HAL
open science

Towards a full emotional system

Khadoudja Ghamen, Alice Caplier

To cite this version:

Khadoudja Ghamen, Alice Caplier. Towards a full emotional system. Behaviour and Information Technology, 2012, 32 (8), pp.783-799. 10.1080/0144929X.2011.624639 . hal-00741441

HAL Id: hal-00741441

https://hal.science/hal-00741441v1

Submitted on 12 Oct 2012

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.

• Article Title : Towards a Full Emotional System.

• Journal name : Behavior and Information Technology.

• The full name and details of the corresponding author : Khadoudja Ghanem, MISC laboratory , university Mentouri, Route Ain El bey 25000 Constantine, Algeria.

Phone number: 00213661375673

Fax: 0021331632376

Email: gkhadoudja@yahoo.fr

• The full names and details of all co-authors of the paper : Alice Caplier, GIPSA laboratory, 961 avenue de la Houille Blanche, F-38410 St Martin d'Hères, France

Phone number: 0033476574363

Fax : 0033476574790

Email: alice.caplier@gipsa-lab.grenoble-inp.fr

Towards a Full Emotional System

Khadoudja Ghanem[1], Alice Caplier[2]

[1]MISC Laboratory, Mentouri University , Rue Ain El Bey, Constantine, Algeria ;
[2]Gipsa Laboratory, 961 avenue de la Houille Blanche, F-38410 St Martin d'Hères, France

Abstract: The present paper proposes a system both able to classify a facial expression in one of the six universal categories namely: (*Joy*, *Disgust*, *Anger*, *Sadness*, *Fear* and *Surprise*) and to assign to each expression its intensity in the range: (*High*, *Medium* and *Low*). This is carried out in two independent and parallel processes. Permanent and transient facial features are detected from still images and pertinent information, about the presence of transient features on specific facial regions and about facial distances computed from permanent facial features, is extracted. Both classification and quantification processes are based on transient and permanent features. The belief theory is used with the two processes because of its ability in fusing data coming from different sensors. The system outputs a recognized and quantified expression. The quantification process allows recognizing a new subset of expressions deduced from the basic ones. Indeed, by associating to each universal expression three intensities low, medium and high we deduce three facial expressions. Finally a set of eighteen facial expressions is categorized instead of the six universal ones. Experimental results are given to show the classification accuracy of the system.

Key Words: Facial expression, Permanent and Transient features, Classification, Expression intensity, belief theory.

1. Introduction

In the field of human computer interaction (HCI), human face is the richest source of nonverbal communication and the most accessible interface displaying human emotion. In order to make this communication as natural as possible, it is necessary to equip computers with an emotional system. The emotional system must not only recognize facial expressions, but must also estimate their intensity.

Intensity refers to the magnitude of the face changes resulting from the activation of facial muscles. The intensity of a facial muscle may be of interest for a variety of reasons. For example, in (Ekman, 1980) 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 between posed and spontaneous smiles (Cohn and Schmidt, 2004).

Researchers in facial expressions field are influenced by Ekman, Friesen and Izard who concentrate generally on six universal expressions: (*joy*, *disgust*, *surprise*, *sadness*, *anger*, *fear*). By estimating expression intensity, we make leave sub expression classes; it means that we recognize other expressions. To each new expression, we associate an appropriate reaction; this action allows creating an "expert system rule". (Table 1) summarizes the new considered set of expressions, and their corresponding reactions.

Table 1. New recognized Expressions deduced by quantification of the universal expressions with the appropriate reactions (Rules Expert system)

Universal Expressions	Intensities	New Recognized Expressions	Corresponding Reactions
Joy	Low	Well being	Act Positively
	Medium	Joy	Roar with laughter
	High	Happiness	Shout and Jump
Disgust	Low	Feel sick	Make grimaces
	Medium	Disgust	Use Disgust words
	High	Bitterness	Vomit
Surprise	Low	Astonished	Ask questions
	Medium	Surprise	Make grimaces
	High	Amazement	To Freeze
Anger	Low	Boredom	Dispute
	Medium	Anger	Shout and Howl
	High	Rage	Strike, Break and Kill
Sadness	Low	Trouble	Insulate itself
	Medium	Sadness	Cry
	High	Abatement	Depress and commit suicide
Fear	Low	Anxiety	Make grimaces
	Medium	Fear	Shout and Cry
	High	Terror	Hide, Flee and Disappear

[Table 1]

The recognition of facial expressions with their intensity is involved in decision-making process to recognize the interlocutor's behavior, either automatically or with the intervention of a human being. It will not influence the dialogue in the same way according to its degree. Thus, a slightly

irritated person will not behave in a so violent way as a furious person against his/her interlocutor. In fact; if *Anger* is expressed with low intensity, the expression is labeled as "Boredom" and the associated reaction is "Dispute", if *Anger* is expressed with medium intensity, the expression is labeled as "Anger" and the associated reaction is "Shout and Howl" and if *Anger* is expressed with high intensity, the expression is labeled as "Rage" or "Furor", and the associated reaction is "Break, Strike or Kill".

It will be a great challenge and of practical significance to develop a computer vision system which can automatically recognize a variety of facial expressions and estimate expression intensity.

One limitation of the existing facial expression recognition methods is that most of the proposed systems perform only facial expression classification into the basic emotion categories defined by Ekman and Friesen (Ekman And Friesen, 1975) (with some exceptions (Tsapatsoulis et al, 2000)) which is a little bit restrictive because humans able of expressing thousands of emotions.

One limitation of the existing facial expression quantification methods is that only few expressions have been quantified (generally: *Joy*, *Anger*, *Sadness* and *Surprise*, rarely the six universal expressions).

Yet, none of these systems performs recognition and quantification of facial expression or action units in parallel to give one result which is a quantified recognized expression.

In this paper, we propose a system which recognizes a considered expression and quantifies it. The system is performed by using the transferable belief model (TBM) which takes into account the ability of humans to express a mixture of emotions, in both classification and quantification process. As a result and by assigning to each universal expression one of the three possible intensities: *High*, *Medium* and *Low*, we get a set of eighteen quantified expressions instead of six.

This paper is an extended version of our previous work described in (Ghanem and Caplier, 2008) where we evaluated the transferable belief model for facial expression classification only. The main contributions of this paper are summarized as follows:

- We evaluate the transferable belief model in the estimation of facial expression intensity;
- We cooperate in creating a rule expert emotional system;
- We formulate the classification process in parallel with the quantification one in order to associate to the considered expression two labels at the same time, a label concerning the expression class (*Joy*, *Disgust*, *Anger*, *Sadness*, *Fear*, *Surprise* or *Neutral*) and a label about the intensity level (*Low*, *Medium* or *High*).

The paper is organized as follows. We present a brief review of related work in section 2. Our main contribution which proposes a new emotional system which gathers both classification and quantification is described in section 3. Section 4 is devoted to the experiments on the (Dafex database). Discussion and conclusion are given in the last section.

2. Previous work

Automatic facial expression recognition has attracted much attention from behavioral scientists, contrary to automatic facial expression quantification. But any interest was related to realize a complete system which recognizes and quantifies a given expression. Here we briefly review some previous works in order to put our work in context.

2.1 Facial Expression Recognition

Different systems were used to classify facial expressions. Three main approaches are proposed for still images, namely: Template-Based-Methods, Neural Based Methods and Rule Based Methods. For more details on the different classification approaches, we can refer to (Pantic, 2000) or (Fasel, and Luettin, 2003). Methods in the Template-Based-Methods category perform expression classification into a single basic emotion category. If a template-based

classification method is applied, the encountered facial expression is compared to the template defined for each expression category. The best match decides the category of the shown expression.

Different techniques like the Linear Discriminant Analysis with the Mahalonobis distance measure (Edwards et al 1998), the Elastic Graph Matching with the personalized galleries (Hong et al, 1998), the Principal Component Analysis with action parameters (Huang and Huang, 1997), the Hopfield neural networks with optical flow (Yoneyama et al, 1997) were used.

The achieved recognition rate varies from 74% to 92.2 % in the case of familiar subjects and from 73 % to 75% in the case of unknown persons.

In general, it is difficult to achieve a template-based recognition of a non-prototypic facial expression. There are a lot of combinations of different facial actions. The problem becomes even more difficult due to the fact that everybody has his/her own maximal intensity of displaying certain facial actions.

Methods belonging to the Neural-Network-Based category are not able to model the doubt between facial expressions, they perform expression classification into a single basic emotion category it means as one of the six universal expression.

For classification of expressions into one of the six basic emotions categories (Zhang, Z. et al, 1998), (Kobayashi and Hara, 1997), (Padgett and Cottrell, 1996), (Zhao and Kearney, 1996), a back-propagation neural network is applied. The inputs of the neural network differ from one method to another. But the output of all neural networks used in each method corresponds to the set of the six universal expressions. In a neural-network-based classification approach, a facial expression is classified according to the categorization process which has been learned by the network during the training phase. The average recognition rate varies from 85% to 100%. The limitation is that those methods were tested only on the set of images used for training. It is not

known how the method will behave in the case of an unknown subject. In addition, most of these methods require manual intervention.

For Rule-Based Methods, the prototypic expressions, which characterize the emotion categories, are first described in terms of facial actions. Then, a candidate expression, also described in terms of facial actions, is compared to each prototypic expression. Just one of the surveyed methods for automatic facial expression analysis from static images applies a rule-based approach to expression classification (Pantic and Rothkrantz, 2000). The average recognition rate was 92 % for the upper face AUs and 86 % for the lower face AUs. The problem is the difficulty to recognize all the 44 action units defined in FACS.

2.2 Facial Expression Quantification

Systems developed to estimate facial expression intensity can be broadly divided into two main approaches: holistic approaches and local feature approaches.

Holistic approaches take into account the information of a face image as a whole (Chandrasiri et al 1999),(Kimura and Yachida, 1997),(Lisetti and Rumelhart, 1998). This allows the classifier to learn the relevant features in the data. However, the required normalization step usually involves the whole image and is generally time consuming. Processing of all pixels in the image is computationally expensive and a large memory space is required. Different methods like SVM and neural networks (Lee, 2003), Hidden Markov Models, dense flow with principal component analysis (PCA) and the sum-of-squared-difference (SSD) (Lien J.J. et al, 1998), elastic net model (Kimura and Yachida, 1997) and Gabor features and artificial neural network (Tian et al, 2000) are used. Intensity is continuously scaled with a value between 0 and 1, for the least and biggest expression intensity, respectively.

Local Feature Approaches track the position of some facial features (such as eyes and mouth) and hypothesize that the relative motion of these features are related to the intensity of the expression(Hong et al, 1998),(Lien et al 1998),(Wang et al 1998). With a local feature method,

the number of input dimensions is significantly reduced, so does the computational complexity and the processing time. In this case, the reliability of the tracker becomes very important. There are only five different AUs (26, 27, 41, 42, 43) for which FACS provides an option to score intensity on a 3-level (or 5-level) intensity scale (low, medium, and high). These AUS are not sufficient to quantify an emotion because these AUs are not the only activated units with all facial expressions. Different Artificial Intelligent (AI) techniques and non-AI techniques were applied to recognize facial actions and their intensity. A Rule-based expert system is applied to encode and quantify the encountered facial actions from the extracted facial expression information (Pantic and Rothkrantz, 2000). The problem with these methods is that the quantification of action units deviates in average of 8% from that done by the FACS coders (Pantic and Rothkrantz, 2000). Another problem is that not all universal expressions are quantified only Joy, Anger, Sadness, Surprise have been considered (Table 2) summarizes classification and quantification methods and their differences with respect to the method we propose in the paper.

Table2. Classification and quantification methods and the difference with the proposed method.

State of the Art Methods	Images	Categories	Accuracy
Classification Methods			
Linear Discriminant Analysis with the Mahalonobis distance measure (Edwards, G.J. and al 1998)	Set of training facial images 25 subjects	Six universal expressions & neutral	74%
Elastic Graph Matching with the personalized galleries (Hong, H. and al, 1998)	Static images from Live video sequence 25 subjects	Six universal expressions	89% -> familiar subjects; 73% -> unknown persons.
Principal Component Analysis with action parameters (Huang C.L and al, 1997)	Static images 9 subjects	Six universal expressions	84,5%
Hopfield neural networks with optical flow (Yoneyama, M. and al, 1997)	Static images	sadness , surprise, anger, and happiness	92,2%
Back-propagation neural network. ((Zhang, Z. and al, 1998), (Kobayashi H. and al, 1997), (Padgett C. and al, 1996), (Zhao J. and al, 1996))	Static images 9 to 15 subjects	Six universal expressions	85% till 100%
Rule-based approach to expression classification	Static images 8 subjects	Recognition Upper AUS and lower AUS	92% 86%

(Pantic, M. and al, 2000)			
Quantification Methods			
SVM + Neural Networks (K.K Lee , 2003)			95.0%, 76.7% and 70.6% for joy, anger and sadness
Dense flow with principal component analysis (PCA), & the sum-of-squared-difference (SSD) (J.J Lien and al, 1998)	Facial expression sequence	integer with values between 0 and 1, for the least and biggest intensities	
Elastic Net Model (S.Kimura and al, 1997)			happiness, anger, surprise
Gabor features and artificial neural network (Y.L Tian and al, 2000)	image sequences of 12 subjects	eye closure AU41, AU42, and AU43 Like FACS	83% is obtained for 112 images from 17 image sequences
AI techniques and non-AI techniques (M. Pantic and al, 2000)	static dual-view images	30 facial actions	90% deviates in average for 8% from that done by the FACS coders
Proposed Method = Classification + Quantification			
Belief theory	Static Images from: Hammal_C (21sub) ; EEbase (42 sub) and Dafex Databases(8 sub).	Three scales: Low , Medium and High intensities	Recognition of six universal expressions + neutral and their intensity

[Table 2]

3. Proposed Approach

3.1 Architecture

An automatic facial expression system is generally made of three steps: face detection, features extraction, and classification. In order to propose a full automatic system, we use the Rowly face detector (Rowley et al, 1998) to detect faces and methods proposed in (Eveno et al, 2001) and (Hammal et al, 2006) to detect permanent facial features.

The architecture of our system is shown in (Figure 1). When an expressive face is presented to the system, permanent facial features are firstly extracted. A reference face image which is an image of the considered face with a neutral expression is needed for facial expression analysis. The neutral image is acquired only one time for all. Characteristic points which represent feature corners (Figure 2) are then extracted from detected facial permanent features. These points allow locating facial transient feature regions and allow computing biometric facial distances. Transient features are detected in facial regions and the nasolabial angle (if any) is calculated.

Based on both permanent and transient features, the classification step is performed in parallel with the quantification one to give at the end the class of the studied expression and its intensity.

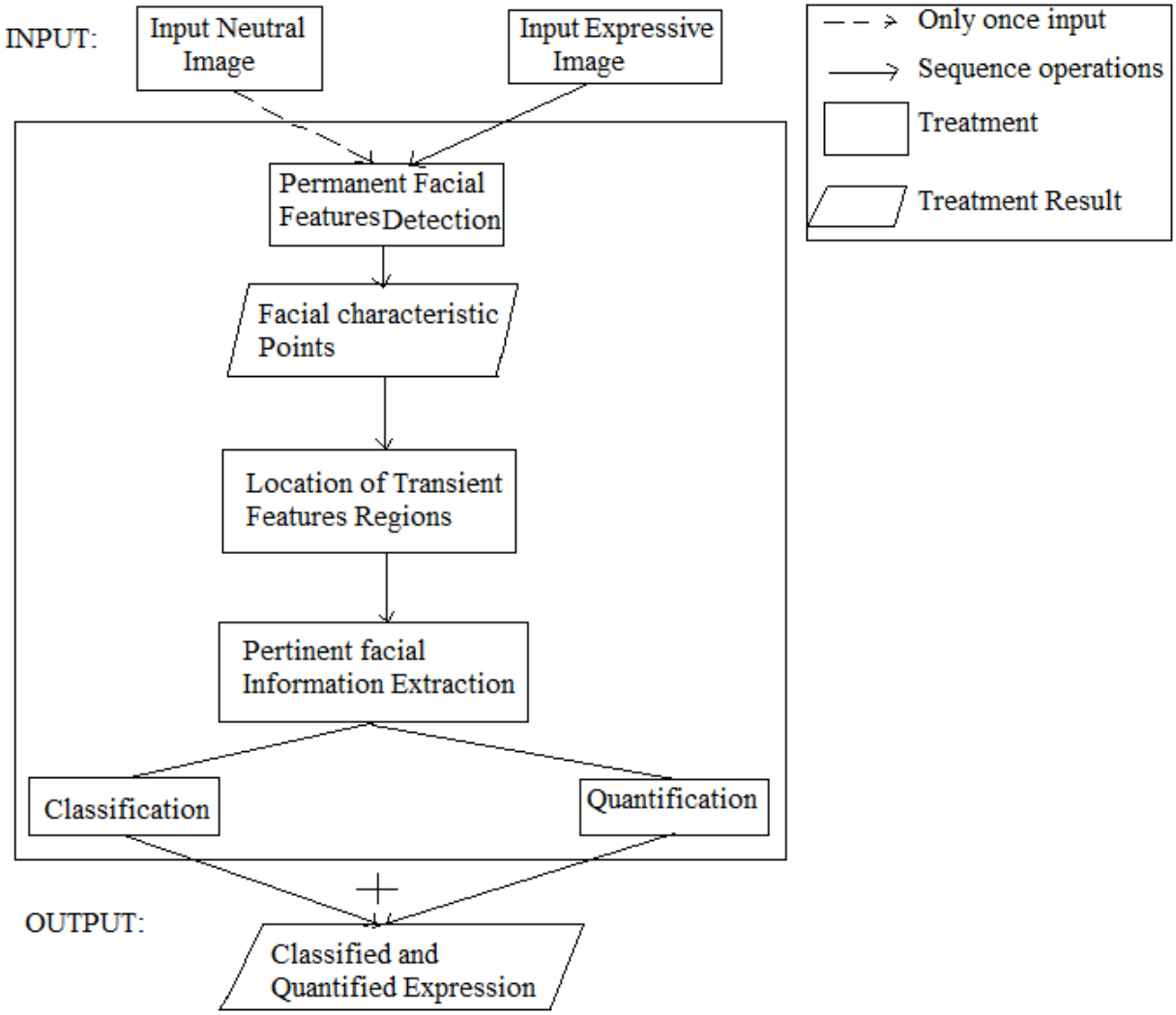


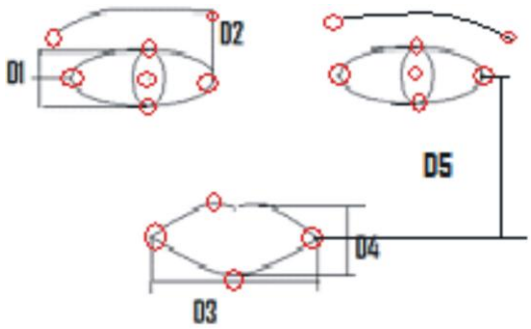
Figure 1: Flowchart of the proposed method. (Page 8).

[Figure 1]

3.2 Permanent Facial Features Segmentation and Transient Features Regions Extraction

At the beginning, the face is detected by the Rowly detector proposed in (Rowley, H.A. et al, 1998). This detector allows fitting a face. Then, permanent features are detected by (Eveno et al, 2001) and (Hammal et al, 2006) algorithms. From these extracted features, 18 characteristic points are defined. These characteristic points are mouth corners, bottom and top of lips, eye

corners, bottom and top of eyelids, the centers of irises and finally eye brows corners (Figure 3). Selected points allow calculating some biometric distances (Figure 2) as well as automatically determining regions of interest where transient features can appear. Nine regions are considered: Forehead, the two regions surrounding eye corners, the nasal root, the two nasolabial regions, the two regions surrounding the mouth, and finally the chin. Before computing region's dimensions, a coefficient (coef) which makes the analysis independent on the variability of face dimensions is calculated. These regions are delimited by characteristic facial points and face limits deduced by Rowly detector (some times formulated by using the coefficient (coef)).



D1: Eye opening distance between upper and lower eyelids
D2: the distance between the interior corner of eye and the interior corner of eyebrow
D3: Mouth opening width, the distance between left and right mouth corners
D4: Mouth opening height , the distance between upper and lower lips,
D5: Distance between eye and mouth corners.

Figure 2 : Characteristic points and facial distances. (Page9).

[Figure 2]

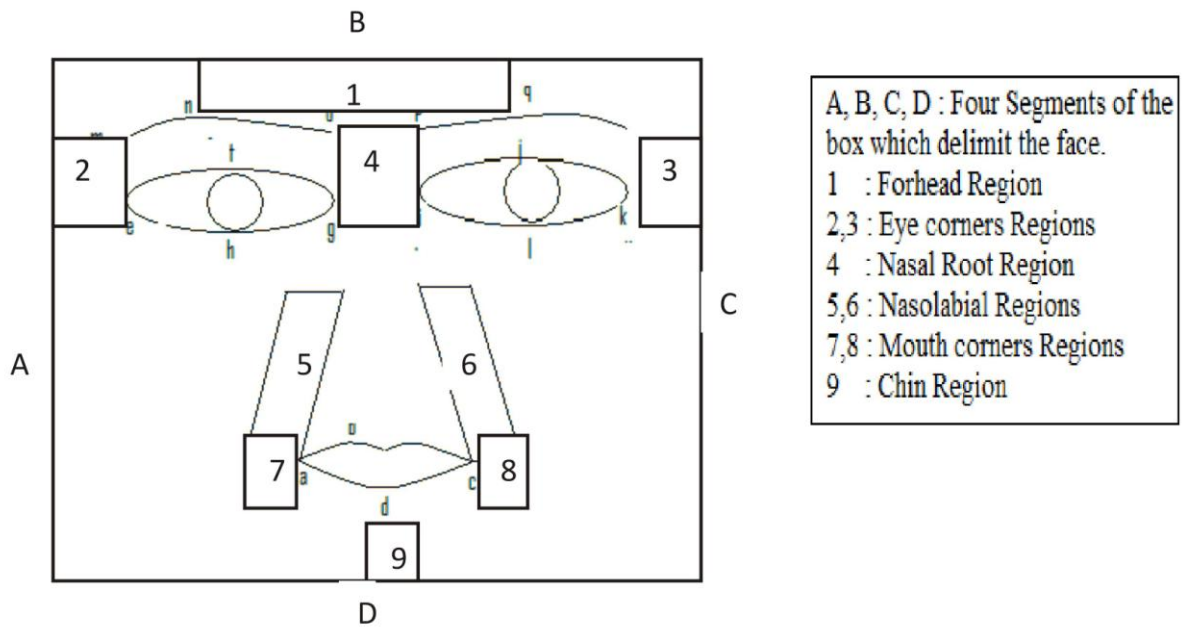


Figure 3 : Detection of Transient and Permanent features. (Page9).

[Figure 3]

3.3 Discriminating Facial information Extraction

Discriminating information is extracted from permanent features as well as from transient features. Concerning information coming from permanent features, five biometric facial distances are computed from the 18 characteristic points of (Figure 2). These distances are then normalized with respect to the distance between the two irises.

To extract information from transient features, a segmentation of the expressive facial image is performed to detect the presence or absence of transient features on the predefined facial regions of (Figure 3). A Canny edge detector (Lyons et al, 1999) associated to a thresholding stage (Yoneyama et al, 1997) is used to detect the presence or absence of such features. We compare the number of edge pixels in a wrinkle area in case of an expressive face with the number of edge pixels in the same area in case of a neutral face (Figure 4). If the ratio is larger than a high threshold (T_{high}), transient features are supposed to be present, and a state “Present” is assigned to the corresponding region. If the ratio is smaller than a low threshold (T_{low}), transient features

are supposed to be absent and the state “Absent” is associated to this region. Otherwise there is a doubt in their absence or presence and the state “Present OR Absent” is associated to it.

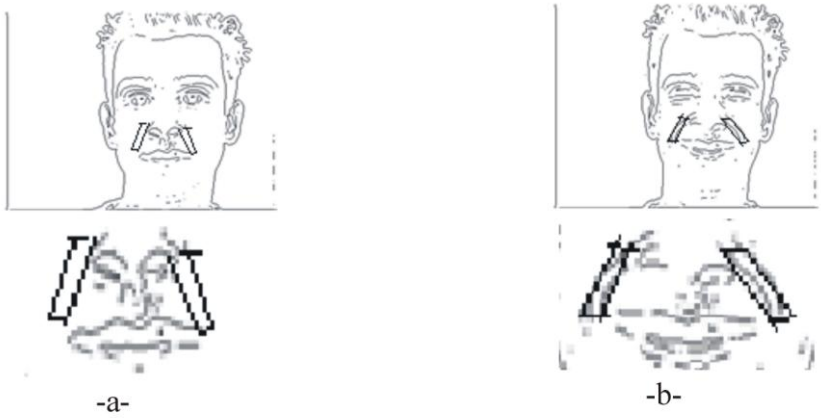


Figure 4 : (a):Neutral Face without TF on the nasolabial region; (Page10).

(b): Expressive Face with TF on the nasolabial region. (Page10).

[Figure 4a; Figure 4b]

Another information conveyed by the nasolabial furrows (if any) which is the angle between the line approximating the nasolabial furrow and the horizontal line connecting both mouth corners, is computed (Figure 5). This angle is normalized with respect to the angle 90° which represents the maximum possible angle reached by any angle formed with any facial expression.

Figure 5 : Detection of nasolabial furrows and computation of the nasolabial angle. (Page10).

[Figure 5]

Results of this step are: five normalized distances, normalized geometric angle (if any) and information about the presence or absence of transient features in each wrinkle region.

3.4 Classification

To classify the given expression in one of the six universal expressions, all extracted data are considered. To do so, the transferable belief model (TBM) is used. The TBM (Girondel et al, 2005), (Denoeux and Smets. 2006), (Mercier, 2006), (Ramasso et al 2007) salient character is the powerful combination operator that allows the integration of information from different sensors. In addition it is able to model intrinsic doubt which can occur between facial expressions in the recognition process (see Figure 6). Moreover, sometimes an emotion is not clearly expressed and then cannot be directly recognized.

Fig 6: Example of doubt between Surprise and Fear. (Page10).

[Figure 6]

3.4.1 Belief Theory principle (Dempster, 1968), (Shafer, 1976)

The Belief Theory is a generalization of the probability theory (Smets, 2000). It has been introduced by Dempster & Shafer and then by Smets. It requires the definition of a set Ω = {E1, E2,…,EN} of N exclusive and exhaustive assumptions. We also consider the power set 2^{Ω} that denotes the set of all subsets of Ω. To each element A of 2^{Ω} is associated an elementary piece of evidence *m(A)* which indicates the confidence that one can have in this proposal. The function *m* is defined as:

$$m: 2^{\Omega} \rightarrow [0,1] \qquad A \rightarrow m(A) \qquad \text{where : } \Sigma m(A)=1 \tag{1}$$

As we have several sources of information, we have to take into account all the available information. The global evidence is obtained using the rule of conjunctive combination or orthogonal sum. In the case of two assumptions, the orthogonal sum is defined in the following way:

$$m = m1 \oplus m2 \qquad m(A) = \sum_{B \cap C = A} m1(B).m2(C) \tag{2}$$

Where A, B and C denote propositions, B $\cap$ C denotes the conjunction (intersection) between the propositions B, C and m1, m2 are two evidences of two different sensors.

Finally, a decision can be taken about the classification. Several criteria can be used: the plausibility which favors the single hypotheses, the belief which favors the mixture of hypotheses and the pignistic probability which only deals with singleton propositions.

3.4.2 Application of the TBM to Facial Expression Classification

3.4.2.1 Definition of the global set of discernment. In our application, the set Ω corresponds to the six facial expressions: {*joy, surprise, disgust, anger, sadness and fear*}. 2^{Ω} corresponds to single expression or combinations of expressions and A is one of its elements.

3.4.2.2 Basic Belief Assignment.

3.4.2.2.1 Transient features information modeling. In our previous work (Ghanem and Caplier, 2008), a study of transient features was done on (Hammal_Caplier database) and (EEbase database), to determine which transient feature corresponds to which expression. (Table 3) reports the logical rules for each interest region:

Table 3. Presence or absence of transient features in each wrinkle region for each facial expression.

Transient features	JOY	SURPRISE	DISGUST	ANGER	SADNESS	FEAR
Chin(1 region)	(0)	(0)	(0)	(1OR0)	(1 OR 0)	(1 OR 0)
Mouth corners (2 Left & right regions)	(0)	(0)	(0)	(0)	(1 OR 0)	(1 OR 0)
Nasolabial (2 Left & right)	(1 OR 0)	(0)	(1 OR 0)	(1 OR 0)	(1 OR 0)	(1 OR 0)
Eyes corners (2 left & right eyes)	(1 OR 0)	(0)	(1 OR 0)	(1 OR 0)	(0)	(0)
Nasal root (1 region)	(0)	(0)	(1 OR 0)	(1 OR 0)	(1 OR 0)	(1 OR 0)
Forehead (1 region)	(0)	(1 OR 0)	(0)	(1 OR 0)	(1 OR 0)	(1 OR 0)

[Table 3]

Lines of (table 3) represent all considered wrinkle regions; columns represent the six universal expressions. ”1” is associated to the considered facial expression, for a wrinkle region if transient features can appear in this region, ”0” otherwise

The proposed model in (Figure 7) is used to compute the basic belief assignment associated to each transient feature region.

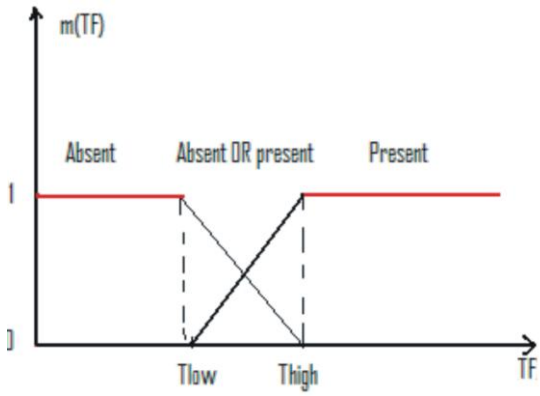


Figure 7: Model of basic belief assignment based for the presence of transient features. (Page12).

[Figure 7]

To formulate the joint Basic Belief Assignment in terms of facial expressions, we use the logical rules (see Table 3) which allows associating the piece of evidence of each symbolic state to the corresponding expression (expression for which the state is reached). The only considered state is the state “present”. For example: m_{chin}(present)=m_{chin}(*Anger* OR *Sadness* OR *Fear*). It means

that the piece of evidence (our belief) associated to the state “present” of the wrinkle region “chin” is equal to the piece of evidence of the expression *Anger* or *Sadness* or *Fear*.

Nasolabial angle is also used to discriminate between expressions. According to Ekman (Ekman et al, 1978), the formed angle with Anger or Disgust expression is due to AU9 or AU10, but the formed angle with *Joy* is due to AU12. Consequently the angle in case of *Anger* or *Disgust* is higher than the angle in case of *Joy* (Figure 8).

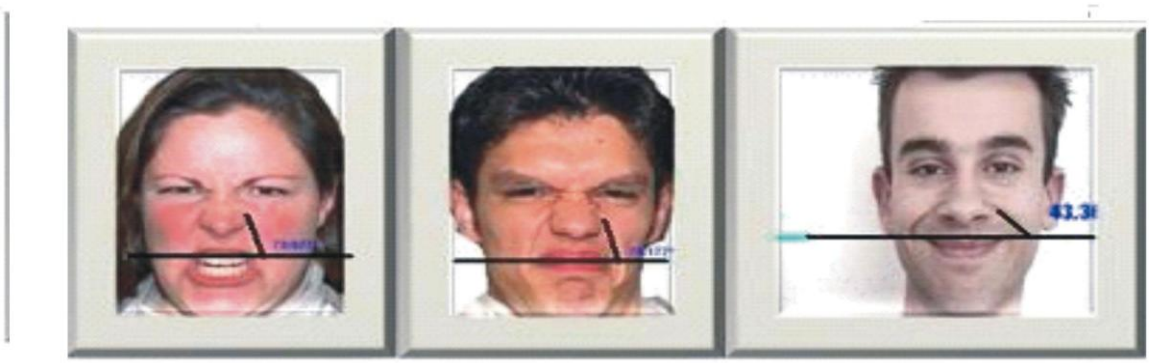
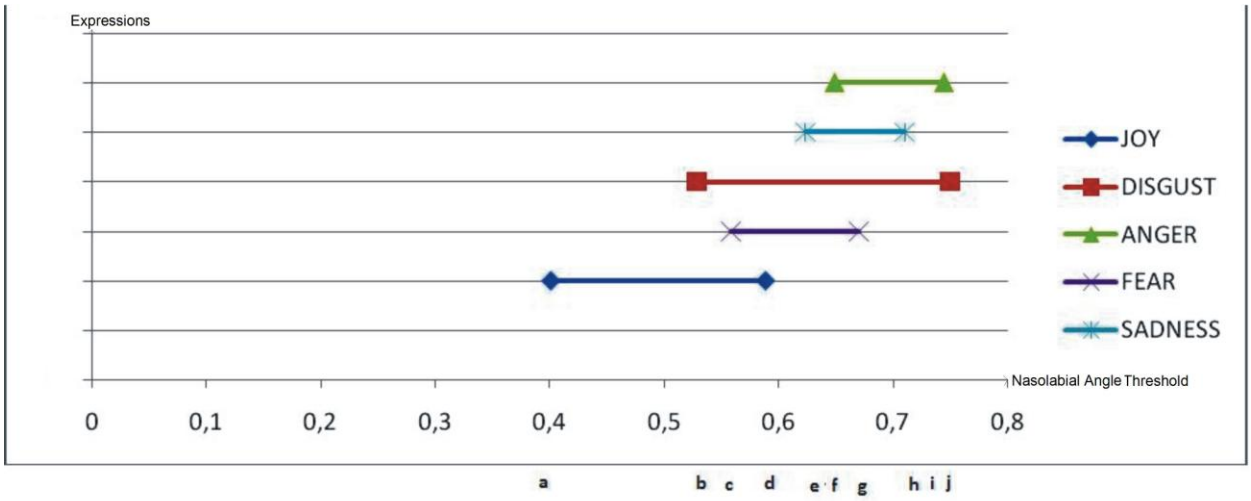


Figure 8: Angle formed by the nasolabial furrows with from left to right: *Anger* (72.6°), *Disgust* (71.2°) and *Joy* (43.4°) (Images coming from Eebase and H_Caplier databases). (Page12).

[Figure 8]

In order to exploit these findings, we measure the angle formed by the nasolabial furrows for 144 images with *Disgust*, *Anger*, *Fear*, *Sadness* and *Joy* expressions from two databases (EEbase) and (Dafex). *Surprise* is not considered here because with this expression we have no transient features on the nasolabial region (see table 3). Estimated angles are normalized with respect to the angle 90°. We estimated for each expression a nasolabial variation range. The threshold values (a, b, c, d, e, f, g, h, i, j) have been derived by statistical analysis on the (EEbase database) for each facial expression. The variation ranges are represented in (Figure 9).



a is the minimum threshold corresponding to *Joy*.
b is the minimum threshold corresponding to *Disgust*.
c is the minimum threshold corresponding to *Fear*.
d is the maximum threshold corresponding to *Joy*.
e is the minimum threshold corresponding to *Sadness*.
f is the minimum threshold corresponding to *Anger*.
g is the maximum threshold corresponding to *Fear*.
h is the maximum threshold corresponding to *Sadness*.
i is the maximum threshold corresponding to *Anger*.
j is the maximum threshold corresponding to *Disgust*.

Figure 9 : Nasolabial angle variation ranges. (Page13).

[Figure 9]

Deduced model in (Figure 10) is used to compute the basic belief assignment associated to normalized angles.

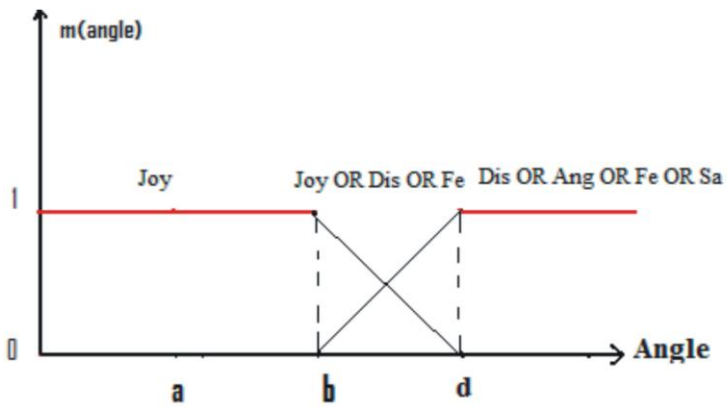


Figure 10 : Model of basic belief assignment for computed nasolabial angle. (Page13).

[Figure 10]

If the normalized angle is less than “b” the expression is *Joy* with a piece of evidence equal to “1” , if it is less than “d”, it can be *Joy* or *Disgust* or *Fear*, else it can be *Disgust*, *Fear*, *Anger* or *Sadness* with a piece of evidence equal to “1”.

3.4.2.2.2 Permanent features information modeling. In (Hammal et al, 2007) authors have modeled permanent features information by using the five specific distances defined in (Figure 2). In their work they proposed a description for each facial expression deduced from MPEG-4 description besides their own observations. Then they fused all available data associated to permanent features to classify the studied expression. In another way, different descriptions of facial expressions are proposed in the literature (Tekalp, 2000), (Carroll and Russell, 1997), (Eibel-Eihesfeldt, 1989), (Pardàs, 2000) and (Tsapatsoulis et al, 2000). In this work, our aim is to make a synthesis of all these descriptions in order to deduce the most common characteristics to each facial expression so a new universal description. And then the new considered distances are fused to transient features information (Nasolabial angle and presence of transient features on some facial regions) to classify the studied expression.

(Table 4) summarizes the evolution of each facial distance with each facial expression according to the different descriptions referenced before.

Table 4. Facial expression Descriptions

	Distance between lids D1	Distance between eye and eye brow D2	Distance between mouth corners D3	Distance between lips D4	Distance between eye corner and mouth corner D5
Joy	Decreases or Relaxes	Relaxes	Increases	Relaxes or Increases	Decreases
Surprise	Increases	Increases or Relaxes	Relaxes or Decreases	Relaxes or Increases	Relaxes or Increases
Disgust	Decreases	Decreases Or relaxes	Increases, relaxes or Decreases	Increases	Increases, relaxes or Decreases
Anger	Decreases or Increases	Decreases Or relaxes	Relaxes or Decreases	Increases, relaxes or Decreases	Relaxes or Increases
Sadness	Decreases Or relaxes	Increases	Relaxes or Increases	Relaxes or Increases	Relaxes or Decreases
Fear	Relaxes or	Relaxes or	Increases,	Relaxes or	Increases,

	Increases	Increases	relaxes or Decreases	Increases	relaxes or Decreases

[Table 4]

Our aim is to find a description of each facial expression by using a minimum of facial distances. To do so, we compared every two expressions in terms of distances. For example, the difference between *Joy* and *Surprise* or *Anger* is that the mouth is opened horizontally with *Joy* and it is opened vertically with *Anger* and *Surprise*.

With the other expressions, we can see that all distances can increase or decrease. For example, in the case of *Disgust*, the mouth can be opened horizontally (like with *Joy*), or can be opened vertically. This is why *Joy* can be easily distinguished from *Surprise* but not from *Disgust*.

In the same way, each expression is compared with all the other ones and the potential differences between expressions have been deduced (see Table 5).

Table 5. Potential differences between universal expressions in terms of facial distance evolution

Expressions	Differences
Joy versus surprise	Joy: D3 ↗ ; Surprise: D3 ↘ or = Joy: D1 ↘ or =; Surprise: D1 ↗
Joy versus anger	Joy: D3 ↗ ; Anger: D3 ↘ or =
Joy versus Sadness	Joy: D2 = ; Sadness: D2 ↗
Surprise Versus Disgust	Surprise: D1 ↗ ; Disgust: D1 ↘
Surprise versus Sadness	Surprise: D1 ↗ ; Sadness: D1 ↘ or =
Disgust versus Sadness	Disgust: D2 ↘ ; Sadness: D2 ↗ or =
Disgust versus Fear	Disgust: D1 ↘ ; Fear: D1 ↗ or =
Anger versus Sadness	Anger: D2 ↘ ; Sadness: D2 ↗ or =
Sadness versus Fear	Sadness: D1 ↘ ; Fear: D1 ↗ or =

[Table 5]

(Table 5) is used to describe expressions with a minimum of discriminating distances (D1, D2, D3). The two distances D4 and D5 are not used in this description because these two distances evolve in the same way with the six universal expressions. (Table 6) sums up these descriptions:

Table 6. New description of facial Expressions

Expressions	D1	D2	D3
Joy	Decreases or relaxes		Increases
Anger		Decreases	Decreases or relaxes
Surprise	Increases		Decreases or relaxes
Disgust	Decreases	Decreases	
Sadness	Decreases	Increases or relaxes	
Fear	Increases or relaxes		

[Table 6]

To associate a basic believe assignment to each considered distance, we use the model proposed in (Eveno et al, 2001). A variable Vi is associated to each distance Di in order to have an intermediate modeling between the numerical values of our characteristic distances and the required expressions. One model is defined for each characteristic distance independently of the facial expression.

Thresholds of each model are estimated by statistical analysis on a training database as done in (Eveno et al, 2001). The piece of evidence mDi(Vi) associated to each proposition given the characteristic distance Di is obtained by the function depicted in (Figure 11).

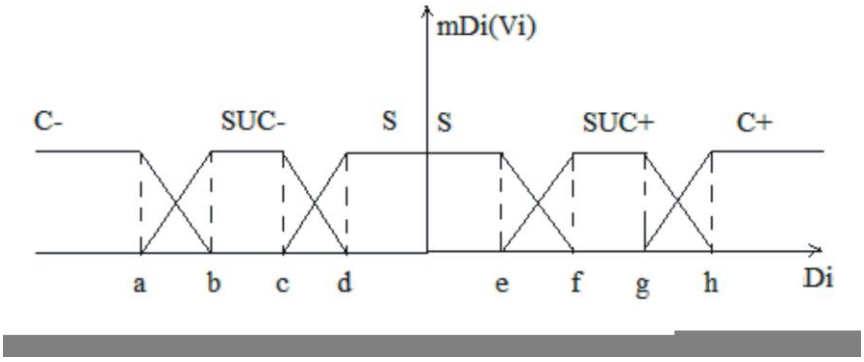


Figure 11: Model of basic belief assignment based on characteristic distance Di for the state variable Vi (Eveno N. and al, 2001). (Page14).

[Figure 11]

Logical rules of the distance D2 for example (Table 7) are used to formulate the basic belief assignment in terms of expression class. This table is deduced from (Table 3) and it allows associating the piece of evidence of each symbolic state to the corresponding class of expressions (class for which the state is reached) for all distances.

Table 7. Logical rules of symbolic states for characteristic distance D2 for each class of expressions.

D2->V2	Joy	Disgust	Anger	Sadness	Fear	Surprise
C+				1U0	1U0	1U0
S					1U0	
C-		1U0	1U0			

[Table 7]

(Table 7) can be interpreted as follows: if D2 decreases, then reached class corresponds to *Disgust* OR *Anger*, if it increases, the reached class is *Sadness*, *Fear* or *Surprise* and if it does not change, the reached class is *Fear*.

3.4.2.3 Data fusion. To take all available information (presence of transient features, nasolabial angle and Di distances) into account, the Dempster rule combination presented in section 3.4.1 is used.

First information about the presence of transient features on the different wrinkle regions is fused. For example, if transient features are detected in the chin wrinkle area and in the mouth wrinkle area, the associated BBAs using the model defined in (Figure 7) are:

$$m_{chin}(\textit{Anger} \text{ OR } \textit{Sadness} \text{ OR } \textit{Fear})=1 \text{ and } m_{mouth}(\textit{Sadness} \text{ OR } \textit{Fear})=1$$

The combination of these two BBAs is equal to:

$$m_{chin,mouth}(\text{Sadness OR Fear})= (m_{chin} \oplus m_{mouth})(\text{Sadness OR Fear})$$

$$=\sum m_{chin}(\text{Anger OR Sadness OR Fear})* m_{mouth}(\text{Sadness OR Fear})=1.$$

Next, this result is combined with information about the nasolabial angle using the model defined in (Figure 10). In the previous example, nasolabial furrows are not detected, this is why this information is not included.

Finally facial distances are joined to the fusion process. In the previous example, D1decreases and D2 increases so that:

mD1(*Disgust* OR *Sadness*)=1 and mD2(*Sadness* Or *Fear* OR *Surprise*)=1.

And the combination of the previous results with these two distances is equal to:

$m_{chin,mouth,D1}$(Sadness)=1 and $m_{chin,mouth,D1,D2}$(Sadness)=1.

The classification of the studied expression in this case is "*Sadness*" with the piece of evidence equal to "1".

3.4.2.4 Decision. The decision is the ultimate step of the classification process. It consists in making a choice between various assumptions and their possible combinations. In our case we choose the belief criterion to make decision because it allows the mixture of hypotheses and we need to model the doubt between propositions. So, if the result of the combination is different from 1, the accepted proposal is the one with maximum value of evidence mass.

3.5 Quantification process

The quantification process of the studied expression is done in parallel with the classification one. The two processes are independent and give at the end two results which are expression category and expression intensity.

Two different sources of data (facial distances and the nasolabial angle) are used in order to make a decision about expression intensity. The TBM is also used for the data combination and the expression intensity estimation

3.5.1 Definition of the global set of discernment

Here, the set Ω corresponds to three intensities: Ω = {ELow, EMedium, EHigh}.

2^{Ω} corresponds to single intensity expression or combinations of intensities and A is one of its elements.

3.5.2 Basic Belief Assignment

3.5.2.1 Permanent features information modeling. In order to quantify a facial expression by using permanent features, only the distances which have changed with the expressive face (from those of the neutral one), are considered. Those distances are compared with thresholds learnt in our precedent works (Ghanem and Caplier, 2008-2), (Ghanem and Caplier, 2008-3). The proposed model in (Ghanem and Caplier, 2008-2) is used. One model (Figure 12) is defined for each considered distance; a state variable Vi is associated to each distance Di (i=1..5) in order to convert the numerical value of the distance to a symbolic state. Vi can take three possible states: Low, Med or High which correspond to three score levels of intensity « ELow, EMedium and EHigh ». Each distance relevant to each expression is classified into one of the three levels or between two levels, with a piece of evidence (BBA: Basic Belief Assignment) associated to each level.

Figure 12: The proposed Model to map changed distances. (Page17).

[Figure 12]

To formulate the joint Basic Belief Assignment in terms of quantified expressions (expressions with intensities), we use (Table 8). This table allows associating a piece of evidence of each symbolic state to the corresponding quantified expression (expression with intensity for which the state is reached). For example: if the associated state to a distance Di is “low” the corresponding expression intensity is “ELow” and so on.

Table 8. Different states taken by a distance and its corresponding expression intensity.

Expression	ELow	EMedium	EHigh	ELowUEMedium	EMediumUEHigh
Vi	low	med	high	lowUmed	medUhigh

[Table 8]

To be more explicit, we consider two distances (D1 and D2), so that:

V1=med and V2=medUhigh ➔mD1(med)=mD1(EMedium);

mD2(medUhigh)=mD2(EMediumUEHigh)

By using the orthogonal sum to join the two distances, we get the associated expression intensity which is *Emedium*.

D1\D2	EMediumUEHigh
EMedium	EMedium

3.5.2.2 Transient features information modeling. The information conveyed by the nasolabial angle associated to a nasolabial wrinkle (if any) is also considered in the quantification process. To do so, we have to map its value in a state in the same way as distance values. We study the normalized angle calculated in section 3.4.2.2.1.

When studying images of (Dafex) and (EEbase) databases, we note that transient features cannot be detected on images with low intensity. When the intensity is medium or high, the calculated angle of *Joy* expression images with medium intensity is higher than the angle with high intensity, but with *Disgust*, *Anger*, *Fear* or *Sadness*, calculated angle of images with medium intensity is lower than the angle with high intensity.

To validate these observations, we estimated the angle formed by the nasolabial furrows for all images of (EEbase database). Calculated angles are normalized with respect to the angle 90 ° (the highest possible angle reached in case of the activation of AU9 or AU10) (Figure 13), (Figure 14). and (Figure 15).

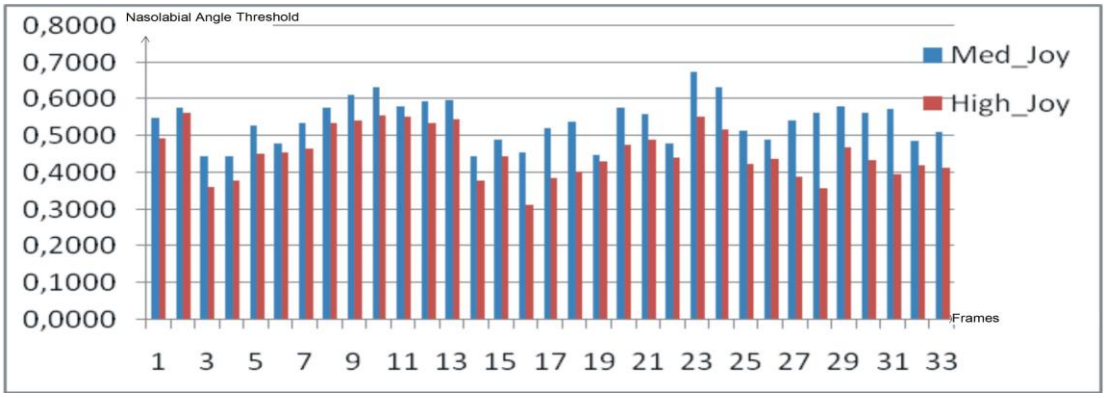
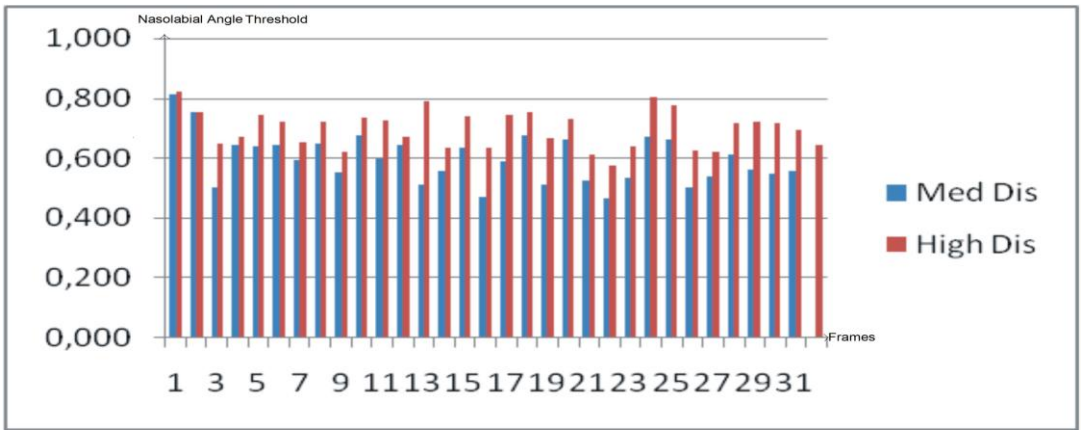


Figure 13: Nasolabial angles with *Joy* expression for medium and high intensities on the EEbase database. (Page18).

[Figure 13]

Figure 14: Nasolabial angles with *Disgust* expression for medium and high intensities on the EEbase database. (Page18).



[Figure 14]

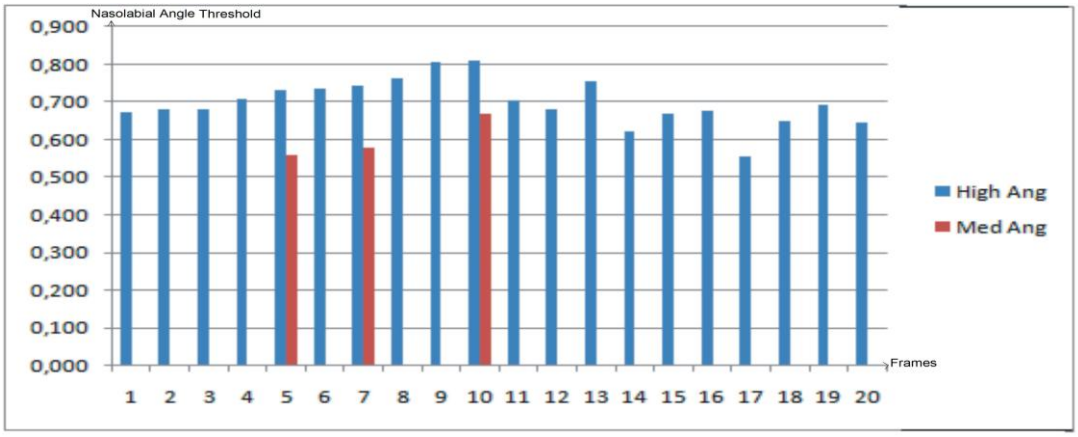


Figure 15: Nasolabial angles with *Anger* expression for medium and high intensities on the EEbase database. (Page18).

[Figure 15]

From (Figure 13), we can see that nasolabial angles formed with *Joy* expression and medium intensity are higher than those formed with high intensity. With (EEbase database), few images (three) with medium anger, present nasolabial angles so from (Figure 14) and (Figure 15), we can see that nasolabial angles formed with *Disgust* or *Anger* expressions and medium intensity are lower than those formed with high intensity. These findings confirm our precedent observations.

Next we calculated several thresholds for each intensity and each expression. The highest thresholds were computed as the average of the highest values corresponding to each expression and the lowest thresholds as the average of the lowest values. These thresholds are represented in (Figure 16).

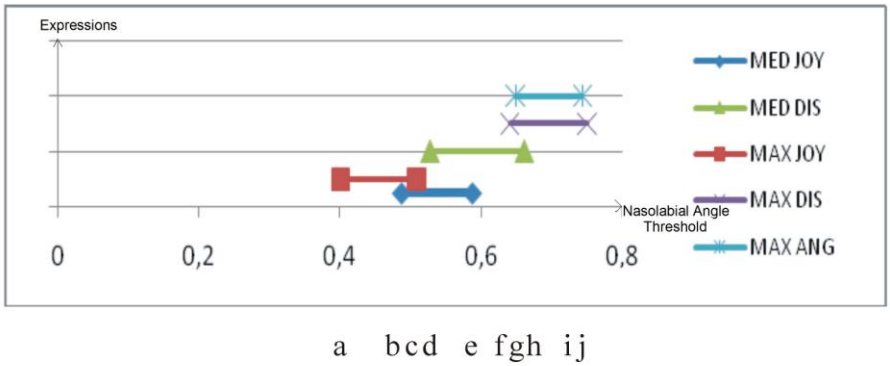


Figure 16: Thresholds representation for *Joy*, *Disgust* and *Anger* expressions. (Page18).

[Figure 16]

From (Figure 16) we can deduce intervals for medium, medium OR high and high intensities for expressions with nasolabial furrows. The considered intervals are given in (Table 9).

Table 9. Intensity classification according to nasolabial angle

	Medium intensity	Medium OR High intensities	High intensity
[a,b]			X
[b,c]		X	
[c,f]	X		
[f,h]		X	
[h,j]			X

[Table 9]

If the nasolabial angle is in the interval [a,b] or in the interval [h,j], the studied expression is an expression with high intensity. If the angle is in [b,c] or in [f,h] there is a doubt between medium and high intensities. If the angle is in [c,f], the intensity is medium.

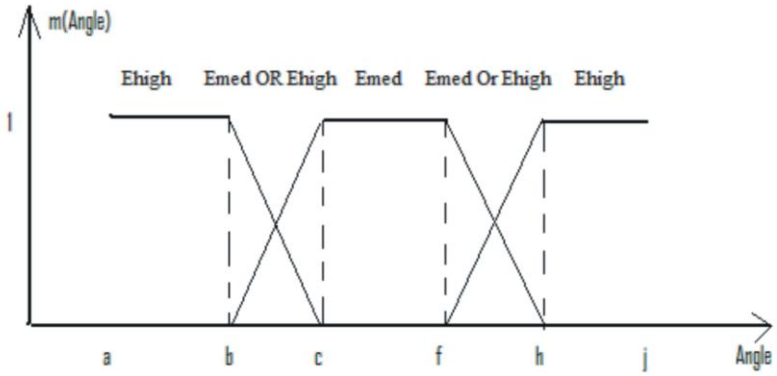


Figure 17: Model to map nasolabial angle. (Page19).

[Figure 17]

The model presented in (Figure 17) is deduced from the angle study. It is defined to map the angle value in one of the possible discrete states. The piece of evidence m(Angle) associated to each proposition is depicted in (Figure 17).

3.5.3 Data fusion and Decision

To make a decision about expression intensity, the global belief (which correspond to the fusion of distances basic belief and nasolabial angle basic belief (if any)) is computed by using the Dempster combination law presented in section 3.4.1.

For example, if we have two changed distances D3 and D5 and if nasolabial furrows are detected:

$$m_{D3}(med)= m_{D3}(Emed)=1 \ , \ m_{D5}(med)= m_{D5}(Emed)=1 \text{ and } m_{Angle}(Emed \text{ OR } Ehigh)=1$$

The combination of the two BBas of the two distances is equal to:

$$m_{D3,D5}(Emed)= (m_{D3} \oplus m_{D5})(Emed)=\sum m_{D3}(Emed)* m_{D5}(Emed)=1$$

Then, the combination of the two distances with the nasolabial angle is equal to:

$$m_{D3,D5,Angle}(Emed)= (m_{D3,D5} \oplus m_{Angle})(Emed)=\sum m_{D3,D5}(Emed)* m_{Angle}(Emed \text{ Or } EHigh)=1$$

With this example the intensity of the studied expression is “medium”. If the global piece of the evidence is different from “1”, the selected proposal is the one with the maximum value of the piece of evidence.

4. Experimental Results

The (Dafex database) has been used to test and validate the proposed system. This database is made of 1008 short videos containing emotional facial expressions of the 6 Ekman’s emotions with three different intensity levels (high, medium, low) plus the neutral expression. The facial expressions were recorded by 8 professional actors (males and females) in two acting conditions (“utterance” and “non utterance”).

Table 10. Classification of facial expressions with intensities

EXPERT / SYSTEM	Joy low (Well being)	Joy_ Med (Joy)	Joy_ High (Happiness)	Dis low (Feel seek)	Dis Med (Disgust)	Dis High (Bitterness)	Ang low (Boredom)	Ang Med (Anger)	Ang_ High (Rage)
Joy low = Well being									
Joy low OR med = Well being or Joy									
Joy med = Joy		37,5%							
Joy med OR high = Joy or happiness		12,5%	12,5%						
Joy high = Happiness			75%						
(JoyORDis) low	70%			37,5%					
(Joy Or Dis) low OR med	30%								
(JoyORDis) med		25%							
(JoyOr Dis med) Or high		12,5%							
(JoyORDis) high									
(AngORDis) low				62,5%			75%		
(AngORDis) low OR med					62,5%		25%		
(AngORDis) med					37,5%			62,5%	
(AngORDis) med OR high						12,5%		37,5%	
(AngORDis) high						87,5%			85,7%
Error		12,5%	12,5%						14,3%
Total Recognized	**100%**	**87,5**	**87,5**	**100**	**100**	**100**	**100**	**100**	**85,7**
Total	100	100	100	100	100	100	100	100	100

EXPERT / SYSTEM	Sad low (Trouble)	Sad_ Med (Sadness)	Sad_ High (Abatement)	Fear Low (Anxiety)	Fear Med (Fear)	Fear High (Terror)	Surp low (Astonished)	Surp Med (surprise)	Surp_ High (Amazement)
Sadness low = Trouble	50%								
Sadness low OR med =Trouble or Sadness	50%								
Sadness med = Sadness		75%							
Sadness med OR high = Sadness or Abatement		25%							
Sadness high = Abatement			100%						
Fear low = Anxiety									
Fear low OR med = Anxiety or Fear									
Fear med = Fear					50%				
Fear med OR high = Fear or Terror					12,5%				
Fear high = Terror						50%			
FearORSurp low				37,75%			25%		
FearORSurp low ORmed				62,5%			75%		
FearORSurp med					12,5%			75%	
FearORSurp medORhi					25%	12,5%		25%	37,5%
FearORSurp high						37,5%			62,5%
Error									
Total Recognized									
Total		100	100	100	100	100	100	100	100

[Table 10]

Classification rates are presented on (table 10) which is divided in two parts because of the number of columns. Columns represent the six universal expressions with three intensities for each expression labeled by an expert and lines represent possible classification combinations between expressions and intensities done by the proposed approach. Cases with possible situations which are not recognized by our approach are ignored.

From (Table 10), we can see that *Joy* can be confused with *Disgust*, that *Disgust* can be confused with *Anger* and that *Surprise* is generally confused with *Fear*. *Sadness* is the only expression which is not confused with any other one.

We note that important rates are given to the doubt between intensities (low-medium, and medium-high) as well as between expressions (Joy-Disgust, Disgust-Anger and Fear-Surprise), we feel that it is better to keep the doubt between two sensors instead of taking the risk of choosing the wrong one. The TBM is actually well adapted for such scenario.

We also observe that, when we have low intensity (labeled by an expert), the system gives as an answer doubt between medium and low (never between low and high), and when we have high intensity, the system gives doubt between medium and high intensities, but if we have medium intensity, the system gives doubt between high and medium or between low and medium.

We can see that expressions with high intensities are well recognized. Consequently, the best recognition rates are obtained with high intensity. This can be explained by the fact emotion with high intensity provide facial deformations with markedly visible deformations which are more easily detected.

5. Conclusion

In this work we have proposed to build a complete emotional system which recognizes an expression and in parallel, estimates its intensity. The system is based on facial deformations of permanent and transient features. The Transferable Belief Model is evaluated in expression classification as well as in intensity estimation. The TBM was evaluated in the case of the fusion of data coming from different sensors (geometric, probabilistic), in the case of imprecise data obtained from automatic segmentation and in modeling doubt.

The other main objective of this work is the set of recognized expressions which is enriched by adding a new set of expressions namely: (Well being, joy, happiness, Feel sick, Disgust, Bitterness, Astonished, Surprise, Amazement, Boredom, Anger, Rage, Trouble, Sadness, Abatement, Anxiety, Fear and Terror). We get eighteen expressions instead of the six universal ones. This finding also contributes in building a “rule expert system” in a future work.

As perspectives, we hope to consider data from temporal information introduced by video sequences which allow studying the dynamic of facial deformations.

References:

Carroll J.M, J.Russell, 1997. Facial Expression in hollywood's Portrayal of Emotion. *J.Personality and social psychology*, 72, 164-176.

Chandrasiri N. P., M.C. Park and T. Naemura and H. Harashima, 1999. Personal Facial Expression Space based on Multi-dimensional Scaling for the Recognition Improvement. Proc.*IEEE ISSPA'99*, Brisbane, Australia, 943-946.

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

Dafex Database: Available from: http://tcc.itc.it/research.i3p/dafex/index.html [Accessed November 2007]

Dempster A. 1968. A generalization of Bayesian inference. *Journal of the royal statistical society*, 30, 205-245.

Denoeux T. and Ph. Smets. 2006. Classification using belief functions : The relationship between the case-based and model-based approaches. *IEEE Trans. on Systems, Man and Cybernetics*. ISSN 1083-4419. 36 (6), 1395-1406

Edwards G.J., T.F. Cootes, and C.J. Taylor, 1998. Face Recognition Using Active Appearance Models. *Proc. European Conf. Computer Vision*, 2, 581-695.

EEBase Database: Available from : http://www.cs.cmu.edu/afs/cs/project/cil/ftp/html/v-images.html [Accessed November 2006]

Eibel-Eihesfeldt I., 1989. Human Ethology. New work: Aldine de Gruvter.

Ekman P. And W.V. Friesen, 1975. Unmasking The Face: A Guide to Recognizing Emotions from Facial Clues. New Jersey: Prentice Hall. Reprint Edition, 2003 Cambridge.

Ekman, P. 1980. Facial asymmetry. *Science*, 209, 833-834.

Ekman P. Wallace V. Friesen, 1978. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Palo Alto, CA: Consulting Psychologists Press.

Eveno N., Caplier A., Coulon P. Y 2001. A New Color Transformation for Lip Segmentation. in Proc. IEEE *4th Workshop on Multimedia Signal Processing*, Cannes, France, 3–8.

Fasel B., Juergen Luettin, 2003. Automatic facial expression analysis: a survey. *Journal of pattern recognition society*, 36 (1), 259 – 275.

Ghanem K., A.Caplier. 2008. Classification of facial Expressions based on Transient Features. *3rd Workshop on Emotion and Computing: Current Research and Future Impact, Kaiserslautern, Germany*. ISSN: 1865-6374, 24-32.

Ghanem K., A.Caplier, 2008-2. Estimation of Facial Expression Intensity Based on the Belief Theory. *VISAPP 2008: Proc. the Third International Conference on Computer Vision Theory and Applications*, Funchal, Madeira,Portugal, 1, 452-460.

Ghanem K., A.Caplier, 2008-3. Estimation of Anger, Sadness and fear expressions intensity based on the Belief theory. ACIT *2008: Proc. 9th International Arab Conference on information Technology*, Hammamet, Tunisia.

Girondel V., A. Caplier, L. Bonnaud, and M. Rombaut, 2005. Belief theory-based classifiers comparison for static human body postures recognition in video. *Int. Jour. of Signal Processing*, 2(1), 29–33.

Hammal_Caplier data base : Available from : http://www.lis.inpg.fr/pages_perso/caplier/english/emotionnelle.html.en/emotionnelle_2.html.en.html [Accessed September 2006]

Hammal Z., N. Eveno, A. Caplier, P.Y Coulon, 2006. Parametric Models for Facial Features segmentation. *Signal processing*, 86, 399-413.

Hammal Z., L. Couvreur, A. Caplier, M. Rombaut 2007. Facial Expression classification: An approach based on the fusion of facial deformations using the transferable belief model.

International Journal of Approximate Reasonning, 46 (3), Elsevier editor, doi: 10.1016/j.ijar.2007.02.003.

Hong H., H. Neven and C. Y. d. Mahurg, 1998. Online Facial Expression Recognition based on Personalized Gallery. *Proc. int . Conf. on Automatic Face and Gesture Recognition*, Nara, Japan, 354-359.

Huang C.L. and Y.M. Huang, 1997. Facial Expression Recognition Using Model-Based Feature Extraction and Action parameters Classification. *J. Visual Comm. and Image Representation*, 8 (3), 278-290.

Kimura S. and M. Yachida,1997. Facial Expression Recognition and its Degree Estimation. *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, 295-300.

Kobayashi H. and F. Hara, 1997. Facial Interaction between Animated 3D Face Robot and Human Beings. *Proc. Int'l Conf. Systems, Man, Cybernetics*, 3,732-3,737.

Lee Ka Keung, 2003 "Real-time Estimation of Facial Expression Intensity", proceedings of the *2003 IEEE International Conference on Robotics & Automation*, Taipei, Taiwan, 14- 19.

Lien J.J., T. Kanade, J.F. Cohn and Ching-Chung Li, 1998. Subtly Different facial Expression Recognition and Expression Intensity Estimation. *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 853-859.

Lisetti C. and D. Rumelhart, 1998. Facial Expression Recognition Using a Neural Network. *Proc. the 11 th Int. Conf. Facial Expression Recognition*, AAAI Press.

Lyons M.J., J. Budynek, and S. Akamatsu, 1999. Automatic Classification of Single Facial Images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21 (12), 1,357-1,362.

Mercier. D. 2006. Information Fusion for automatic recognition of postal addresses with belief functions theory. PhD thesis, University of Technologies of Compiegne.

Padgett C. and G.W. Cottrell, 1996. Representing Face Images for Emotion Classification. *Proc. Conf. Advances in Neural Information Processing Systems*, 894-900.

Pantic M. and L.J.M. Rothkrantz, 2000. Automatic Analysis of Facial Expressions: The State of the Art. *IEEE transactions on pattern analysis and machine intelligence*, 22 (12), 1424 - 1445.

Pantic M., 2000. An Expert System for Recognition of Facial Actions and Their Intensity. *Proceedings of AAAI Int'l Conf. Applications of AI (IAAI'00).* Austin, USA, 1026 - 1033.

Pardàs M., 2000. Extraction and Tracking of the Eyelids. *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Istambul, Turkey, 4, 2357-2360.

Ramasso E., C. Panagiotakis, M. Rombaut, and D. Pellerin. 2007. Human action recognition in videos based on the transferable belief model - application to athletics jumps. *Journal of Pattern Analysis and Applications*, 11 (1), 1-19, DOI: 10.1007/s10044-007-0073-y.

Rowley H.A., S. Baluja, and T. Kanade, 1998. Neural Network-Based Face Detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20 (1), 23-38.

Shafer G. 1976. A Mathematical Theory of Evidence. *Princeton, NJ, Princeton University Press.*

Smets PH., 2000. Data Fusion in the Transferable Belief Model. *Proceedings of the 3rd International Conference on Information Fusion, FUSION'2000*, France, PS21–PS33.

Tekalp M. 2000. Face and 2d mesh animation in mpeg-4. *In Signal Processing: Image Communication*, 15, 387-421.

Tian Y.-L., 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.

Tsapatsoulis N., K. Karpouzis, G. Stamou, F. Piat, S.A Kollias, 2000. A fuzzy system for emotion classification based on the MPEG-4 facial definition parameter set. *Proc. 10th European Signal Processing Conference*, Tampere, Finland.

Wang M., Y. lwai and M. Yachida, 1998. Recognizing Degree of Continuous Facial Expression Change", *Proc. Fourteenth Int. Conf. On Pattern Recognition*, 2, 1188-1190.

Yoneyama M., Y. Iwano, A. Ohtake, and K. Shirai, 1997. Facial Expressions Recognition Using Discrete Hopfield Neural Networks. *Proc. Int'l Conf. Information Processing,* 3, 117-120.

Zhang Z., M. Lyons, M. Schuster, and S. Akamatsu, 1998. Comparison between Geometry-Based and Gabor Wavelets-Based Facial Expression Recognition Using Multi-Layer Perceptron. *Proc. Int'l Conf. Automatic Face and Gesture Recognition,* 454-459.

Zhao J. and G. Kearney, 1996. Classifying Facial Emotions by Backpropagation Neural Networks with Fuzzy Inputs. *Proc. Conf. Neural Information Processing*, 1, 454-457.

Tables:

List of Figures: