INTENSITY ESTIMATION OF UNKNOWN EXPRESSION
BASED ON A STUDY OF FACIAL PERMANENT FEATURES DEFORMATIONS

KHADOUDJA GHANEM*, ALICE CAPLIER**, M.K. KHOLLADI*

*Laboratory MISC, Mentouri University , Rue Ain El Bey, Constantine, Algeria
**Gipsa Laboratory, 46 avenue Félix Viallet, F-38031 Grenoble, France
gkhadoudja@yahoo.fr

ABSTRACT

In this work we report on the progress of building a system that enables the intensity estimation of unknown expression based on a study of the degree of facial permanent features deformations from still images. The facial changes can be identified as facial action units which correspond to the movement of muscles. We analyze subtle changes in facial expression by interpreting the movement of the muscle by its corresponding distances computed from characteristic facial points. All changed distances, are compared with corresponding Thresholds, to be mapped to symbolic states that qualitatively encode how much a given distance differs from its corresponding value in the neutral state. The Transferable Belief Model is used to fuse all data which correspond to the whole of changed distances. Expression intensity is quantified as: High, medium or low. Different raisons are done to prove that is better to estimate expression intensity of unknown expression than of known one.

KEY WORDS: Facial expression, expression intensity, belief theory.

1   INTRODUCTION

Facial expressions correspond to facial changes in response to a person's internal emotional states, intentions, or social communications. It can vary in intensity and its analysis includes both measurement of facial motion and recognition of expression. The most approaches developed in facial expression analysis field were interested by expressions recognition. Then, few researchers were interested by estimating the intensity of a recognized expression. In this work we are interested by the opposite analysis, we are interested by estimating expression intensity in order to recognize the facial expression in future work. The intensity of a facial expression may be of interest for a variety of reasons. For example, in [18] Ekman found that the intensity of zygomatic major muscle action was correlated with retrospective self-reports about the intensity of happiness experienced. It means that by estimating intensity, we can recognize the facial expression. Besides that, the velocity of smile onsets in relation to intensity also appears to differ markedly between posed and spontaneous smiles [19].

The FACS manual of Ekman [1] is the first reference in using point intensity scale to describe intensity variation of action units. Several computer vision researchers proposed methods to represent intensity variation automatically. Mase and colleagues [3] used optic flow to estimate activity in a subset of facial muscles, Essa and pentland [4]

extended this approach by representing intensity variation in smiling using optical flow in a detailed anatomical and physical model of the face. Kimura and Yachida [5] and Lien et al [6] quantified intensity variation in emotion specified expression and in action units, respectively. Bartlett and colleagues [7] tested their algorithms on facial expressions that systematically vary in intensity as measured by manual FACS coding. Although they failed to report results separately for each level of intensity variation, their overall findings suggest some success. Tian and colleagues [8] may be the only group to compare manual and automatic coding of intensity variations. Using Gabor features and an artificial neural network, they discriminated intensity variation in eye closure as reliably as did human coders. These findings suggest that it is feasible to automatically recognize intensity variation within types of facial actions.

The main problem of all these methods as well as the method proposed in our precedent work[14], is the necessity of recognizing the facial expression before estimating its intensity which is not always obvious. Generally, only six universal facial expressions are recognized. Sometimes we need to estimate intensity of positive or negative expressions just to know the degree of the positivity or negativity (the mood) of a person.

In the present work, we propose a method which estimate the intensity of unknown expression based on the analysis

of all possible changes which occur on the face. Movement of all activated facial muscles are interpreted in terms of distances computed from characteristic facial points. All changed distances, are compared with corresponding Thresholds, to be mapped to symbolic states that qualitatively encode how much a given distance differs from its corresponding value in the neutral state. The result of the fusion data which correspond to all changed distances and done by the Transferable Belief Model is the scoring of the intensity quantified as: High, medium or low intensities.

The remainder of the paper is organized as follows: In section 2, the proposed method is presented. The Belief theory principle which is used to fuse extracted data from facial images, is briefly described in section 3. The facial data evaluation is discussed in section 4 and finally, section 5 provides concluding remarks and a short overview of the future work.

2 RELATED WORK

The method consists in extracting data from still images. The considered data correspond to the measurement of facial deformations. These data are computed from characteristic points of facial permanent features which are eyes, eyebrows and mouth. This step is not the main goal of our work. This is why we have selected the characteristic points of permanent features manually.

2.1 Measurement of facial deformations

When expressing an emotion, deformations appear on two regions of the face [1] and [9] :

2.1.1 Upper Deformations

Upper deformations are generally caused by the 12 upper Action units which are:

Upper Face Action Units					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
Inner Brow Raiser	Outer Brow Raiser	Brow Lowerer	Upper Lid Raiser	Cheek Raiser	Lid Tightener
*AU 41	*AU 42	*AU 43	AU 44	AU 45	AU 46
Lid Droop	Slit	Eyes Closed	Squint	Blink	Wink

Figure 1: Upper Face Action Units.

All these action units describe motions and deformations of two permanent features which are eyes and eyebrows, Au1,AU2 and AU4 concern the motion of eyebrows. AU5, AU6, AU7, AU41, AU42, AU43, AU44, AU45 and AU46

describe the motion of eyelids.

According to the FACS system, only AU41, 42 and 43 which concern the closing eye, can be scored on intensity.

To interpret Action units in terms of distances, we have based our work on [10]. So we consider the distance"D1"(distance between two eyelids) to compute how much the eye is opened or closed (see figure 3).

The information provided by AU1, 2, 4 which concern the movement of the brow is not materialized; this is why we consider the distance"D2"'distance between eye and brow corners) to compute how much the brow moves (see figure 3).

2.1.2 Lower deformations

Lower deformations are generally caused by the 18 lower action units which are:

Lower Face Action Units					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
Nose Wrinkler	Upper Lip Raiser	Nasolabial Deepener	Lip Corner Puller	Cheek Puffer	Dimpler
AU 15	AU 16	AU 17	AU 18	AU 20	AU 22
Lip Corner Depressor	Lower Lip Depressor	Chin Raiser	Lip Puckerer	Lip Stretcher	Lip Funneler
AU 23	AU 24	*AU 25	*AU 26	*AU 27	AU 28
Lip Tightener	Lip Pressor	Lips Part	Jaw Drop	Mouth Stretch	Lip Suck

Figure 2: Lower Face Action Units.

All these action units describe motions and deformations of mouth. The action units from AU9 to AU20, concern horizontal motion of the lips, and the AU22 to AU28, describe the vertical motion of the lips.

According to the FACS system, only the AU25, 26 and 27 which correspond to the vertical opening of the mouth, can be scored on intensity. We associate the distance "D4" (distance between lips)to compute how much the mouth is opened vertically (see figure 3).

If we consider only this information, we can lose information about the horizontal opening of the mouth. For this purpose we consider the distance "D3" (distance between mouth corners)to compute how much the mouth is opened horizontally (see figure 3).

In order to bind upper deformations with lower ones, we add another distance "D5" which corresponds to the distance between the eye (upper feature) and the mouth (lower feature). At the end we get five distances to measure the degree of permanent features deformations like it is shown in figure 3:
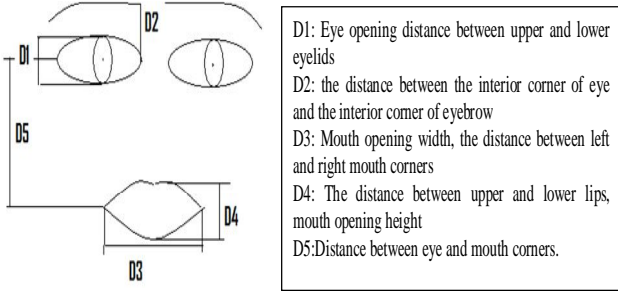
Figure 3: Computed Distances extracted from facial characteristic points.

All computed distances are normalized with respect to the distance between the centers of both irises. This makes the analysis independent on the variability of face dimensions and on the position of the face with respect to the camera.

2.2 Definition of Symbolic States

We associate a state variable V_i ($1 \leq i \leq 5$) to each characteristic distance D_i in order to convert the numerical value of the distance to a symbolic state. The analysis of each variable shows that V_i can take three possible states, $\Omega' = \{$low, medium, high$\}$; $2^{\Omega'} = \{$low, medium, high, lowUmedium, mediumUhigh$\}$ where lowUmedium states the doubt between low and medium, and mediumUhigh states the doubt between medium and high. We assume that impossible symbols (for example lowUhigh) are removed from $2^{\Omega'}$.

2.3 Modeling Process

The modeling process aims at computing the state of every distance D_i and at associating a piece of evidence. To carry out this conversion, we define a model for each distance using the states of $2^{\Omega'}$ (Figure 4).
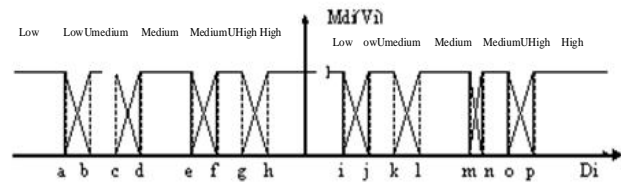

Figure 4: Proposed model.

One model is defined for each characteristic distance independently of the facial expression. If the calculated distance increase , we consider the right half part of the model from "I" to "p" thresholds, and if the calculated distance decrease, we consider the left half part of the model from "a" to "h" thresholds.

Each D_i represent a distance between two characteristic points of permanent features, it can increase or decrease, when expressing a surprise, the distance D1 which describe the opening of the eye, evolves from a distance corresponding to the neutral state (minimum value), to a larger distance corresponding to expressive state (middle or maximum value) so that the state variable V1 evolves from the state (low) to a higher state (medium) via an undetermined region lowUmedium or to a significantly higher state (high) via an undetermined region mediumUhigh.

In the same way, when expressing disgust, the distance D1 which describe the opening of the eye, evolves from a distance corresponding to the neutral state (maximum value), to a smaller distance corresponding to expressive state (middle or minimum value) so that the state variable V1 evolves from the state (high) to a lower state (medium) via an undetermined region highUmedium or to a significantly lower state (low) via an undetermined region mediumUlow. For each value of Di, the sum of the pieces of evidence of Di states is equal to 1.

$$m_{Di} : 2^{\Omega'} \to [0, 1]$$
$$V_i \to m_{Di}(V_i) \qquad (1)$$

The piece of evidence $m_{Di}(V_i)$ is obtained by the function depicted in Figure 4.

2.4 Definition of Thresholds

Thresholds {a,b,…. p} of each model state are defined by statistical analysis on (Hammal_Caplier) Database. The database contains 21 subjects. The database have been divided into a learning set called HCEL (13 subjects and 4 expressions: Joy, Surprise, Disgust and neutral, 4680 frames) and a test set called HCET (8 subjects and 4 expressions: Joy, Surprise, Disgust and neutral, 3840 frames). The learning set is then divided into expressive frames noted HCELe and neutral frames HCELn. The minimum threshold "a" is averaged out over the minimum values of the characteristic distances from the HCELe database. Similarly, the maximal threshold "p" is obtained from the maximum values. The middle thresholds "h" and "I" are defined respectively as the mean of minimum and maximum of the characteristic distances from the HCELn. The threshold "b" is the median of the characteristic distances values for facial images assigned to the highest state. "g" is the median of the characteristic distances values for facial images assigned to the lowest state. The intermediate threshold "d" is computed as the mean of the difference between the limit thresholds "a" and "h" divided by three augmented by the value of the threshold "a". Likewise the threshold "e" is computed as the mean of the difference between the limit thresholds "a" and "h" divided by three reduced by the value of the threshold "h". The thresholds "c" and "f" are computed as the mean of thresholds "b" and "d" respectively "e" and "g". Thresholds from positive part of the proposed model are computed

similarly.

After computing thresholds, the five distances are compared to these thresholds, to associate a state from $2^{\Omega'}$={low, medium, high, lowUmedium, mediumUhigh} to each distance. Only changed distances with respect to the neutral state are considered.

To take on the count all changes appearing on the face with unknown expression, we proceed to a data fusion by using the Transferable Belief Model.

3 THE TRANSFERABLE BELIEF MODEL

Initially introduced by Dempster [11] and Shafer [12] and enriched by Smets [13], the belief theory considers a frame of discernment Ω of N exhaustive and exclusive hypotheses characterizing some situations. This means that the solution of the considered problem is unique and that it is obligatorily one of the hypotheses of Ω. This approach takes into account the uncertainty of the input information and allows an explicit modeling of the doubt between several hypotheses. It requires the definition of a Basic Belief Assignment (BBA) that assigns an elementary piece of evidence m(A) to every proposition A of the power set 2^Ω. The function m is defined as:

$$m : 2^\Omega \rightarrow [0, 1]$$
$$A \rightarrow m(A), \ \sum m(A) = 1, \ A \subseteq \Omega \qquad (2)$$

In our application, the assumption "low" corresponds to the minimum or low expression intensity; "medium" corresponds to the medium intensity and "high" corresponds to the maximum or high intensity. 2^Ω corresponds to single expression intensities or to combinations of expression intensities, that is 2^Ω = {low, medium, high, (lowUmedium), (mediumUhigh),…}, and A is one of its elements. The salient character of the transferable belief model is the powerful combination operator that allows the integration of information from different sensors. The Basic Belief Assignment (BBA) associated to each characteristic distance, can be viewed as independent sources of information that scores their belief in a proposition given some observations. These BBAs are combined to take into account all the available information about the facial expression using the Dempster combination law (conjunctive combination). Given the BBAs mDi and mDj of two characteristic distances, the joint Basic Belief Assignment mDij is given using the conjunctive combination (orthogonal sum) as:

$$m_{Dij} (A) = (m_{Di} \oplus m_{Dj})(A)$$
$$= \sum m_{Di}(B)m_{Dj} (C) \qquad (3)$$
$$B \cap C = A$$

A, B and C denote propositions and B ∩ C denotes the conjunction (intersection) between the propositions B and C.

In our case, Ω' = {low, medium, high}; so Vi can take its value from the set: $2^{\Omega'}$ = {low, medium, high, lowUmedium, mediumUhigh}. To formulate distance states by expression intensity we can use table 1:

Table 1: Different states taken by a distance and its corresponding Expression intensity.

Expression	ELow	Emedium	Ehigh	ElowUEmedium	EmediumUEhigh
Vi	low	medium	high	lowUmedium	mediumUhigh

The meaning of this table is that the piece of evidence associated to the state low of the characteristic distance D is equivalent to the piece of evidence of the expression Elow:

V=low ; mD(low)=mD(Elow);

In the same way: V=medium ; mD(medium)=mD(Emedium);

V=high ; mD(high)=mD(Ehigh);

V=lowUmedium ; mD(lowUmedium)=mD(ElowUEmedium);

V=mediumUhigh ; mD(mediumUhigh)=mD(EmediumUEhigh).

To be more explicit, if we consider two distances (D1 and D2) in respect of the neutral state, so that V1=medium and V2=mediumUhigh => mD1(medium)=mD1(Emedium), mD2(mediumUhigh)=mD2(EmediumUEhigh)

We can use the orthogonal sum to join the two distances:

D1\D2	EmediumUEhigh
Emedium	Emedium

So, for the intensity of the expressed emotion with these two changed distances, we can conclude that it is an expression with "medium intensity". Sometimes, the empty set φ can appear and allows handling conflicts between incoherent sources. Since any expression must have an intensity, the underlying expression intensity is assigned to middle intensity, noted Emedium and its piece of evidence is equal to the resulting piece of evidence of the empty set because it cannot be a high intensity for the reason that one or more of the changed distances has reached at the maximum the mean level, and it cannot be low intensity for the reason that one or more of the changed distances has reached at least the mean level.

4   EXPERIMENTS AND RESULTS

Because of lack of intensity databases, only three face databases are tested in order to evaluate the efficiency of the proposed approach.

4.1   Databases

4.1.1   *Caplier _H database [15]*

It is composed of video recordings of 21 subjects with different gender and ethnicity performing 4 different facial expressions, namely Joy, Surprise, Disgust each one beginning and ending by Neutral expression. Because of the difficulty to simulate Sadness, Anger and Fear facial expressions for non-actor subjects, these expressions were not recorded in the database and not considered in the experiments. Each video sequence has been recorded at 25Hz image rate and is 5s long(100frames). The subject starts with a Neutral expression, performs the required facial expression and returns to the Neutral expression. The whole database was manually labeled, meaning that a human expert assigned a facial expression to each image.

For the expertise step of the Belief theory, we have considered 10 subjects for each expression. 10 cases of low intensity which correspond to the first frame of the video recording where a human expert can distinguish the first changes on the face; 10 cases with high intensity which correspond to the apex of each expression; 10 cases with medium intensity taken from the video recordings corresponding to the face changes from the expressionless to the expression with maximal intensity and 10 other cases of medium intensity taken from the video recordings corresponding to changing back to an expressionless face. The dimensions of each image are 320X240.

4.1.2   *EEBase Database [16]*

It is composed of 43 subjects, with different gender and ethnicity , 24 males and 19 females, of which 6 Africans males,4 Africans females, and 5 females Asiatic, and for each subject, we have 16 frames in neutral, joy, disgust, sadness, anger, surprise and fear expressions. For all expressions we have two intensities medium and high, except for surprise we have only one intensity : high. So we have 260 images with high intensity, and 197 images with medium intensity. The dimensions of each image are 506X650.

4.1.3   *Dafex Database [17]*

The DaFEx database consists of 1008 short video clips, lasting between 4 and 27 sec., each showing a facial expression corresponding to one of the 6 Ekman's emotions [8], [9] – happiness, surprise, fear, sadness, anger and disgust – plus the neutral expression. The expressions were acted by 8 Italian professional actors (4 male and 4 female), who recorded each emotion at 3 intensity levels (low,

medium and high), and in 2 recording conditions ("Utterance" and "No utterance"). In the "Utterance" condition, the actor produced the emotional expression while uttering a phonetically rich balanced sentence ("In quella piccola stanza vuota c'era però soltanto una sveglia:" In that little empty room there was only an alarm clock). In the "No-utterance" condition emotions were acted without uttering any sentence. The entire set of emotions was recorded by every actor more than once: four times in the "Utterance" condition and twice in the "No-utterance" condition.

To get still images from these videos, we have considered the apex of each video and for each actor.

After that, we have computed all considered distances D1 to D5 for each image, and then associate a state to changed distances, at the end we have applied the belief theory to proceed to the fusion of data. The dimensions of each image are 360X288.

4.2   Results when applying the Belief theory

4.2.1   *Results on Dafex and Caplier_H databases*

The performances of the resulting classification system are evaluated on all actors of the Dafex database. Before giving final results and to be more explicit we give an example of extracted data while associating a state to each computed distance for an actor with different intensities for different expressions (cf. fig.5):

Figure 5: Images expressing six expressions with three intensities for each expression of Actor 4 from Dafex Database used to give table 2

Figure 5 presents 18 images of an actor from the Dafex database. From the left to the right, and from the top to the bottom, we can see the six universal expressions: Anger, Disgust, Fear, Happy, Sadness and Surprise, each expression is given with three intensities: High, Medium and low.

Extracted data from these 18 images corresponding to changed distances are given in table 2.

Table 2: states of distances (Extracted from an actor from Dafex database) used to recognize expression intensity by the B.T

Frame	D1	D2	D3	D4	D5	Recognition by B.T	Reality in the Database²
A4HAP	HIGH		MEDIUM U HIGH	HIGH	HIGH	HIGH	Hap_HIGH
A4ANG		HIGH		MEDIUM U HIGH		HIGH	Ang_HIGH
A4DIS	HIGH	MEDIUM U HIGH		HIGH		HIGH	Dis_HIGH
A4SUR	HIGH	HIGH		HIGH		HIGH	Sur_HIGH
A4SAD	HIGH	HIGH			HIGH	HIGH	Sad-HIGH
A4FEA	MEDIUM U HIGH	HIGH		MEDIUM U HIGH		HIGH	Fea_HIGH
A4HAP	MEDIUM		MEDIUM	MEDIUM	MEDIUM	MEDIUM	Hap_MEDIUM
A4ANG	LOW U MEDIUM			MEDIUM		MEDIUM	Ang_ MEDIUM
A4DIS	LOW U MEDIUM			MEDIUM		MEDIUM	Dis_ MEDIUM
A4SUR	MEDIUM	HIGH		MEDIUM		ERROR	Sur_ MEDIUM
A4SAD	MEDIUM	MEDIUM			MEDIUM	MEDIUM	Sad- MEDIUM
A4FEA	MEDIUM			MEDIUM		MEDIUM	Fea_ MEDIUM
A4HAP	LOW		LOW	LOW U MEDIUM		LOW	Hap_LOW
A4ANG		LOW		LOW U MEDIUM		LOW	Ang_ LOW
A4DIS	LOW			LOW		LOW	Dis_ LOW
A4SUR	MEDIUM	MEDIUM		MEDIUM		MEDIUM	Sur_ LOW
A4SAD	LOW	LOW				LOW	Sad- LOW
A4FEA	LOW	LOW		MEDIUM		ERROR	Fea_ LOW

The first column represents the studied frames which correspond to frames in figure 5. From column two to column six, we can have all considered distances computed from characteristic points. If the computed distance has changed we associate to it a state from $2^{\Omega'}$. The seventh column represents the estimated intensity of the studied face using the belief theory model and the last column gives the reality of the intensity as it is labeled in the database. As an example, with the first frame (A4HAP) four distances have changed (D1, D3, D4, D5), the associated states of each distance correspond to (High, MediumUHigh, High, High). When applying the fusion of all these changed distances by the Dempster law, we get the estimation of the studied intensity which is High intensity. We can observe that the result is true because the reality of this intensity as it is labeled in the database is high. For the frame labeled 'A4SUR', changed distances are (D1, D2 and D4), the associate states are (Medium, High and Medium). The fusion data gives an error; the reality in the database is Medium. Final results are given in Table 3:

Table 3: Classification rates of the system for Dafex and Hammal_caplier databases

	Int.LOW	Int.MEDIUM	Int.HIGH
Recognized intensity	56/78 71,79%	72/78 92,31%	59/78 75,64%
LOWUMEDIUM	17/78 21,79%		
MEDIUMUHIGH			12/78 15,38%
ERROR	5/78 6,41%	6/78 7,69%	7/78 8,97%
Total recognized	93,58%	92,31%	91,02%
TOTAL	100%	100%	100%

Rates of recognized low intensity, medium and high are given in the first row, rates of doubt between low and medium or medium and high are given in the second and third rows, errors found in each level of intensity are given in the next row, in the last one we have the global rates.

It can be observed that medium intensity yields good classification rates (without doubt with other intensities) compared with high and low intensities. This can be explained by the fact that it is the easiest intensity to simulate for all actors.

Classification rates of low intensity decreases in favor of the doubt between low and medium intensities. This can be explained by the fact that low intensity is the more difficult intensity to simulate because the simulation must be sufficiently small so that the intensity is considered to be minimal and sufficiently large so that the expression is recognized.

Another doubt is detected between medium and high intensities, indeed, it is better that the classification system accepts doubt about two intensities and does not try to discriminate them. The Transferable Belief Model is actually well adapted for such a scenario.

Errors occur because of the variability of the actors to express an emotion. It can be observed too, that the most changes appeared on the face when expressing an emotion, are based on two or three distances, like it has been proved in our precedent work [14].

Generally, recognized rates for the three intensities are almost equal.

In another hand, we have noted that, the recognition intensity is better with unknown expressions than with known ones, because when we want to estimate intensity of known expression, we look for all considered distances associated to each expression (surprise : D1,D2,D4; Joy : D1,D3,D5; Disgust : D1,D2,D4; anger : D1,D2,D4; Fear:D1,D2,D4 and Sadness : D1,D2,D4) [14], but with an unknown expression, we look only for changed distances , and not all distances. For example, for image in figure 6, only two distances D1, D2 change so that:

V1=low U medium and V2=medium => mD1(low U medium) =mD1(Elow U Emedium)=1;

mD2(medium)=mD2(Emedium)=1;

=> mD1(Elow U Emedium) ⊕ mD2(Emedium)= mD12(Emedium)=1

=> Expression intensity=medium.

However, if we know that expression is "Surprise" we consider three distances D1, D2, and D4 [14], and then we get an error because V4=low and the joint distances mD12(Emedium) ⊕ mD4(Elow)= φ .

We also observed that if we have 0 changed distances, we are in the neutral state, so no intensity to estimate.

Figure 6: Surprise expression with two changed distances.

4.2.2   *Results on Eebase Database:*

The main problem with this database is the absence of images with low intensity and the bad labeling of most of images with medium intensity. These images are labeled as images with medium intensity and they have been quantified as images with high intensity. For that reason the recognition rate of medium intensity has decreased compared to the same rate with the other databases (see table 4) in favor of high intensity. This can be explained by the fact that most of images are badly labeled like it is shown in figure 7:

Table 4: Classification rates of the system for EEbase database.

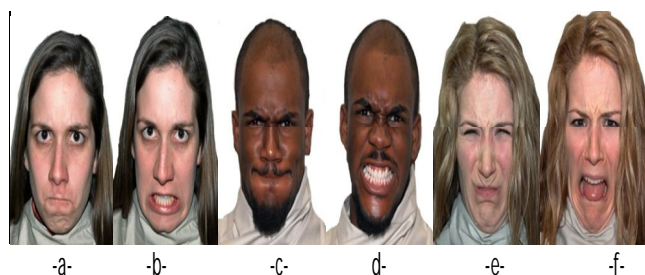|	Int.LOW	Int. MEDIUM	Int. HIGH
Recognized intensity	Images not disponible	120/197 61%	224/260 86,15%
LOW		2/197 1%	/
MEDIUM			6/260 2,31%
HIGH		42/197 21,32%	/
LOWUMEDIUM		8/197 4,06%	/
MEDIUMUHIGH		18/197 9,13%	17/260 6,54%
ERROR		7/197 3 ,55%	13/260 5%
Total recognized		74,10%	92,69%
TOTAL		100%	100%

Figure 7: Badly labeled images : (a,c,e) are labeled as images with Medium intensity?; (b,d,f) are labeled as images

With high intensity (what is the difference?), all these images are quantified as images with high intensity.

However, classification rate of high intensity is very comparable with the one of the other databases.

Other observations have been noted with this database about the different raisons which prove that is better to estimate intensity of unknown expression: We noted that some expressions are expressed differently from one subject to another. For example in case of anger we can open the mouth vertically or tighten completely the lips like in figure 8, the description is not the same, but the intensity is given correctly. With fear, we can open the mouth horizontally, or open it vertically, in both descriptions; intensity is given correctly like it is shown in figure 9. For that raison, it is better to estimate intensity of unknown expression than of known one.
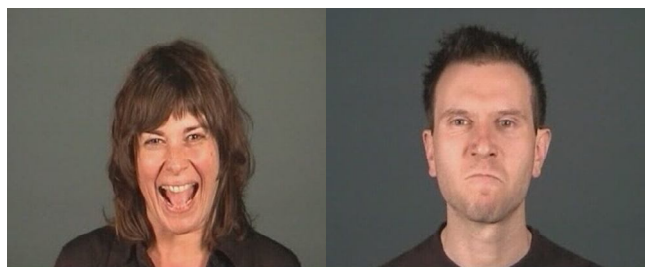

Figure 8: Two Different descriptions of Anger with high intensity.

Figure 9: Two different descriptions of Fear with high intensity.

Sometimes, when the intensity decreases, the number of changed distances decreases too. For example, for disgust

or fear expression with high intensity there are three changed D1,D2,D4 but with low intensity there are only two changed distances D1,D2 (see figure 10). This is another raison, for that it is better to estimate intensity of unknown expression than of known one (like it has been mentioned bellow).

Figure 10: images with decreasing the number of changed distances when

Decreasing expression intensity (a,c): high intensity, (b,d): medium intensity.

We noted that sometimes, from intensity to another, changed distances are not the same like it is shown on figure (11). This is another raison, for that it is better to estimate intensity of unknown expression than of known one(like it has been mentioned bellow).

Figure 11 : Changed distances in different intensities

(D1,D2,D5 : with medium intensity; D1,D2,D4 : with high intensities).

Finally, we noted that when estimating intensity, and when we have D3 and D4 changed, if D3 has states "mediumUhigh" or "high", D4 cannot have the same state because the two distances D3 and D4 are very dependent. So when the mouth is largely open horizontally and reaches its maximum, it cannot be largely open vertically at the same time while keeping the distance D3 at its maximum. If we keep the states of the two distances we get a wrong estimation of intensity. This is why we do not consider D4, when the state of D3 is "mediumUhigh" or "high", this case have been observed especially with joy expression.

CONCLUSION:

In this paper, we have presented a new method to recognize intensity of human facial expression. What is interesting in

this work is that it is not necessary to recognize the expression in order to quantify its intensity. An extraction of data followed by an understanding of changes appeared on the face can quantify the real intensity of expressions. This method takes into account the most important changes which appear on human face when expressing any emotion. By interpreting these changes to distances, results given by our method have proved that the most important factor to estimate expression intensity is the degree of geometrical deformation of facial structures which are interpreted by the proposed distances (D1, D2, D3, D4 and D5). Since the Transferable Belief Model has proved its ability to deal with imprecise data, and its interest to model the doubt between expression intensities, it is used for the fusion of available information to provide more reliable decisions. Different raisons from the reality are done to prove that is better to estimate expression intensity of unknown expression than of known one.

In the future work, we aim to recognize facial expressions based on the recognition of expression intensity and to discriminate between posed and spontaneous expressions.

REFERENCES

[1] P.Ekman, Wallace V. Friesen, Joseph C. Hager, '"Facial action coding system: A technique for the measurement of facial movement" Psychology, Physiological 1978, 2002. ISBN 0-931835-01-1 1.

[2] K.Mase, "Recognition of facial expression from opticzal flow". IEICE Trans. E, vol. 74,n°10 , pp 3474-3483 1991.

[3] I. Essa and A. Pentland, 1997 "Coding, analysis, interpretation, and recognition of facial expressions". IEEE Trans. on Pattern Analysis and Machine Intell., 19(7):757–763.

[4] S. Kimura and M. Yachida, 1997 "Facial expression recognition and its degree estimation". In Proc. Of the Int. Conf. on Computer Vision and Pattern Recognition, 295–300.

[5] J.J. Lien, T. Kanade, J. Cohn, and C. Li, 1998 "Subtly different facial expression recognition and expression

[6] Intensity estimation". In Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition, 853–859.

[7] M. Bartlett, J. Hager, P.Ekman, and T. Sejnowski, 1999. "Measuring facial expressions by computer image analysis". Psychophysiology, 36:253–264.

[8] Y.-L. Tian, T. Kanade, and J. Cohn, Sept 2000 "Eye-state action unit detection by Gabor wavelets". In Proc. of Int. Conf. on Multi-modal Interfaces, 143–150.

[9] J.N Bassili. Emotion Recognition: The role of facial movements and the relative importance of upper and lower areas of the face. Journal of personality and social psychology. vol. 61, no. 11, pp. 298- 307, 1979.

[10] Hammal Z., Caplier A. and Rombaut M. "A fusion process based on belief theory classification of facial basic emotions". Proc. the 8th International Conference on Information fusion, (ISIF), Philadelphia, PA, USA, 2005.

[11] A. Dempster, 1967 "Upper and lower probability inferences based on a sample from a finite univariate". Biometrika, 54: 515-528.

[12] Shafer, Glenn, 1976. A Mathematical Theory of Evidence. Princeton University Press.

[13] [13] P. Smets, R. Kennes, 1994 "The Transferable Belief Model". Artificial Inteligence, 66 (2): 191-234.

[14] K.Ghanem, A.Caplier "Estimation of Facial Expression Intensity Based on the Belief Theory" VISAPP

[15] 2008: Proc. the Third International Conference on Computer Vision Theory and Applications, Funchal, Madeira,Portugal, - Volume 1 452-460,2008.

[16] Hammal_CaplierDatabase : :

[17] http://www.lis.inpg.fr/pages_perso/caplier/english/emotionn elle.html.en/emotionnelle_2.html.en.html.

[18] EEBase data base http://www.cs.cmu.edu/afs/cs/project/cil/ftp/html/v-images.html.

[19] Dafex Database: http://tcc.itc.it/research.i3p/dafex/index.html.

[20] Ekman, P. Facial asymmetry. Science, 209, 833-834. (1980).

[21] Cohn, J. F., & Schmidt, K. L. The timing of facial motion in posed and spontaneous smiles. International Journal of Wavelets, Multiresolution and Information Processing, 2, 1-12 (2004).