



**HAL**  
open science

## Impact of the face registration techniques on facial expressions recognition

B. Allaert, J. Mennesson, Ioan Marius Bilasco, C. Djeraba

► **To cite this version:**

B. Allaert, J. Mennesson, Ioan Marius Bilasco, C. Djeraba. Impact of the face registration techniques on facial expressions recognition. *Signal Processing: Image Communication*, 2018, 61, pp.44-53. 10.1016/j.image.2017.11.002 . hal-01644769v2

**HAL Id: hal-01644769**

**<https://hal.science/hal-01644769v2>**

Submitted on 6 Mar 2024

**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.

# Impact of the face registration techniques on facial expressions recognition

B.Allaert<sup>a,\*</sup>, J.Mennesson<sup>a</sup>, IM.Bilasco<sup>a</sup>, C.Djeraba<sup>a</sup>

<sup>a</sup>Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISTAL -  
Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France

---

## Abstract

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.

*Keywords:* Facial expressions, Head pose variations, Large displacements

---

## 1. Introduction

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

5 Several systems evaluate their performances on image collections, where facial expressions  
6 are played by actors, in order to obtain exaggerated facial deformations (acted expressions). Sev-  
7 eral approaches [1, 2, 3] obtain very good results in these settings. However, these collections do  
8 not fairly represent in-the-wild challenges, where expressions are natural (spontaneous expres-  
9 sions), and problems like head pose variations and large displacements are frequent, as illustrated  
10 in Figure 1. To answer these challenges, recently created collections [4, 5, 6] are mainly related to  
11 interaction situations where people are free of their movements. They are more challenging due  
12 to misalignment in faces, primarily caused by head motions, but also, spontaneous expressions.

---

\*Corresponding author

*Email addresses:* benjamin.allaert@ed.univ-lille1.fr (B.Allaert), jose.mennesson@univ-lille1.fr  
(J.Mennesson), marius.bilasco@univ-lille1.fr (IM.Bilasco), chabane.djeraba@univ-lille1.fr  
(C.Djeraba)

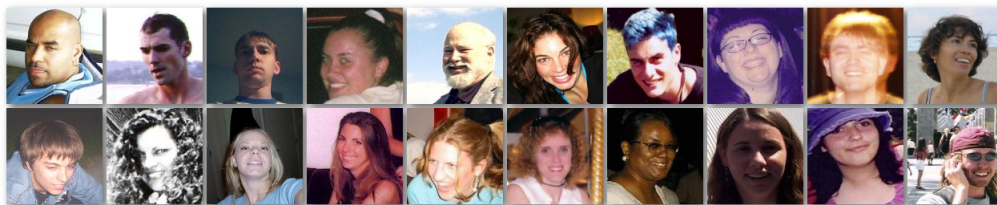


Figure 1: Faces captured in-the-wild, from GENKI-4K database [7].

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

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

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

30 In this study, we address two challenges : head pose variations and large displacements in  
 31 facial expressions recognition, denoted HPV and LD, respectively. In section 2, we discuss the  
 32 impact of HPV and LD on facial expressions recognition. In section 3, representative frameworks  
 33 of automatic facial expression analysis systems are introduced. Representative databases used for  
 34 facial expressions recognition are reviewed in section 4. A focus on these two challenges and the  
 35 performances of several approaches are compared. A common experimental framework using  
 36 a newly created data collection covering simultaneously free (camera in front of the subject)  
 37 and constrained (camera attached to the head) facial expressions is proposed in section 5. A  
 38 series of experiments are presented in section 6, in order to quantify the performance degradation  
 39 induced by HPV and LD considering representative state-of-the-art approaches. In section 7, we  
 40 summarize the limits of existing methods and data collections, as well as the benefits brought by  
 41 the proposed experimental framework.

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

43 In interaction situations, facial expression analysis has to deal with HPV and LD challenges.  
 44 LDs involve translation, cinematic blur and scale changes, whereas, HPVs involve 3D-rotations  
 45 (in-plane and out-of-plane). A first encountered issue with HPV is that most of the state-of-  
 46 the-art approaches which give the best results in expression recognition are not invariant under

47 3D geometric transformations, thus computed features for the same face and the same expres-  
 48 sion vary depending on LD and HPV. For example, it is obvious that histogram-like [8, 11] or  
 49 dynamic texture features computed from equal-sized facial grids are not invariant under transla-  
 50 tions, rotations and scale changes. Figure 2 shows an overview of a generic workflow often used  
 51 in facial features extraction. Faces are divided into a regular grid of  $m \times n$  local regions from  
 52 which features can be extracted. Finally, features are concatenated into one-row vector which  
 53 depicts the facial expression. HPV induces misalignment of the face (no correspondence of major  
 54 facial components in each block, across the same facial image from a different point of view)  
 55 and may results in mismatching between extracted features.

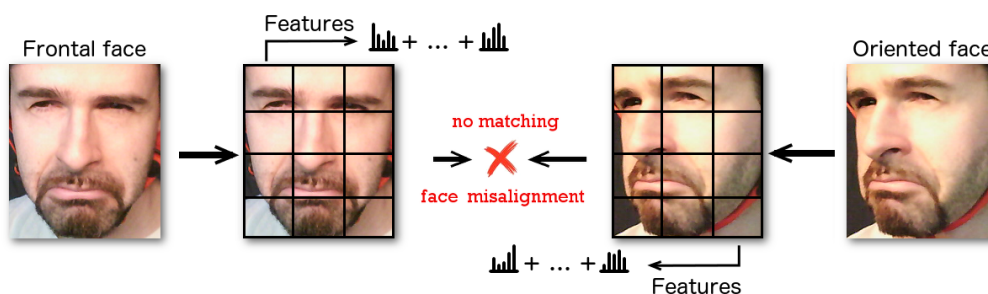


Figure 2: Example of misalignment of the face in the presence of head pose variations.

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

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

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

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

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



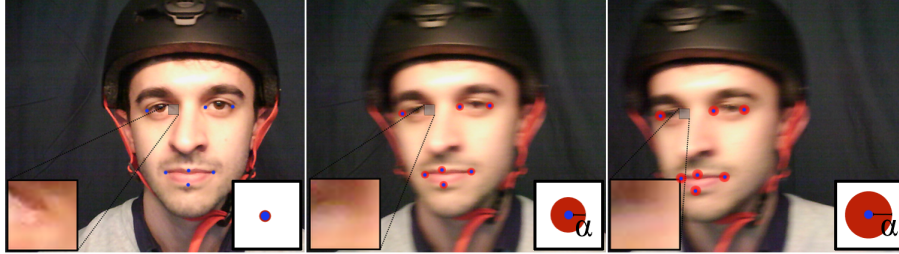


Figure 3: Poor estimation of landmarks location due to the blur effect caused by LD.  $\alpha$  corresponds to the level of uncertainty concerning the landmarks location, which decreases sharply with the quality of the picture.

### 76 3. Automatic facial expression analysis

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



Figure 4: Generic facial expression analysis framework.

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

#### 87 3.1. Face registration approaches

88 The face is usually detected and registered in order to establish the correspondence of major  
 89 facial components such as eyes, nose and mouth across different facial images. This aims at  
 90 guaranteeing invariance to geometric transformations. In the following we discuss the benefits  
 91 and limitations of various techniques such as eye-based registration (Eyes), as well as, more  
 92 evolved techniques such as shape-based registration (Shape) or 3D model-based registration (3D  
 93 Model).

94 *Eyes registration.* Eyes registration is the most popular strategy in near frontal-view databases  
 95 [2, 8, 9]. Eyes are detected and images are aligned and scaled with regard to the inter-pupilar  
 96 distance and orientation. Eyes are the most reliable facial component to be detected and suffer

97 little changes in the presence of expressions. The limitation of this approach is that eyes must  
 98 be well-detected. Usually, when out-of-plane rotations appear, the eyes quickly disappear and  
 99 additional deformations are induced, avoiding the detection of eyes.

100 *Shape registration.* Shape registration is based on 2D facial landmarks and aims at increasing ro-  
 101 bustness to HPV. Extensions considering more landmarks is supposed to provide greater stability  
 102 in case of individual poor landmark detections. Some approaches [1, 16] only rely on landmark  
 103 points located near the center of the face. The inner landmarks are mostly used to detect the  
 104 face and estimate the head pose. However, these points are affected by facial deformations in the  
 105 presence of facial expressions. Other approaches also take into account the contour of the face  
 106 in order to exploit the information related to the geometry of the face [17]. The outer landmarks  
 107 are less affected by facial deformations due to facial expressions, but they are difficult to locate  
 108 in case of out-of-plane rotations. We can say that most of 2D-feature-based methods are suitable  
 109 for the analysis of near frontal facial expressions in the presence of limited head motions. But,  
 110 they do not cope well with difficulties brought to occlusions and out-of-plane rotation. Indeed,  
 111 an image acquisition system provides only the projection of the observed scenes in a 2D plane.  
 112 The projection only captures information available in front of the camera and loses out-of-plane  
 113 information. Figure 5 illustrates a poor estimation of facial landmarks due to a yaw out-of-plane  
 rotation, where the left part of the face disappears progressively as the face rotates.

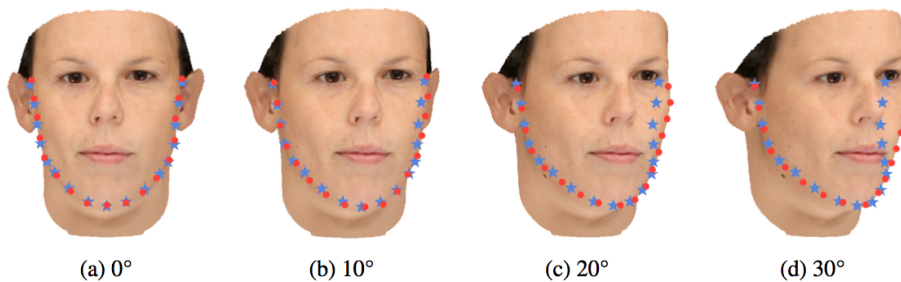


Figure 5: Similarity errors of 2D (red) and 3D (blue) facial contour landmarks under different angles [18].

114

115 *3D model registration.* Recent approaches propose robust face registration based on 3D to gener-  
 116 ate a natural face image in frontal pose. Compared to 2D approaches, 3D approaches reduce  
 117 the deformation of the face when facial expressions occur. Among these approaches, Zhu et al.  
 118 [19] propose a robust face registration approach based on a 3D Morphable Model (3DMM). To  
 119 build a 3D face model from a 2D face image, they estimate the depth of the external face region  
 120 and the background.

121 **Pose registration** Landmarks are detected using facial alignment techniques from the 2D  
 122 face. The authors apply landmark marching in order to solve the issue illustrated in Figure 5.  
 123 Corrected landmarks on the boundary of the face are used as facial anchors. Facial anchors  
 124 correspond to specific facial points that are used in order to align the 2D face on a 3D morphable  
 125 model (constructed from large training data). A fitted 3D face is then generated and 3DMM  
 126 coherently registers the face in front of the camera and preserves the appearance and the shape

127 of the face. However, in case of high HPV, some regions can be hidden due to self-occlusions,  
 128 as illustrated in Figure 6.

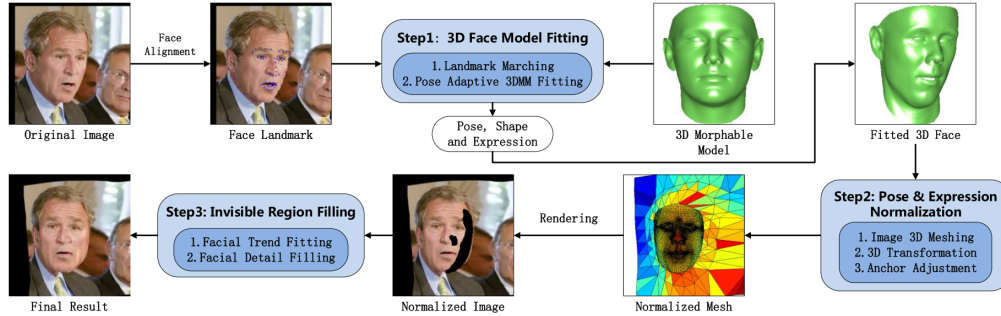


Figure 6: Face registration approach based on a 3D Morphable Model (3DMM), extracted from [19].

129 **Filling of occluded regions** Bad filling of the occluded region leads to large artifacts af-  
 130 ter registration and deteriorates recognition performance. To deal with self-occlusions, several  
 131 approaches use the facial symmetry or in-painting approaches [20]. The quality of these ap-  
 132 proaches depends on the size of the occluded face region and they are often not well-suited for  
 133 use in unconstrained conditions (e.g illumination variation, occluding objects such as glasses and  
 134 hair). Recent 3D approaches (such as [21]) use measurements over multiple frames to refine the  
 135 rigid 3D shape and estimate hidden facial parts (assuming that the hidden facial part was visible  
 136 in the previous frames).

137 In the next section, we present the most significant facial feature extraction approaches that  
 138 have been proposed in the literature.

### 139 3.2. Facial Feature extraction approaches

140 In the literature, there are two types of methods used to analyze facial expressions : methods  
 141 based on facial static characteristics and methods based on dynamic characteristics, each of them,  
 142 applied locally or globally on the face.

143 *Static texture features.* Most of facial expression recognition approaches are based on the tex-  
 144 tural information [2, 8, 13, 22]. One of the most popular method is the Local Binary Pattern  
 145 (LBP) [23]. For every pixel of the image, its gray-scale value is compared with those of the eight  
 146 surrounding pixels. The value of each neighbor is set to 0 if its gray-scale value is smaller than  
 147 the value of the central pixel and to 1 otherwise. To reduce the dimensionality of the problem  
 148 further, images are usually divided into a regular grid of  $m \times n$  local regions from which LBP  
 149 histograms can be extracted. Then, they are concatenated into a single histogram.

150 *Dynamic features.* The dynamic characterization of the facial texture can be achieved either by  
 151 considering the changes in terms of temporal texture characteristics (extending static character-  
 152 istics to temporal domain) or evaluating the changes in terms of perceived motion (by means of  
 153 dense optical flow fields):

154 **Region-based temporal texture features** Dynamic texture is an extension of texture char-  
155 acterization to the temporal domain. Ambadar et al. [24] prove the importance of the dynamic  
156 texture for facial expression recognition as it allows a better analysis of physical deformation of  
157 face. Zhao et al. [8] propose an extension of the original LBP operator to the spatio-temporal  
158 domain called Volume Local Binary Patterns (VLBP). VLBP considers a block of video frames  
159 as a single 3 dimensional array of grayscale values. A simplified, more practical version of the  
160 approach was proposed by its creators to make it more attractive for further usage called Local  
161 Binary Patterns from Three Orthogonal Planes (LBP-TOP) [8]. LBP-TOP applies LBP on every  
162  $xy$ ,  $x_t$ , and  $y_t$  slices separately. Then it averages the histograms over all slices in a single plane  
163 orientation, and concatenates the resulting histograms of the three dimensions. With LBP-TOP it  
164 is possible to combine motion and appearance analysis in one operator : the features histogram.

165 **Optical flow features** Optical flow measures the relative motion between two successive  
166 images in a sequence. It is used to analyze facial expression [3, 25] and obtains good perfor-  
167 mances. Optical flow are dense and features encoding their local or global characteristics are  
168 extracted in order to exploit the encoded motion information. For instance, Histogram of Ori-  
169 ented Optical Flows (HOOF) feature [3] is successfully used in order to encode the distribution  
170 of optical flows and extract global movement characteristics. HOOF feature encodes the dense  
171 optical flow fields by cumulating directions binned with regard to the horizontal axis and by  
172 weighting their magnitude. The weighting step aims at minimizing the noise impact on the  
173 global feature. However, high HPV involves an important loss in terms of facial information and  
174 it reduces the recognition rate of facial expression algorithms. Indeed, occluded face areas of the  
175 current picture are defined by a set of pixels who disappear in the next picture when out-of-plane  
176 rotations occur. These pixels have no correspondence within the next picture. This results in  
177 motion that is not directly observable in these regions. Recent approaches use the boundaries of  
178 the face (which have a high probability of being occluded in case of HPV) in order to reduce the  
179 noise induced by motion discontinuities [26]. Therefore, they use fill-in methods based on the  
180 motion of the neighboring regions and the physical constraints of the face (wrinkles, shape, ...)  
181 [27].

182 Although dynamic textures approaches perform well in near frontal view, facial expression  
183 recognition based on dynamic textures, when HPV and LD occur, is still a challenging problem.  
184 Indeed, in these context stationary dynamic textures must be well-segmented in space and time.  
185 The performances of these approaches depend heavily on the quality of the face registration  
186 approaches to reduce facial deformations.

187 In the next section, we analyze how HPV and LD challenges are highlighted in several  
188 databases commonly used to validate facial expression analysis approaches.

#### 189 4. Facial expression databases

190 Most of facial expression systems evaluate their performances in controlled settings [28, 29],  
191 where the face pose is static. In these settings, expressions are exaggerated and often played  
192 by actors in order to induce important deformations on the face. In contrast, some data collec-  
193 tions are recorded in more natural interaction contexts, where the subject has full freedom of  
194 its movement and facial expressions are spontaneous [30, 5, 6]. These databases, which pro-  
195 pose more natural interactions, yield more often problems related to high HPV and LD. Some  
196 acted databases [28, 29] have extended their data collections in order to offer a more challenging  
197 context for approaches aiming to improve their robustness to in-the-wild conditions.

198 The most commonly used databases for facial expressions analysis are shown in Figure 7.  
 199 Figure 7 shows that the complexity of the different databases increases depending on the type  
 200 of expression (acted to spontaneous). Concerning the presence of LD and HPV, an indicator of  
 201 intensity between one and three stars (\*) depicts the ratio of data which contains LD, or HPV.

	CK+	MMI	GEMEP	DISFA	SEMAINE	RECOLA
Details	CK+ [28]	MMI [29]	GEMEP [4]	DISFA [31]	SEMAINE [5]	RECOLA [6]
Expressions	Acted	Acted	Acted	Spont.	Spont.	Spont.
Presence of HPV	-	*	**	*	**	***
Presence of LD	*	*	***	**	***	***

Figure 7: Commonly used databases for facial expression analysis.

202 Evolving challenges proposed in these databases reflect that the facial expression analysis  
 203 in an interaction situation is a complex issue. In these contexts, the presence of HPV and LD  
 204 challenges the current approaches.

205 Features previously discussed have been applied to several databases and the results are re-  
 206 ported in Figure 8. Results on CK+ [28] and MMI [29] show that facial expression analysis  
 207 achieve excellent performances under controlled settings where the pose is static and expres-  
 208 sions are acted. As illustrated in Figure 8, registration approaches based on 3D models and 2D  
 209 face shape are intensively used to analyze facial expressions in uncontrolled contexts. Despite the  
 210 fact that these approaches provide better performances, recognition rates are still very low with  
 211 regard to performances observed under controlled settings. However, in more natural interaction  
 212 contexts like in GEMEP [4], DISFA [31] and SEMAINE [5], a significant drop in performance  
 213 can be observed. To better visualize the performances of each registration approach on the vari-  
 214 ous databases, we included in the lower part of Figure 8 two related graphics. A comparison of  
 215 the approaches using average recognition or classification rates are given in Figure 8-A. Figure  
 216 8-B shows the performances of the approaches estimated using Person’s cross correlation. Each  
 217 color is associated with a pair of a registration approach and a database.

218 Recent databases, like SEMAINE [5] or RECOLA [6], include free head movements. Still,  
 219 it is difficult to study the impact of head movements on facial expression recognition as many  
 220 other parameters are changing within or between the existing databases. Hence, it is difficult to  
 221 quantify the impact of issues related to LD and HPV, as well as, the registration techniques on  
 222 the recognition performances. Basically, we are missing the equivalent near-frontal view data in  
 223 order to measure effectively the induced deformations while correcting LD and HPV effects.

## 224 5. Synchronous acquisition system

225 In order to quantify the impact of free head movements on expression recognition perfor-  
 226 mances, we propose an innovative acquisition system that collects data simultaneously in pres-  
 227 ence and absence of head movement. Experiments are then conducted in order to estimate the  
 228 impact of HPV and LD on the recognition process.

References	Features	Registration	Databases		Performances
Zhao <i>et al.</i> [8]	LBP-TOP	Eyes	CK+	■	ar:95.2%
Happy <i>et al.</i> [2]	Salient LBP Patches	Eyes	CK+	■	ar:94.14%
Allaert <i>et al.</i> [25]	Optical flow & Geometry	Eyes	CK+	■	cr:95.34%
Koelstra <i>et al.</i> [3]	FFDs	Eyes	MMI	■	cr:94.3%
Jiang <i>et al.</i> [1]	LPQ-TOP	Shape	MMI	■	cr:94.7%
Rivera <i>et al.</i> [16]	DNG	Shape	MMI	■	ar:97.6%
Jiang <i>et al.</i> [22]	LPQ	N/A	GEMEP	■	cr:66%
Yang <i>et al.</i> [11]	LBP,LPQ	Shape	GEMEP	■	ar:84%
Sandbach <i>et al.</i> [13]	LBP	Eyes	DISFA	■	cc:0.342
Cruz <i>et al.</i> [32]	LPQ	Shape	SEMAINE	■	ar:55%
Nicolle <i>et al.</i> [33]	Appearance & Geometry	Shape	SEMAINE	■	cc:0.46
Chen <i>et al.</i> [34]	3D Facial Shape	3D Model	SEMAINE	■	cc:0.51

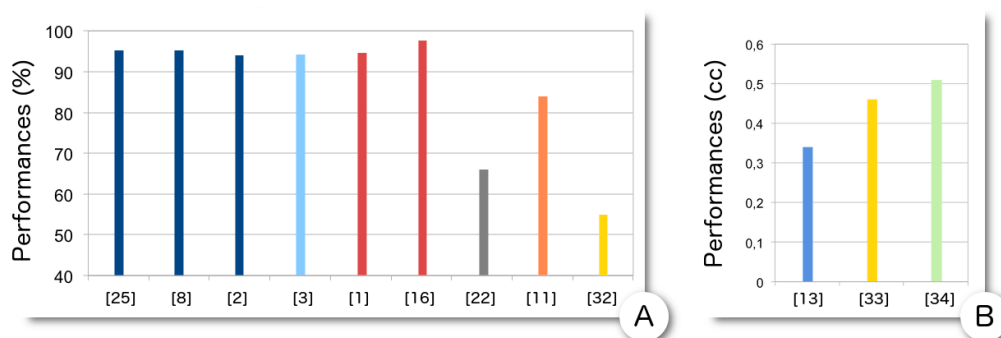


Figure 8: Recent methods to facial expression analysis in the literature. (cc: Person's cross correlation, ar: average recognition rate, cr: classification rate).

229 To address this issue, we propose a new acquisition system called : Simultaneous Natural  
230 and Posed Facial expression (SNaP-2DFe) allowing the study of the HPV and LD impact on  
231 expression recognition methods.

### 232 5.1. Acquisition system

233 Each facial expression is recorded simultaneously using a two-camera system : one camera  
234 is fixed on a helmet, while the other is placed in front of the user at near-range distance. The  
235 helmet camera provides data similar to CK+ [28] and MMI [29] databases, where little or no head  
236 movements occur. The frontal camera provides data similar to RECOLA [6] and SEMAINE [5]  
237 databases, as subjects are freely moving their head. Our database enhances measuring the impact  
238 of head-movements relying on the information returned by the frontal camera, compared to the  
239 helmet camera.

240 The helmet is equipped with eight LEDs, which ensure homogeneous illumination on the  
241 face, even when the head is moving. It also includes a "9DOF Razor IMU" board by SparkFun,  
242 which contains a 3-axis gyroscope/accelerometer/magnetometer and a micro-controller perform-  
243 ing sensor fusion. Finally, it includes a camera located fifty centimeters in front of the face and  
244 maintained by an aluminum rail in order to ensure global stability.



245 We use a counterweight that enhances the user’s comfort and guarantees that the helmet does  
 246 not shift position while the user moves. It is important to guarantee that the helmet is stable in  
 247 order not to disrupt the user experience during the recording session. We verify that the helmet is  
 248 stable by computing the mean difference (in pixels) of facial landmark locations from the helmet  
 249 camera under neutral expression between different head poses. We have obtained very similar  
 250 values regardless of the head movement. When the head is not moving we have obtained an error  
 251 of 1.74 pixels. When the head is executing a diagonal movement we have obtained 1.87 pixels.  
 252 In case of a Pitch, a Yaw or a Roll movement we have obtained respectively errors of 1.77 pixels,  
 253 1.77 pixels and 1.95 pixels. When the head is executing a translation an error of 1.71 pixels has  
 254 been reported. With regard to the values, in our understanding, errors stem primarily from the  
 255 instability of landmarks location detection and not the instability of the helmet.

256 The capturing system is illustrated in Figure 9. Each participant was instructed to wear  
 257 a helmet fitted with a camera (Camera 1) and to sit in front of a projection screen at about one  
 258 meter away from the fixed camera (Camera 2). We recorded images using two Creative Live cam  
 259 inPerson HD (Full HD 1080p at a frame rate of 30 fps) and with an uniform background. The  
 260 capturing system is illustrated in the left part of Figure 9. The right part of Figure 9 represents  
 261 image samples of facial expressions where the subject performs a pitch movement. The first two  
 262 lines correspond to selected synchronous frames in time. The first line corresponds to the helmet  
 263 camera (Camera 1) and the second line, to the fixed camera (Camera 2). The curve in the bottom  
 264 right part of the Figure 9 represents the yaw, pitch and roll values obtained by the gyroscope  
 265 during the session.

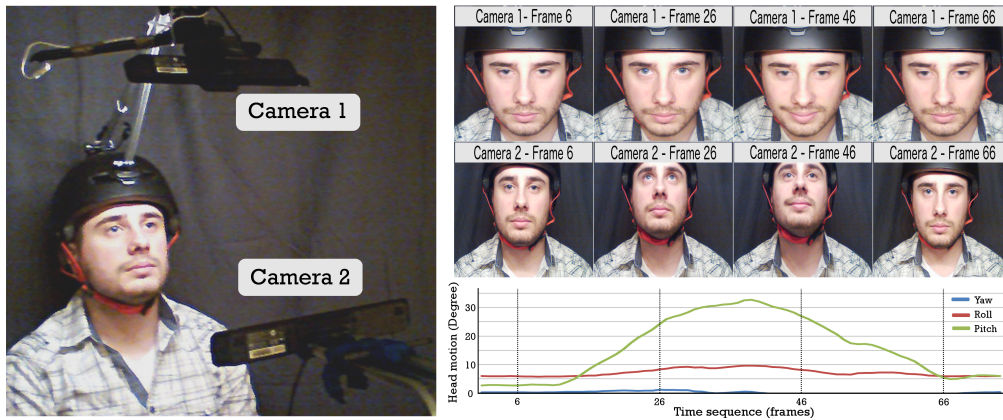


Figure 9: SNaP-2DFe system setup and example images of facial expressions recorded during a pitch movement.

266 *5.2. Data acquired*

267 Our preliminary database includes 840 samples collected from 10 subjects. Each video cor-  
 268 responds to a combination of one facial expression and a sequence of uniform head movements.  
 269 In each sequence, the user follows a specific pattern of movement that corresponds to one of the  
 270 following animations : one translation on  $x$  ( $T_x$  - corresponding to LD motions), combined with  
 271 three rotations (roll, yaw, pitch - corresponding to HPV motions).

272 In static sequences, the user does not move the head, in order to collect data for head-posed  
 273 facial expression analysis. In diagonal sequence, the user moves the head while combining

274 translations and rotations. Movement ranges are large: translations go up to more than 150mm  
275 in any axis from the starting point, and rotations can reach 40 degrees.

276 For each subject, we have six animations combined with seven expressions (Neutral, Hap-  
277 piness, Sadness, Anger, Fear, Surprise, Disgust) from two cameras, which makes a total of 84  
278 videos for each subject (42 with HPV and 42 without HPV).

279 The subjects were instructed to express emotions. Yet, as the subjects were not actors, in  
280 some situations, expressions recorded show spontaneous expressions characteristics: low inten-  
281 sity, limited facial deformations, various ways of expressing a given expression. In some other  
282 situations, the subjects were acting a different expression than the one they were asked for. How-  
283 ever, this is not an issue at this point, as our main concern is to evaluate classification results  
284 comparing the near frontal-view data and the HPV-LD data of the same underlying expression.

285 The data collected is freely available for research purposes and can be downloaded on de-  
286 mand from <http://www.cristal.univ-lille.fr/FOX/>.

287 In order to assess the impact of face registration approaches on the recognition of facial ex-  
288 pressions, the next section discusses the results of different registration methods on the collected  
289 SNaP-2DFe database.

## 290 6. Experimentation

291 Several experiments are conducted in the following. Firstly, we measure the ability of the  
292 registration techniques to simulate frontal pose images (like the ones produced by the helmet  
293 camera) from the static camera. As a reminder, the helmet-camera is a fixed frontal camera,  
294 where no head motion appears except the facial expressions. The characteristics of the face are  
295 stable during the sequence. This means that no registration step is necessary. In the experiments  
296 presented in Section 6.1, we evaluate the capabilities of face registration approaches to reduce  
297 the discrepancies between faces from frontal and non-frontal settings.

298 Secondly, we study the ability of registration techniques to preserve the original facial defor-  
299 mations produced by the underlying expression. In Section 6.2, expression recognition classifiers  
300 are trained from the helmet-camera images and we measure the ability of registration techniques  
301 to bring non-frontal images in frontal settings with regard to the frontal classifiers. In Section 6.3,  
302 we conduct a series of experiments, where the classifiers are trained from the registered images,  
303 in order to evaluate if the deformations, induced by the registration, preserve distinctive features  
304 for expression classification. Finally, we conduct an experiment in order to evaluate the impact  
305 of the registration deformation with regard to specific expressions in Section 6.4.

### 306 6.1. Evaluation of registration quality

307 In order to clearly illustrate the quality of the registration process we provide a qualitative  
308 and a quantitative evaluation. The qualitative evaluation illustrates visually the deformation in-  
309 duced by the registration process, whereas the quantitative evaluation measures the geometric  
310 and structural similarity between the registered face (from the fixed camera) and the near-frontal  
311 view face (from the helmet camera).

#### 312 6.1.1. Qualitative evaluation

313 Figure 10 shows a qualitative comparison of three face registration techniques on different  
314 head poses extracted from SNaP-2DFe database (e.g. frontal pose, translation on x (Tx), roll,  
315 pitch, yaw and diagonal). To deal with near frontal face, Eyes registration is more adapted



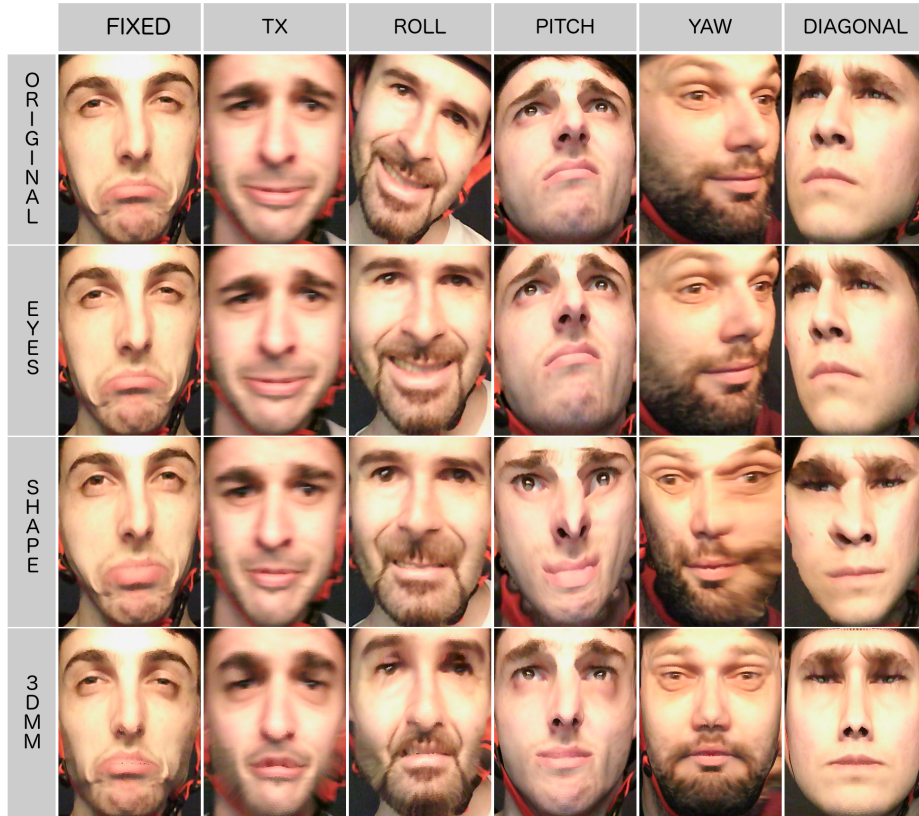


Figure 10: Example of different facial registration approaches on each animation.

316 because this registration will not cause facial deformations and the locations of feature points are  
 317 rather stable. However, severe out-of-plane rotation downgrade the precision of feature points  
 318 localization process. This causes strong face deformations. In this case, recent approaches based  
 319 on 3D face model seem better suited than other approaches. 3DMM method illustrated here is  
 320 based on facial symmetry reconstruction [19]. Thanks to the reconstruction of occluded face  
 321 regions, this approach allows rebuilding faces in the presence of out-of-plane rotations.

### 322 6.1.2. Quantitative evaluation

323 *Experimental Setup.* In order to evaluate the quality of face registration process, we use the  
 324 structural similarity index method (SSIM) [32]. SSIM compares local patterns of pixel intensi-  
 325 ties that have been normalized for luminance and contrast. The face geometry delivers good  
 326 information for some facial expressions, but fails in detecting subtle motions, that can be de-  
 327 tected only by observing skin surface changes. From our point of view, SSIM is appropriate to  
 328 measure the errors of the face registration for facial expression recognition. SSIM formula is  
 329 based on three comparison measurements : luminance ( $l$ ), contrast ( $c$ ) and structure ( $s$ ). The  
 330 measure between two local regions  $x$  and  $y$  of common size  $N \times N$  is :

$$SSIM(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_x\sigma_y + c_2)(2cov_{xy} + c_3)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)(\sigma_x + \sigma_y + c_3)}. \quad (1)$$

where  $\mu_x$  and  $\mu_y$  are the average of  $x$  and  $y$ ,  $\sigma_x^2$  and  $\sigma_y^2$  are the variance of  $x$  and  $y$ , and  $cov_{xy}$  is the covariance of  $x$  and  $y$ . Three variables  $c_1 = (k_1L)^2$ ,  $c_2 = (k_2L)^2$ ,  $c_3 = c_2/2$  stabilize the division with weak denominator, where  $L$  corresponds to the dynamic range of the pixel-values and  $k_1 = 0.01$  and  $k_2 = 0.03$ .

SSIM is applied on each animation of SNaP-2DFe, after using different registration approaches, based on : Affine Transformation on eyes location (Eyes), facial shape deformation using moving least squares [33] (Shape) and 3D Morphable Model [19] (3DMM). Results are reported in Table 1.

*Result Analysis.* 3DMM registration approach gives the best overall results, with mean similarity of 61.33% over all animations. Eyes and Shape registrations are very similar. Eyes registration approaches suit better in-plane geometric transformation (fixed, translation and roll) than Shape registration. Both present limitations due to the fact that they only exploit the visible 2D information, whereas 3DMM registration achieves better results in out-of-plane conditions (pitch, yaw, diagonal). In all cases, registration approaches improve the SSIM metrics in challenging conditions even though it does not guarantee a perfect match between the two cameras.

Registration	Fixed	Tx	Roll	Pitch	Yaw	Diagonal	Mean
None	53.30	47.88	48.68	46.34	46.54	51.14	48.98
Eyes	58.26	55.06	55.29	54.57	52.01	55.79	55.16
Shape	55.25	53.49	52.29	57.40	54.44	57.67	55.09
3DMM	<b>64.72</b>	<b>61.81</b>	<b>58.17</b>	<b>60.52</b>	<b>58.47</b>	<b>64.29</b>	<b>61.33</b>

Table 1: SSIM (in percentage) applied on each animation, with different face registration approaches.

Considering the results, face registration approaches may not ensure a perfect similarity between faces. But still, we expect that they encode expression-related artifacts that might still differentiate between expressions.

In the next section, we study the impact of the facial deformation induced by face registration approaches to facial expression recognition.

## 6.2. Evaluation of registration impact on expression recognition

We provide experimental results about the impact of facial registration on expression recognition performance when free head movements occur.

In this context, we try to measure the capacity of the registration method to induce facial deformation that can cope with classifiers learnt from the near-frontal recordings provided by the helmet camera.

*Experimental Setup.* The next series of experiments are conducted using LIBSVM [34] with the Radial Basis Function kernel for classification. Each expression is classified into one of the seven classes : Neutral, Anger, Fear, Disgust, Happiness, Sadness, and Surprise. Assuming the face region is well aligned after applying different face registration approaches, we use a 5\*5 facial

361 block based approach to extract the features. We consider commonly used static (LBP) and  
 362 dynamic (LBP-TOP, HOOF) texture features for classifying expressions. The implementation  
 363 of LBP and LBP-TOP were taken from [8]. HOOF feature extraction was reimplemented by  
 364 us considering the algorithm described in [35]. Expression recognition rates are computed by  
 365 employing the 10-fold cross-validation protocol.

366 In the first experiment, near-frontal faces recorded with the helmet camera are used for the  
 367 training step. This first experiment allows to evaluate the performance of the different approaches  
 368 in good conditions (not involving HPV or LD). The characteristics of the face are stable during  
 369 the sequence, hence no registration step is necessary. All registration approaches were applied  
 370 on each animation. The expression recognition classifier is trained using images captured in  
 371 near-frontal settings using the helmet camera.

372 *Result Analysis.* The results for the different configurations : fixed versus helmet camera, no  
 373 registration versus various registration approaches on the fixed camera are given in Table 2. A  
 374 first look at the results obtained in Table 2 shows that the originating camera and face registration  
 375 approaches have significant impact on the performances. In the following we discuss the impact  
 376 of the features, the registration method as well as the originating camera.

	Helmet camera	Fixed camera			
Method	Original data	Original data	Eyes	Shape [33]	3DMM [19]
LBP	75.52	<b>30.55</b>	47.46	<b>47.76</b>	<b>51.34</b>
LBP-TOP	78.34	19.44	49.12	44.62	46.93
HOOF	<b>83.21</b>	17.38	<b>50.01</b>	42.16	48.73

Table 2: Facial expression recognition rates while using different face registration approaches.

377 **Impact of the originating camera** Results reported in first column of Table 2 show that  
 378 the state-of-the-art methods are suitable for the analysis of facial expressions when the head is  
 379 not moving. However, in the presence of HPV and LD, images provided by the fixed camera are  
 380 poorly classified as shown in the second column (Fixed Camera - Original Data).

381 **Impact of the features** The results obtained with the helmet camera show that dynamic  
 382 texture features such as LBP-TOP or HOOF are more efficient than LBP. The HOOF approach  
 383 obtains better performances than LBP-TOP where little or no head movement appears and proves  
 384 that optical flow approaches are better suited to the facial expression analysis. However, the  
 385 experiment shows a drastic fall in performances on the original data from the fixed camera. In  
 386 this context, recognition rates measured with dynamic texture features have suffered more than  
 387 others. Overall, these methods are much more sensitive to the presence of HPV and LD. It is  
 388 important, therefore, to ensure that the face is aligned in order to maintain the benefits brought  
 389 by the dynamic texture features.

390 **Impact of the registration method** When considering results obtained using various reg-  
 391 istration methods on the images captured with the fixed camera (last three columns), it can be  
 392 easily seen that the performances are very similar. Each column corresponds to the expression

393 classification rate obtained after applying a different registration approach, considering respec-  
 394 tively : affine transformation on eyes location (Eyes), facial shape deformation using moving  
 395 least squares [33] (Shape) and 3D Morphable Model [19] (3DMM).

396 The use of registration techniques improves significantly the performances of facial expres-  
 397 sion analysis when free head movements occur. Considering the results in Table 2, the Eyes  
 398 registration seems to be the most successful strategy in terms of sustainability and effectiveness  
 399 with regard to the needs of facial expression classification method. Despite the gain obtained  
 400 from both Shape and 3DMM registration approaches when using LBP, these registration tech-  
 401 niques appear to be less suited with dynamic texture features.

### 402 6.3. Evaluation of preserving distinctive facial expression deformations

403 In the following, we evaluate the impact of specific head motion patterns on the expression  
 404 recognition rates when using LBP features. Implicitly, we quantify of capacity of registration  
 405 techniques to preserve distinctive facial deformations in case of various movements typologies.

406 *Experimental setup.* The selected registration techniques have been applied and compared on  
 407 each class of animations. We evaluate the impact of face registration approaches on facial ex-  
 408 pressions recognition rates and we identify the strengths and weaknesses of each. The results  
 409 are given in Table 3. The training was performed on the registered images captured by the fixed  
 410 camera. Hence, the trained classifier took into account the deformation induced by the registra-  
 411 tion.

Registration	Fixed	Tx	Roll	Pitch	Yaw	Diagonal
Original data	45.23	38.09	32.47	37.71	33.33	14.26
Eyes	<b>52.38</b>	33.33	<b>47.61</b>	30.95	40.47	26.19
Shape	47.61	35.71	42.85	30.95	38.02	11.90
3DMM	48.02	<b>39.27</b>	40.21	<b>40.74</b>	<b>41.08</b>	<b>34.96</b>

Table 3: Recognition rates of facial expression classification using LBP features, after face registration step.

412 *Results analysis.* The results in bold in Table 3 show the best results per registration approach  
 413 obtained for specific movement patterns. While the head does not perform out-of-plane rotations,  
 414 as in Fixed and Roll settings, Eyes registration provides the best results. Whereas, in case of out-  
 415 of-plane rotations the 3DMM registration performs better.

416 The Eyes registration is the most suitable in frontal (fixed) settings. Indeed, in near-frontal  
 417 view condition, a simple in-plane rotation aligns the face. This solution preserves the geometry  
 418 of the face. However, this method is not working well in case of LD or HPV. Thanks to the  
 419 reconstruction of the occluded face region, 3DMM approach obtains the best results when out-of-  
 420 plane rotations occur. However, the reconstruction system is based on a face mirroring technique,  
 421 which sometimes has negative impact on the induced facial expression.

### 422 6.4. Evaluation of per-expression registration impact

423 We have conducted a complementary study about measuring the impact of the registration  
 424 techniques with regard to the underlying expression.

425 *Experimental setup.* In the light of the previous results (see Table 2), we have selected 3DMM  
 426 and Shape registrations, as well as, LBP as texture features for studying the recognition rate  
 427 variations when considering various expression classes. We have constructed ROC curves using  
 428 10-fold cross-validation protocol considering the whole dataset, as well as, independent Neutral,  
 429 Anger, Surprise, Happiness, Sadness, Fear and Disgust partitions. Training was conducted on  
 430 the whole dataset, as well as on each expression-related partition resulting in eight different  
 431 classifiers.

432 *Results analysis.* Figure 11 shows ROC curves corresponding to each expression score calcu-  
 433 lated with LBP from the helmet camera (blue) and the fixed camera after different registrations  
 434 (red : 3DMM and green : Shape). With regard to the Mean Curve, faces obtained by registra-  
 435 tion show lower performances than the faces acquired by the helmet camera (see Mean Curve in  
 436 Figure 11).

437 Results show that some expressions suffered severely from the registration process. Expres-  
 438 sions like Anger, Surprise are less impacted by facial registration. This is probably due to the  
 439 fact that face registration process induces less facial deformations around regions (such as eye-  
 440 brows) used in the recognition process. Disgust and Fear expressions show similar behavior as  
 441 Anger and Surprise, but the 3DMM registration technique seems more robust. Expressions like  
 442 Happiness, Neutral and Sadness seem more impacted by the registration as regions outside the  
 443 landmarks are affected (such as upper cheeks for Happiness). The drop in performances in the  
 444 case of Neutral expressions underlines the fact that the deformations produced by the registration  
 445 induce "false" facial expression recognition.

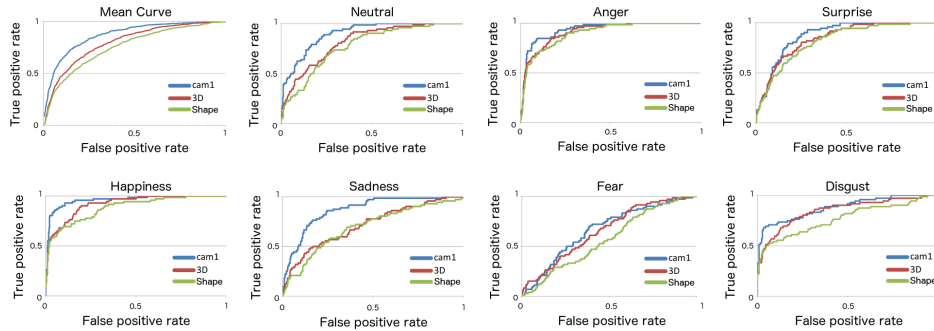


Figure 11: ROC curves for the 10-fold cross-validation protocol for LBP.

446 In the light of results from these experiments, the facial registration approaches improve the  
 447 facial expression recognition in case of HPV and LD. Yet, a lot of improvements still have to  
 448 be made in order to obtain comparable performances in the two settings : near-frontal views vs  
 449 unconstrained head movements. The choice of face registration approach is heavily dependent  
 450 on the type of head motion variation occurring in the video sequence. Mainly used for face  
 451 recognition, these registration approaches do not ensure a convenient alignment persevering the  
 452 expression of the face over time.

453 In the following section, we summarize the contributions of this study. From literature re-  
 454 sults (see section 3), and results obtained in our experimental settings we discuss perspectives  
 455 concerning LD and HPV.

## 456 7. Conclusion and future works

457 In this paper we have addressed the study of the impact of registration techniques on expres-  
458 sion recognition performances. Registration techniques are employed in order to handle HPV  
459 and LD for facial expression recognition. When analyzing the facial characteristics (texture, ge-  
460 ometry, motion) for expression recognition, facial distortions due to misalignment degrade the  
461 performances of the system. Removing distortions is a complex task. Most of the time it has a  
462 negative impact on the coherency of facial characteristics (texture, geometry, motion) [14].

463 3D Model registration is constantly improving with regard to in-the-wild challenges, but  
464 there is still no solution to ensure a satisfactory face registration while maintaining the facial  
465 expression. Indeed, the use of face registration techniques does not seem adequate to preserve  
466 the features encoding facial expression deformations. The loss in terms of precision when con-  
467 sidering free head movements is partly due to the noise induced by the face registration process  
468 itself.

469 In this paper, we propose an innovative acquisition system, in order to quantify the impact  
470 of free head movements on expression recognition performances. Experiments on the impact  
471 of well known head pose registration techniques (Eyes, Shape or 3DMM) on facial expression  
472 recognition are reported. The results show that the face registration and the facial expression  
473 recognition approaches are heavily dependent on the type of head motion variation. When con-  
474 sidering static approaches (such as LBP), in the presence of in-plane rotations, registration tech-  
475 niques based on landmarks (such as Eyes or Shape) preserves better the underlying expression.  
476 However, when out-of-plane rotations occur, registration techniques based on the reconstruction  
477 of 3D models seem more accurate as they preserve the underlying expressions better. Approaches  
478 using dynamic features (such as LBP-TOP, HOOF) are more efficient in terms of facial expres-  
479 sion analysis for frontal poses. However, these approaches do not handle well face registration  
480 techniques (Eyes, Shape or 3DMM).

481 Out-of-plane rotations affect in a strong manner, the expressions recognition process. Sup-  
482 porting out-of-plane rotations can be achieved either by incrementing data (as in Deep-learning  
483 methods) or by registering the face representation to near frontal views. Although the first ap-  
484 proach seems more popular at the present time, we truly believe that progress can be made in  
485 the latter by creating innovative face registration techniques that preserve facial expression. The  
486 SNaP-2DFe database can jointly be used to propose and evaluate innovative registration tech-  
487 niques while reinforcing the facial expression recognition or the head pose estimation methods.

488 As an alternative to registration techniques, we think that solutions from the field of dense  
489 optical flow should be explored. The enhancements of post-filtering solutions capable of regis-  
490 tering the movement, by filtering out the head movement and keeping only the inner-facial local  
491 movement could be done. The database proposed here may serve future works in this direction.  
492 The helmet camera provides the ground truth movements while the static camera provides the  
493 challenging data from where the head movements should be subtracted by future post-filtering  
494 solutions.

## 495 8. References

### 496 References

- 497 [1] B. Jiang, M. Valstar, B. Martinez, M. Pantic, A dynamic appearance descriptor approach to facial actions temporal  
498 modeling, *IEEE Transactions on Cybernetics* 44 (2) (2014) 161–174.

- 499 [2] S. Happy, A. Routray, Automatic facial expression recognition using features of salient facial patches, *IEEE Transactions on Affective Computing* 6 (1) (2015) 1–12.
- 500
- 501 [3] S. Koelstra, M. Pantic, I. Patras, A dynamic texture-based approach to recognition of facial actions and their  
502 temporal models, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 32 (11) (2010) 1940–  
503 1954.
- 504 [4] T. Bänziger, M. Mortillaro, K. Scherer, Introducing the geneva multimodal expression corpus for experimental  
505 research on emotion perception., *Emotion* 12 (5) (2012) 1161.
- 506 [5] G. McKeown, M. Valstar, R. Cowie, M. Pantic, M. Schröder, The semaine database: Annotated multimodal records  
507 of emotionally colored conversations between a person and a limited agent, *IEEE Transactions on Affective Com-  
508 puting* 3 (1) (2012) 5–17.
- 509 [6] F. Ringeval, A. Sonderegger, J. Sauer, D. Lalanne, Introducing the recola multimodal corpus of remote collaborative  
510 and affective interactions, in: *International Conference on Automatic Face and Gesture recognition (FG)*, IEEE,  
511 2013, pp. 1–8.
- 512 [7] <http://mplab.ucsd.edu>, The MPLab GENKI Database, GENKI-4K Subset.
- 513 [8] G. Zhao, M. Pietikainen, Dynamic texture recognition using local binary patterns with an application to facial  
514 expressions, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 29 (6) (2007) 915–928.
- 515 [9] K. Sikka, T. Wu, J. Susskind, M. Bartlett, Exploring bag of words architectures in the facial expression domain, in:  
516 *European Conference on Computer Vision (ECCV)*, Springer, 2012, pp. 250–259.
- 517 [10] E. Krumhuber, A. Kappas, A. Manstead, Effects of dynamic aspects of facial expressions: a review, *Emotion  
518 Review* 5 (1) (2013) 41–46.
- 519 [11] S. Yang, B. Bhanu, Facial expression recognition using emotion avatar image, in: *International Conference on  
520 Automatic Face and Gesture recognition (FG)*, IEEE, 2011, pp. 866–871.
- 521 [12] C. Liao, H. Chuang, C. Duan, S. Lai, Learning spatial weighting via quadratic programming for facial expression  
522 analysis, in: *Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, 2010, pp. 86–93.
- 523 [13] G. Sandbach, S. Zafeiriou, M. Pantic, Markov random field structures for facial action unit intensity estimation, in:  
524 *International Conference on Computer Vision Workshops (ICCVW)*, 2013, pp. 738–745.
- 525 [14] S. Chew, P. Lucey, S. Lucey, J. Saragih, J. Cohn, I. Matthews, S. Sridharan, In the pursuit of effective affective com-  
526 puting: The relationship between features and registration, *IEEE Transactions on Systems, Man, and Cybernetics,  
527 Part B: Cybernetics* 42 (4) (2012) 1006–1016.
- 528 [15] C. Herrmann, C. Qu, D. Willersinn, J. Beyerer, Impact of resolution and image quality on video face analysis, in:  
529 *Advanced Video and Signal Based Surveillance (AVSS)*, 2015 12th IEEE International Conference on, IEEE, 2015,  
530 pp. 1–6.
- 531 [16] A. R. Rivera, O. Chae, Spatiotemporal directional number transitional graph for dynamic texture recognition, *IEEE  
532 Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 37 (10) (2015) 2146–2152.
- 533 [17] A. Dhall, A. Asthana, R. Goecke, T. Gedeon, Emotion recognition using phog and lpq features, in: *International  
534 Conference on Automatic Face and Gesture recognition (FG)*, IEEE, 2011, pp. 878–883.
- 535 [18] C. Qu12, E. Monari, T. Schuchert, J. Beyerer21, Adaptive contour fitting for pose-invariant 3d face shape recon-  
536 struction, *British Machine Vision Conference (BMVC)*.
- 537 [19] X. Zhu, Z. Lei, J. Yan, D. Yi, S. Li, High-fidelity pose and expression normalization for face recognition in the  
538 wild, in: *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 787–796.
- 539 [20] S. Li, X. Liu, X. Chai, H. Zhang, S. Lao, S. Shan, Morphable displacement field based image matching for face  
540 recognition across pose, in: *European Conference on Computer Vision (ECCV)*, Springer, 2012, pp. 102–115.
- 541 [21] L. A. Jeni, J. F. Cohn, T. Kanade, Dense 3d face alignment from 2d videos in real-time, in: *International Conference  
542 on Automatic Face and Gesture recognition (FG)*, Vol. 1, IEEE, 2015, pp. 1–8.
- 543 [22] B. Jiang, B. Martinez, M. F. Valstar, M. Pantic, Decision level fusion of domain specific regions for facial action  
544 recognition, in: *International Conference on Pattern Recognition (ICPR)*, IEEE, 2014, pp. 1776–1781.
- 545 [23] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with  
546 local binary patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 24 (7) (2002) 971–  
547 987.
- 548 [24] Z. Ambadar, J. Schooler, J. Cohn, Deciphering the enigmatic face the importance of facial dynamics in interpreting  
549 subtle facial expressions, *Psychological Science* 16 (5) (2005) 403–410.
- 550 [25] B. Allaert, I. M. Bilasco, C. Djeraba, Consistent optical flow maps for full and micro facial expression recognition,  
551 *International Conference on Computer Vision Theory and Applications (VISAPP)*, 2017.
- 552 [26] P. Sundberg, T. Brox, M. Maire, P. Arbeláez, J. Malik, Occlusion boundary detection and figure/ground assignment  
553 from optical flow, in: *Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2011, pp. 2233–  
554 2240.
- 555 [27] J. Revaud, P. Weinzaepfel, Z. Harchaoui, C. Schmid, Epicflow: Edge-preserving interpolation of correspondences  
556 for optical flow, in: *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp.  
557 1164–1172.

- 558 [28] P. Lucey, J. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, The extended cohn-kanade dataset (ck+):  
559 A complete dataset for action unit and emotion-specified expression, in: International Conference on Computer  
560 Vision and Pattern Recognition Workshop (CVPR), IEEE, 2010, pp. 94–101.
- 561 [29] M. Pantic, M. Valstar, R. Rademaker, L. Maat, Web-based database for facial expression analysis, in: International  
562 Conference on Multimedia and Expo (ICME), IEEE, 2005, pp. 5–pp.
- 563 [30] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, J. Movellan, Automatic recognition of facial actions in  
564 spontaneous expressions, *Journal of Multimedia* 1 (6) (2006) 22–35.
- 565 [31] S. M. Mavadati, M. H. Mahoor, K. Bartlett, P. Trinh, J. F. Cohn, Disfa: A spontaneous facial action intensity  
566 database, *IEEE Transactions on Affective Computing* 4 (2) (2013) 151–160.
- 567 [32] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural  
568 similarity, *IEEE Transactions on Image Processing* 13 (4) (2004) 600–612.
- 569 [33] S. Schaefer, T. McPhail, J. Warren, Image deformation using moving least squares, in: *ACM Transactions on*  
570 *Graphics (TOG)*, Vol. 25, ACM, 2006, pp. 533–540.
- 571 [34] C.-C. Chang, C.-J. Lin, Libsvm: a library for support vector machines, *ACM Transactions on Intelligent Systems*  
572 *and Technology (TIST)* 2 (3) (2011) 27.
- 573 [35] R. Chaudhry, A. Ravichandran, G. Hager, R. Vidal, Histograms of oriented optical flow and binet-cauchy kernels  
574 on nonlinear dynamical systems for the recognition of human actions.