtain an invariance under geometric transformations, a pre-processing step which consists in registrating faces is proposed in [12,13]. Face registration aims to find the transformation (or the deformation) which reduces the discrepancies between two or more faces. These approaches modify facial characteristics (texture, geometry, motion) while reducing variations in translation, rotation and scale changes. However, registration induces artifacts which have a negative impact on the consistency of facial characteristics [14].

Another issue is encountered with LD which corresponds to important head motions between two frames. In the presence of LD, a blur effect appears on the face. This noise causes texture changes. Face registration suffers significantly under motion blur [15]. Indeed, most representative face registration approaches are built on features (i.e facial landmarks), and their robustness is heavily dependent on the image quality and resolution. Hence, the performances of the registration approaches may be less efficient when head motions occur. Fig. 3 shows an example of mis-estimation of facial landmarks due to the blur effect caused by LD, which deteriorates the face registration.

In brief, the presence of HPV and LD brings several challenges in the facial expression analysis:

- facial misalignment due to head pose variations
- preservation of initial facial expression during face registration process
- blur effect due to rapid movements resulting in poor landmark locations

In the next section, we discuss solutions to the challenges listed above.

3. Automatic facial expression analysis

Automatic facial expression analysis is a complex task as the face shape varies considerably from one individual to another. Furthermore, HPV and LD generate various face appearances for the same person. Such variations have to be addressed at different stages of an automatic facial expression analysis system. The generic facial expression analysis framework is illustrated in Fig. 4. First, the face is located in the frame and a registration step may be applied to remove the head motion and inter-subject differences. Next, the face is analyzed to estimate the remaining deformation caused by facial expressions. Then, features are extracted, and these features are used in the classification part of the system.

In the remainder of this section, we discuss the impact and the way HPV and LD are dealt with face registration and facial feature extraction processing stages.

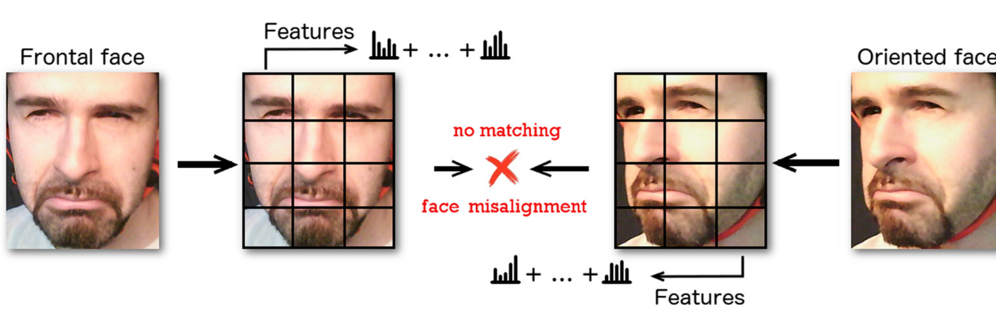

Fig. 2. Example of misalignment of the face in the presence of head pose variations.

Download English Version:

https://daneshyari.com/en/article/6941646

Download Persian Version:

https://daneshyari.com/article/6941646

Daneshyari.com