We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Bio-Inspired Architecture for Clustering into Natural and Non-Natural Facial Expressions

Claudio Castellanos Sánchez, Manuel Hernández Hernández
and Pedro Luis Sánchez Orellana

Additional information is available at the end of the chapter

1. Introduction

Over the past years there has been an increasing surge of interest in automated facial expression analysis but there are several situations where the analysis of the expressions cannot be successfully done with the different classical approaches due to the variations in the environmental conditions (like illumination changes).

A way to face these variations in the environmental conditions is by looking at the existing systems that efficiently locate, extract and analyse the face information. An example of these systems is the brain of the human being.

According to several researchers [2, 27, 36], emotion processing is done in the amygdala, an area located within medial temporal lobes in the brain. Experiments carried out by [35] demonstrated that face recognition process can be achieved with the simple facial features (eyebrows and eyes vs eyebrows) of threatening and friendly faces. These experiments confirm that emotion processing involves other areas like the visual cortex, e.g. V1 [32]. Later, [17] demonstrated that fast, automatic, and parallel processing of unattended emotional faces provide important insights into the specific and dissociated neural pathways in emotion and face perception.

Although, there are different areas involved in the analysis of facial expressions in the brain of the human being, there is no consensus about how this process is done.

Clues from the psychological point of view, like that proposed by Ekman and Friesen [9], point out the importance of the symmetry in this process and the facial asymmetry (with left part of the face stronger than that of right part) is apparent only with deliberate and non-spontaneous (non-natural) expressions. According to this work, the symmetrical expressions show a true or natural emotion. Nevertheless, the asymmetrical expressions show a false or non-natural emotion.

In this research, we take advantage of psychological point of view to interpret our results issued by symmetrical and asymmetrical face measures when we model the visual areas of

brain of the human being. Furthermore, we are inspired by early visual areas for processing the visual stream. Then, the interaction between neurons of visual cortex lead us to introduce our proposed architecture based on the detection of asymmetrical facial expressions and their clustering into natural/non-natural expressions.

In this chapter, we will present an architecture composed of five stages in order to cluster into natural/non-natural facial expressions in a sequence of images.

The first stage process a sequence of images for eyes and mouth corners detection inspired by the sensibility of cone-cells in the eyes. With these corners we build an anthropometrical grid of six vertical bands.

The second stage is a bio-inspired treatment through three steps:

- Contrast and orientation detection inspired in the simple neurons in V1 (primary visual cortex) that are modelled by Gabor-like filters.
- Integration and smoothing contours due to the performance of complex neurons in V1.
- The coherence and integration activity of neurons in MT (middle temporal area) inspire us to detect the motion through a temporal interaction of our consecutive complex neurons.

The third stage computes the ratio between the quantity of active neurons and the maximum quantity of neurons in each of six bands built in the first stage. A process inspired in the visual stream integration in IT (infero-temporal area).

The fourth stage compares each symmetrical band with an empirical threshold κ obtained in our experiments.

The end stage clusters in natural/non-natural expressions. This stage uses the results of the fourth stage with a bio-inspired neural network model.

In the following sections, we will present an overview of the state of the art in research of facial expressions. In addition, we present our proposed architecture to analyse the asymmetrical responses of the faces from a sequence of images. We also mention the experiments, discuss the results and finally comment on the future work.

2. State of the art

The analysis of facial expressions constitute a critical and complex part of our non-verbal social interactions. Therefore, over the past years there has been an increasing surge of interest to create tools that allow to analyse the facial expression automatically. This analysis is a complex task; on one hand facial expression structure of human beings are different from one another. On the other hand, faces are analysed from sequences of images captured in different environments with variations of lighting, pose and scale changes that difficult the extraction of their structure. Traditionally, the analysis of facial expressions has been focused on identify emotions; however, both deformation and face information can be originated by other factors (like verbal communication or fatigue) rather than emotions. The emotions can be evaluated from two main perspectives: psychological, physiological and neurophysiological points of view versus computational point of view.

2.1. Psychological and physiological perspective

From the psychological and physiological perspective, the emotions can be analysed considering the symmetry of the expressions. According to Ekman and Friesen [9] through the analysis of symmetry it can be evaluated the veracity (naturalness) of an emotion. In the brain, the information about emotions is processed in the amygdala [2, 27, 36], an area located within medial temporal lobes in the brain. By using fMRI some researchers [1, 13, 32] have discovered that in this area the processing is modulated by the significance of faces, particularly with fearful expressions, also with others like gaze direction. Visual search paradigms have provided evidence the enhanced capture of attention by threatening faces. Experiments carried out by [35] demonstrated that the recognition process can be achieved with simple facial features (expressions of eyebrows and eyes together vs only eyebrows) of threatening and friendly faces. These experiments confirm that emotion processing involves other areas like the visual cortex, e.g. V1 [32].

2.2. Neurophysiological perspective

According to neurophysiological studies [15, 16], the task of facial expression recognition also occurs in some regions in the visual cortex (ventral and dorsal pathways). The neurons in these areas respond mainly to gesticulations and identity. In the case of gesticulations, the neurons in the superior temporal sulcus involve the dynamics of the face. In the case of identity, the neurons in the inferior temporal gyrus imply the analysis of invariant facial features. In both cases the processing is modulated by attentional mechanisms [37]. Later, [17] demonstrated that fast, automatic, and parallel processing of unattended emotional faces, provides important insights into the specific and dissociated neural pathways in emotion and face perception.

In the figure 1 we show the main pathways of processing visual stream in the cortex of the human being. The dorsal pathways analyse the motion and localisation of information, while ventral pathways analyse information about form and colour. Both pathways have the same source, V1. The visual cortex is the last specialised layer of evolution in mammals; the limbic system (not shown in the figure) is the first and primitive layer in the mammals where the amygdala is situated in the temporal lobes. Then, it is covered by visual cortex.

2.3. Computational perspective

From the computational point of view, facial expression recognition implies the classification of facial motion and facial structure deformation into abstract representations completely based on visual information. To address this analysis it is necessary to perform two main tasks: face detection and facial feature extraction. In the first one, techniques have been developed to deal with the extraction and normalization of the face considering variations in the pose [10, 28] and illumination [3, 12]. However, according to [11] the main effort has been focused to feature facial extraction where techniques like appearance-based models [6, 18, 19] work with significant variations in the acquired face images without normalization.

Basically, the facial feature extraction has been approached by two main streams [11, 23]: facial deformation extraction models and facial motion extraction models. Motion extraction approaches focus on facial changes produced by facial expressions, whereas deformation-based methods contrast the visual information against face models to extract features produced by expressions and not by age deformation like wrinkles. The main
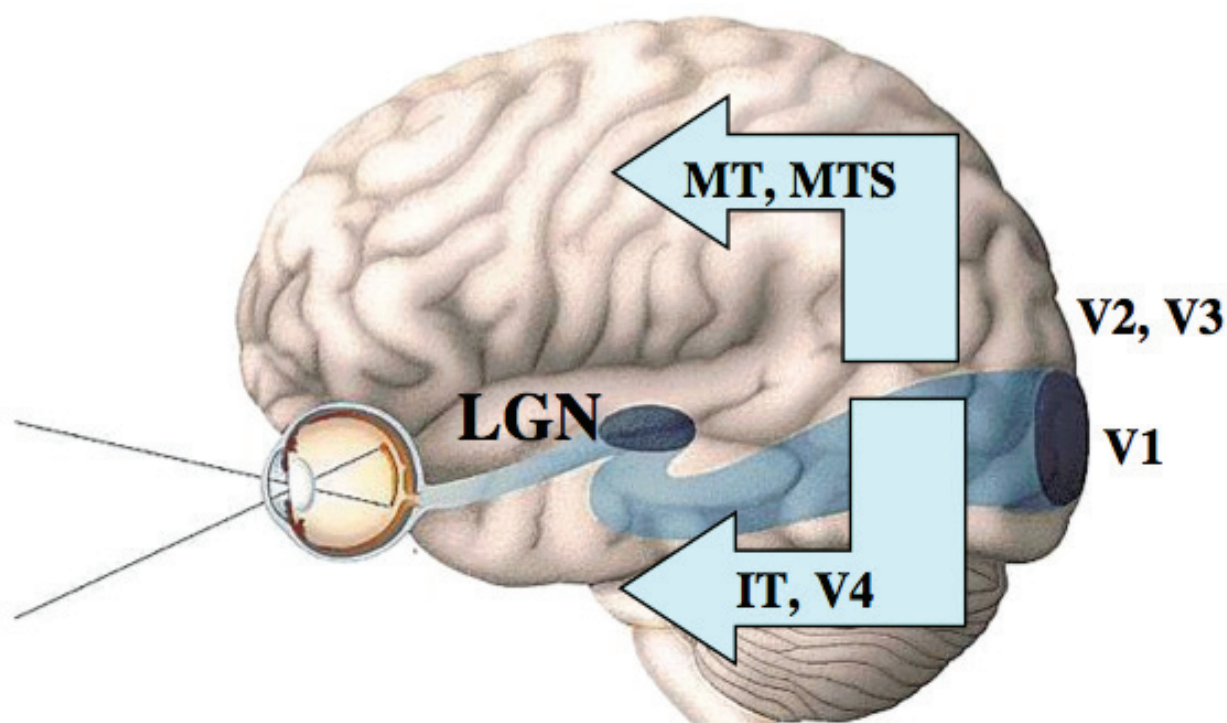
Figure 1. The visual pathway. The information comes in from the retina, next it is integrated in LGN, treated in V1, and finally processed in two different pathways : dorsal (MT, MTS, etc.) and ventral (V4, IT, etc.). The amygdala are almond-shaped groups of nuclei located deep within the medial temporal lobes and perform a primary role in the processing and memory of emotional reactions.

difference between these two methods is that deformation-based methods can be applied to single images as well as image sequences, and motion methods need a sequence of images. In both cases the facial features may be processed holistically (the faces are analysed as a whole) or locally (focusing on features from specific areas). In the case of deformation methods, the extraction of features relies on shape and texture changes through a period of time. On one hand the Holistic approaches either use the whole face [7, 25] or partial information about regions in the face like mouth or eyes [14, 21, 31]. For the motion extraction methods, the features are extracted by analysing motion vectors: holistically [8, 29], and locally [26, 34, 39] obtained.

3. Proposed bio-inspired architecture

A general diagram of the proposed methodology is shown in the figure 2. In the beginning, a sequence of images or video is captured in RGB format. Next, a processing based on green and red colours is realized to detect eyes and mouth corners. In this stage, we generate a grid of 6×5 based on coordinates of eyes and mouth corners. In this chapter, we only chose the six vertical bands for our experiments, the other combinations for analysis of the antropomorphical grid are subjects our future analysis.

Our bio-inspired treatment obtain the simple and complex neuron responses following the function of basic neurons of visual primary cortex (V1). Next, a temporal integration in complex neuron responses simulates the neurons in middle temporal cortex (MT).

Then, we extract the active neurons for each vertical band and compute their ratio of active quantity neurons. Next, each vertical band is compared to their symmetrical band.

Finally, we detect the asymmetries for each pair of vertical bands and cluster it in three classes applying the self-organizing maps (SOM).

In the next subsections we give a brief explanation of the steps of our approach where we assume that the face has been located correctly.

3.1. Description of input image sequences

In our proposed approach, the images of each sequence are in RGB colour, because we use the colour during the pre-processing to detect eyes and mouth corners. The process is described below. We suggest sampling capture to 25 frames per second and 1024×768 pixels. High sampling and high resolution is desirable.

3.2. Pre-processing

In this stage we realize three steps which are (1) the eyes and mouth detection, (2) we build the grid and finally we utilize only six regions, the six vertical bands (6-RVB) to analyse it (with left side of the stronger than right one). These steps will be briefly explained in the following paragraphs.

In the first step, let $I = \{I_1, I_2, I_3, \ldots I_t\}$ be an image face sequence where the eyes and mouth corners are correctly detected using red and green band colours as the sensibility of cones cells in our eyes. We work on the rectangle that contains the face; next, we define the point corresponding to the center of the face (x_c and y_c) and two parameters (width and height) for the size of the face. We use these anthropometric measures in the order to define three regions of interest that probably contain the mouth and eyes (left and right). For each region we find the corresponding coordinates of detected corners in both eyes and mouth and these coordinates are used in the next step.

In the second step, we use the coordinates of eyes and mouth corners to build an antropomorphical grid on the face. With the processing we can obtain twenty-five small regions, we also split the central part in two regions such as can see in the figure 2 obtaining thirty small regions (the step anthropometrical segmentation using black line doted).

Finally in this stage, we use only six symmetrical columns ($R_1, R_2, R_3, R_4, R_5, R_6$) of proposed grid to analyse expressions in the face.

3.3. Bio-inspired treatment

The facial expressions carry much more information that cannot be extracted from traditional facial expression analysis. So we propose, from the bio-inspired point of view, a methodology to analyse the symmetry of facial expressions. The objective with this is to achieve both a perspective, according to brain mechanisms, of the interpretation of facial motion relevance and a methodology that takes advantage of the tolerance to illumination change of the brain mechanisms. This can be divided into three main processing steps:
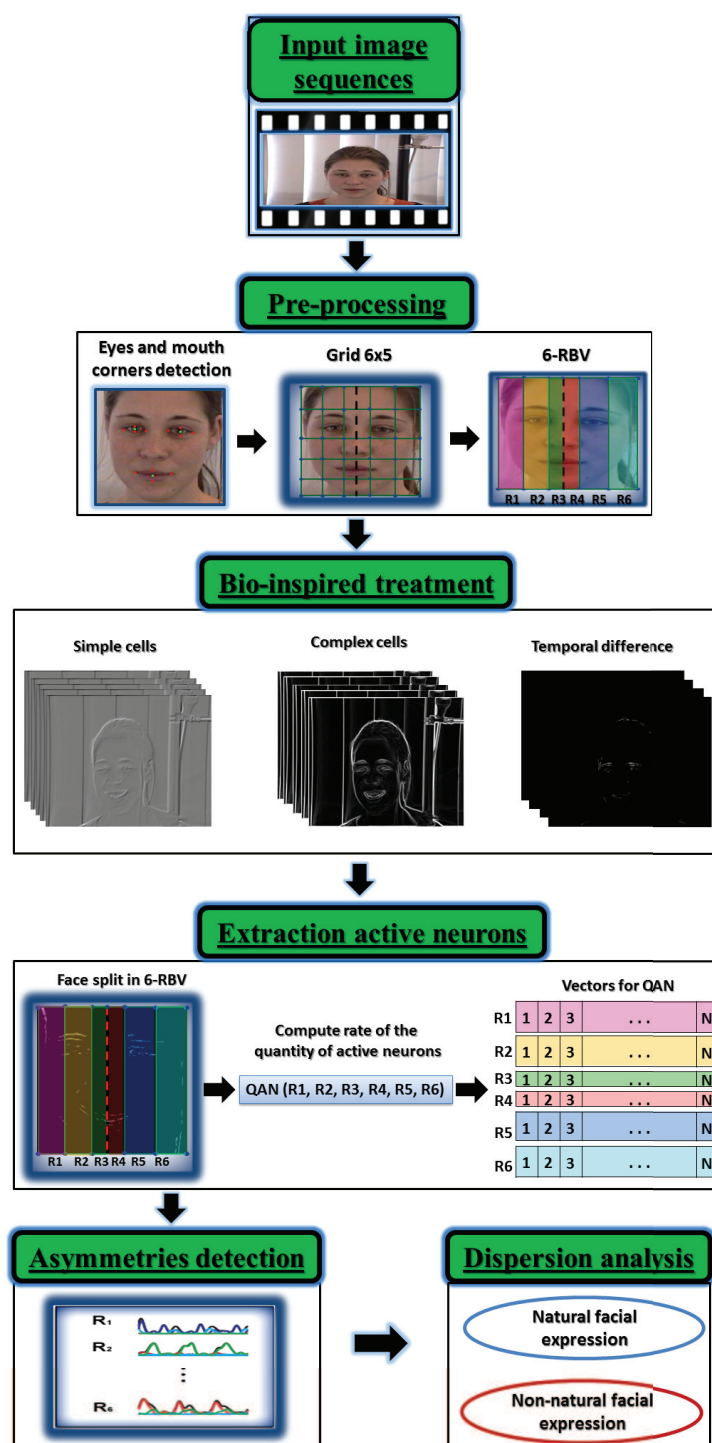
Figure 2. General architecture of bio-inspired proposed model. After video acquisition, a pre-processing generates an anthropometrical grid. Next, a treatment inspired in V1 determines the active neurons. Then, we compare two symmetrical regions: hears, eyes and nose to extract the vector composed of ratio of quantity active neurons. Finally, the asymmetries detection are evaluated with SOM to cluster into natural or non-natural facial expressions.

- Inspiring in simple neurons of V1 (primary visual cortex) and modelled by Gabor-like filters, we extract the simple active neurons in eight different orientations and two different

phases.

- Two simple neurons with different phase are merged to generate a complex neuron.
- A temporal integration between two consecutive complex neurons models MT neurons (Middle temporal area). The neurons with a major response to a threshold are considered as active neurons.

In the first proposed step, considering the processing done by the visual cortex, the first neurons that receive the stimuli from the eyes are the simple cells. Physiological evidence [5, 38] affirms that neuronal populations in the primary visual cortex (V1) of mammals exhibit contrast normalization. Neurons that respond strongly to simple visual stimuli, such as sinusoidal gratings, respond less well to the same stimuli when they are presented as part of a more complex stimulus which also excites other, neighbouring neurons. The behaviour of these neurons show a preference to specific orientations, which can be modelled computationally by the oriented Gabor-like filters [4, 22]:

$$S_{\theta,\varphi}(x,y,t) = \sum_{t=0}^{t} I(x,y,t) * G_{\theta}(\hat{x}, \hat{y}, \varphi) \tag{1}$$

where $S_{\theta}(x,y,t)$ are simple cells which is the result of the convolution between a pool of Gabor functions ($G_{\theta}(\hat{x}, \hat{y})$), in our case with 8 orientation and 2 phases, and the image t of the image sequences, $(\hat{x}, \hat{y})$ are the rotational components and (x,y) a position of the image. The Gabor function $G_{\theta}(\hat{x}, \hat{y})$ is defined as:

$$G_{\theta}(\hat{x}, \hat{y}, \varphi) = exp\left(-\frac{\hat{x}^2 + \gamma^2\hat{y}^2}{2\sigma^2}\right) cos\left(2\pi\frac{\hat{x}}{\lambda} + \varphi\right) \tag{2}$$

where $\hat{x} = x\cos\theta + y\sin\theta$, $\hat{y} = -x\sin\theta + y\cos\theta$, λ is the length-width, θ represent the orientation, φ the phase, σ the standard deviation and γ represent the relation of aspect.

In the second step, the responses of two diffent phases of simple cells $S_{\theta}(x,y,t)$ are then integrated in the complex cells by using a non-linear model that allows to merge the responses. Then, the complex cells responses are estimated by

$$C_{\theta}(x,y,t) = \sqrt{S_{\theta,\frac{\pi}{2}}(x,y,t)^2 + S_{\theta,-\frac{\pi}{2}}(x,y,t)^2} \tag{3}$$

where $C_{\theta}(x,y,t)$ represent complex cells responses, $\frac{\pi}{2}$ and $-\frac{\pi}{2}$ are the symmetric and anti-symmetric phases, respectively. Then, the eight orientations are integrated of the complex cells responses $C_{\theta}(x,y,t)$ for obtain one matrix of the map active neurons complex $C(x,y,t)$ for the image t.

Finally the third step, the neurons in the human brain are connected to MT neurons allowing a temporal processing that is defined as the temporal integration between $C(x,y,t)$ and $C(x,y,t-1)$ obtaining the map of active complex neurons by

$$D(x,y,t) = C(x,y,t) - C(x,y,t-1) \tag{4}$$

where $D(x,y,t)$ is the result of active complex neurons in the image t, $C(x,y,t)$ and $C(x,y,t-1)$ are the current image and the previous image, respectively. We use the information $D(x,y,t)$ to compute the number of active neurons.

3.4. Extraction of active neurons

In this stage, we realize the extraction of active complex neurons of the active map complex neurons ($D(x, y, t)$). For this, we use the six vertical bands and the map of active neurons. The 6-RVB overlapping in the map of active neurons. So, we compute the quantity of active neurons (QAN) for each region ($R_1, R_2, R_3, R_4, R_5, R_6$) and find the region with the maximum quantity of active neurons (QAN). The QAN is the number of active complex neurons in a region of the image.

So, we obtain six vectors of the quantity of active neurons (QAN): one for the QAN obtained and one for QAN expected for a region. The last one is the total possible QAN in a region. The first one computes only the active present neurons such that the ratio is always between 0 and 1. This information is used to compare two symmetric regions to detect temporal asymmetries.

3.5. Asymmetry detection

Next stage, we detect the asymmetries for each region (6-RVB) of the map of active complex neurons. To obtain or detect asymmetries in each image during two seconds using 6-RVB we follow the next steps:

- First step, we compute the difference between the symmetric regions. For example, the difference between region R_1 and R_6, that are the extreme regions.

- Then, compute average and standard deviation for each difference of the regions.

- So, we calculate a threshold (a minimum and a maximum) using the average and standard deviation.

- Then, we detect the asymmetries for each difference of regions in the image sequences, using the "rule": for each difference of regions, we verify that the difference is outside the threshold (i.e., minor to minimum and major to maximum). We count each outsider in the sequence, Out.

- Finally, we compute the Out for each pair of regions in the images sequence.

The detected asymmetries (Out) in the facial expression during the image sequences are weighed for each pair of symmetrical regions (R_1 with R_6, R_2 with R_5 and R_3 with R_4). For each database, we proposed an empirical threshold κ that we apply to all image sequence in the same database. This κ considers various aspects as the sampling frequency (we used two seconds in our experiments), the personal conditions, head motion and luminance-capture conditions.

3.6. Clustering in natural/non-natural expressions

For each different database we obtain a κ adapted to number of images in each sequence. So, we used the equation 5 for cluster the facial expressions in the natural/non-natural expressions:

$$C_\kappa = \frac{C_A}{C_E} \frac{2F}{T\kappa} \qquad (5)$$

where $\frac{C_A}{C_E}$ is the ratio of quantity of active neurons (QAN) in a region, T is the total number of images in a typical sequence in the database to compute κ, F is the capture frequency for the

sequences in this database and the κ index. With C_κ values we cluster in natural expression and non-natural expression.

These values are sending to SOM 1D that performs 5000 epochs, with a $\eta = 1.000$ to $\eta = 0.0001$ and they were analysed statistically.

4. Experimental results

We take six columns of proposed face grid. The symmetrical regions are: R1 with R6 (ears), R2 with R5 (eyes) and R3 with R4 (nose).

The tests for our approach took only the colour sequence of images for the CK+ database while for FG-Net and LTI-HIT we take all sequences of images. The last database was split per question in each interview (sequence of images).

Figure 3 shows a sequence of image of CK+ database for a non-natural expression. There are 11 asymmetries in both ears and nose regions, and 9 asymmetries in eyes region, the weighted sum is $10 = 11 * 0.3 + 11 * 0.2 + 9 * 0.5$. But this ratio is higher than fixed threshold $\kappa = 7$, then this sequence is non natural. We tested all the experiments for two seconds independently of sampling frequency (see table 1). We fix κ according table 1 for test our all available databases. In our experiments this threshold tolerates the generated asymmetries by illumination, environment, and personal conditions.

We test the databases Cohn-Kanade (CK+) (only colour images) [20], FG-Net [33], and LTI-HIT as shown in the table 2. We confirm that FG-Net is the most natural database. Our classification error was of 12% in average.

4.1. Description of databases

We used three databases to test our proposed approach: FG-Net, CK+, and LTI-HIT.

4.1.1. Cohn-Kanade (CK+) database

In 2000 and 2010, The Institute of Electrical & Electronics Engineers at Grenoble, France, created the CK and CK+, respectively (the same CK database but 107 new sequences of which 33 are in color), consists of 123 persons with 7 different expressions per person: anger, contempt, disgust, fear, happy, sadness and surprise. The image sequences vary in duration (i.e. 10 to 60 frames) and incorporate the onset of a neutral frame at the moment of peak formation of the facial expressions. This database was captured to 30 frames per second with a resolution 640×480. The final frame of each image sequence was coded using FACS (Facial Action Coding System) which describes person's expression in terms of action units (AUs) [20]. Participants were instructed by an experimenter to perform a series of 23 facial displays including single action units and specified emotion like expressions of joy, surprise, anger, disgust, fear, and sadness. Each display began and ended in a neutral face.

4.1.2. FG-Net database

In 2006, Facial Expressions and Emotions from the Munich Technical University created the FG-Net database. This is an image database containing face images showing a number of persons performing six different basic emotions defined by Ekman & Friesen: anger, fear,
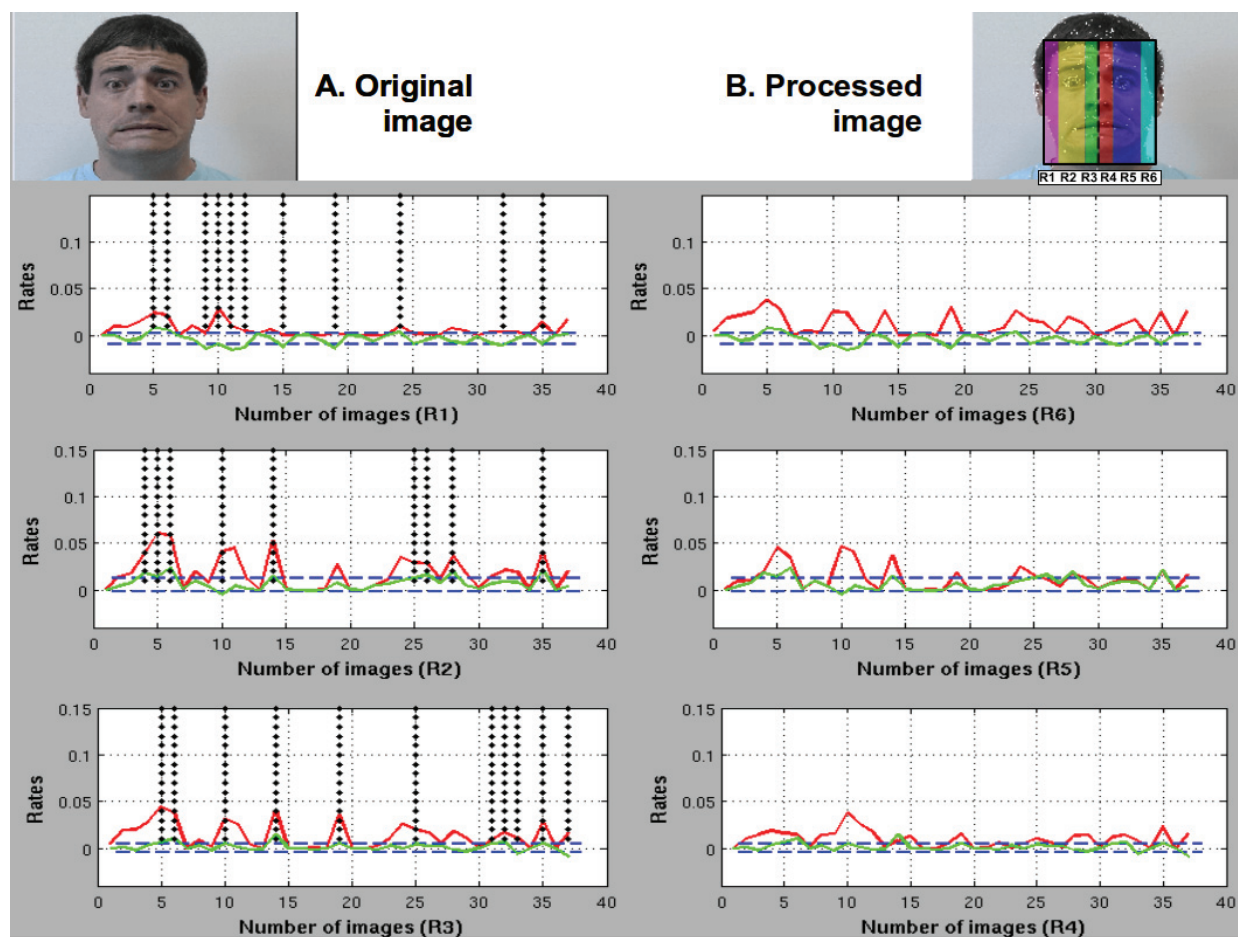
Figure 3. Asymmetry in vertical regions for a sequence of images in CK+. Horizontal red (black) line represents the data; horizontal green (gray) line, the difference between symmetrical regions; the separation between horizontal dashed lines is the symmetrical tolerance, and vertical dotted lines show the asymmetry (we show it only in the left graphics).

surprise, disgust, sadness and happiness. They also add the neutral expression for each person. This database consists of 19 persons with 21 sequences for each person, then it contains 399 different sequences. The sequences of images were captured to 25 frames per second with a resolution of 640×480 pixels [33]. This database allows to observe people react as natural as possible. As a consequence, it was tried to wake real emotions by playing video clips or still images after a short introductory phase instead of telling the person to play a role. This is in contrast to the asymmetry of luminance of the face, which includes not only asymmetry introduced by lighting, but also asymmetry introduced by the face itself.

4.1.3. LTI-HIT database

Finally, in 2010, the Children's Hospital of Tamaulipas created the LTI-HIT database. This database is taken from interviews that consists of 52 persons (11 men and 41 women) out of them 5 persons were put into one type of environment and 47 in the other. Each interview has between 27 and 33 different questions. The participants were between 18 and 66 years old. Each sequence of images were captured to 30 frames per second with resolution of 720×480 pixels and duration of 1.6 to 2.5 minutes per interviews. The interviews were made in natural conditions.

The tests for our approach took only the colour sequence of images for the CK+ database while for FG-Net and LTI-HIT we take all sequences images. The last database was split per question.
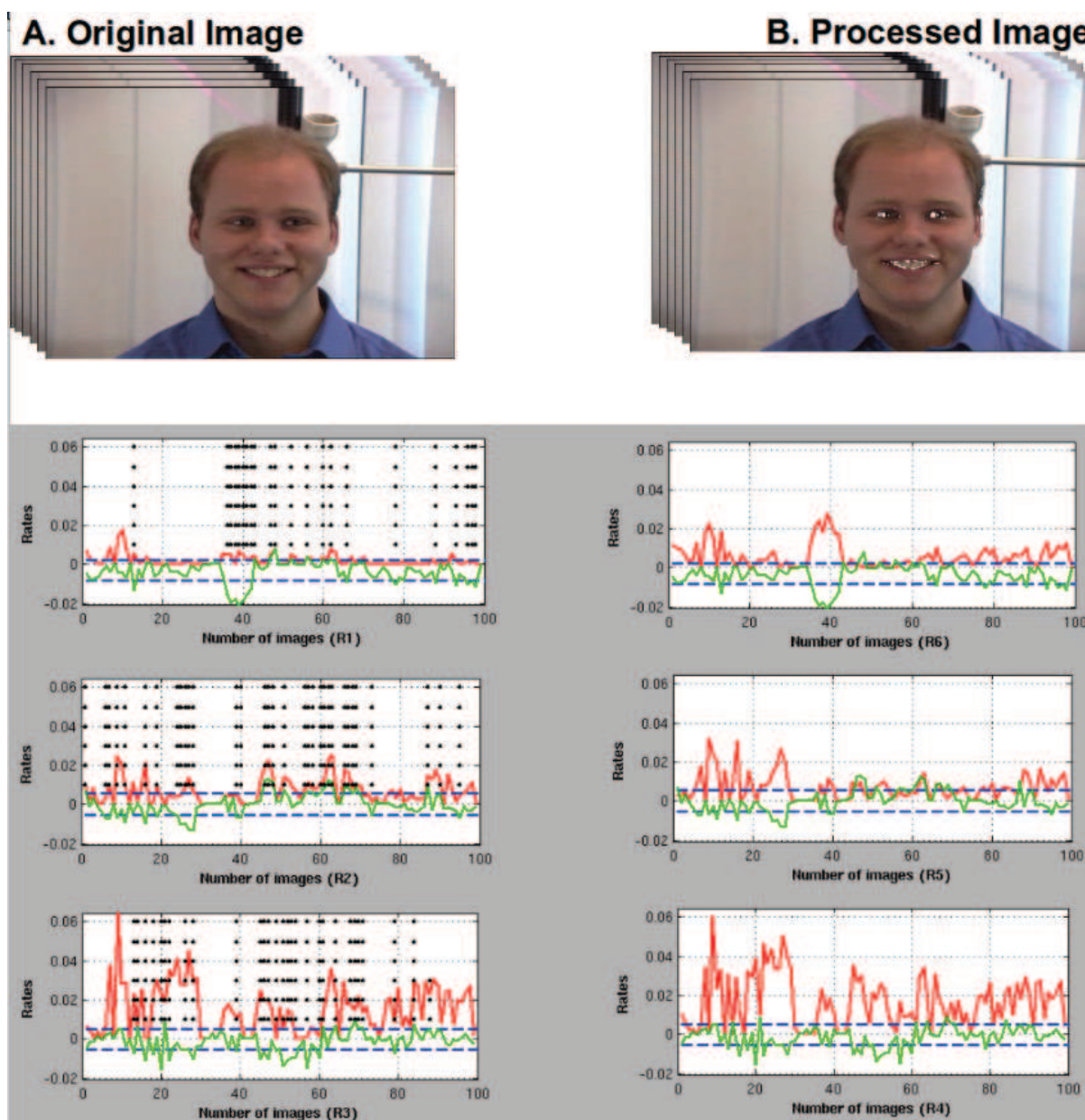

Figure 4. Results for face asymmetry detection in a sequence of images for FGNet. In this case, the boy is suspect to deceitful clues.

4.2. Remarks

The three databases are very different in its acquisition, illumination and manipulation of capture conditions. For weight these databases, we propose the κ index according to table 1.

In natural conditions, facial changes on the left side are only about 2% greater than overall right-side changes [30]. We establish a level of 1 to 5 for the personal conditions (hair-style,

Database (DB)	PC	HM	LC	κ index
CK+	3	1	2	6
FG-Net	3	3	3	9
LTI-HIT	3	4	4	11

Table 1. κ index is the value for each of the three databases. PC, personal conditions (facial expression); HM, head movement, and LC, luminance conditions. Each feature is between 0 to 5.

skin colour and tolerable static face asymmetry), and, in average, we choose a level 3 to model all personal conditions (PC) for our available databases.

The head motion (HM) is very different too between these three databases. In CK+ all persons are very controlled and the original sequences were cut to show only the evident expression. Again, the FG-Net sequences show activities before and after an evident expression. Finally, the LTI-HIT sequences were not prepared for the major expression. It shows the natural reactions of the interviewed persons. Then, we assign to HM variable the values 1, 3 and 4 to each database, respectively.

The luminance and capture conditions (LC) is too different. CK+ was created in a studio, FG-Net in a controlled environment and LTI-HIT in an uncontrolled environment. Then, we assign to LC variable 2, 3 and 4 to each database, respectively.

	Natural		Non-natural	
DB	DA	SOM	DA	SOM
CK+	42.42%	27.27%	33.33%	63.64%
FG-Net	33.60%	56.65%	39.68%	29.63%
LTI-HIT	34.55%	35.37%	38.70%	52.51%

Table 2. Natural/non-natural percentages according to κ index. The quantity of active neurons is manipulated into two manipulations: statistical analysis of active neurons (DA) and self-organizing maps (SOM). The percentages correspond to the obtained results with $12\% \pm 8\%$ error with two expert responses.

The figure 4 shows another example of test for a sequence of images in FGNet database. The vertical dotted lines show the asymmetries (only shown the left side). The horizontal red lines are the ratios of quantity, while the horizontal green lines are the difference between symmetrical regions. Finally, the horizontal blue lines are tolerable threshold obtained from this sequence of images. In this case, the boy is suspected to deceitful clues because all the graphics show a higher value (22, 33 and 28 for each pair of vertical band) and the weighed sum is $29.3 = 22 * 0.2 + 33 * 0.5 + 28 * 0.3$ that is higher than $\kappa = 18.0$ index.

The table 2 resume the obtained ratios for natural/non-natural facial expressions with statistical analysis and SOM. If we see SOM percentages, these percentages show FG-Net as greater in natural facial expressions than the other two databases, CK+ also has greater percentage followed by LTI-HIT with non-natural facial expressions.

For statistical analysis we have the opposed results. This difference is due to concentration of values and the dispersion. With a simple statistical analysis, all responses are separated based
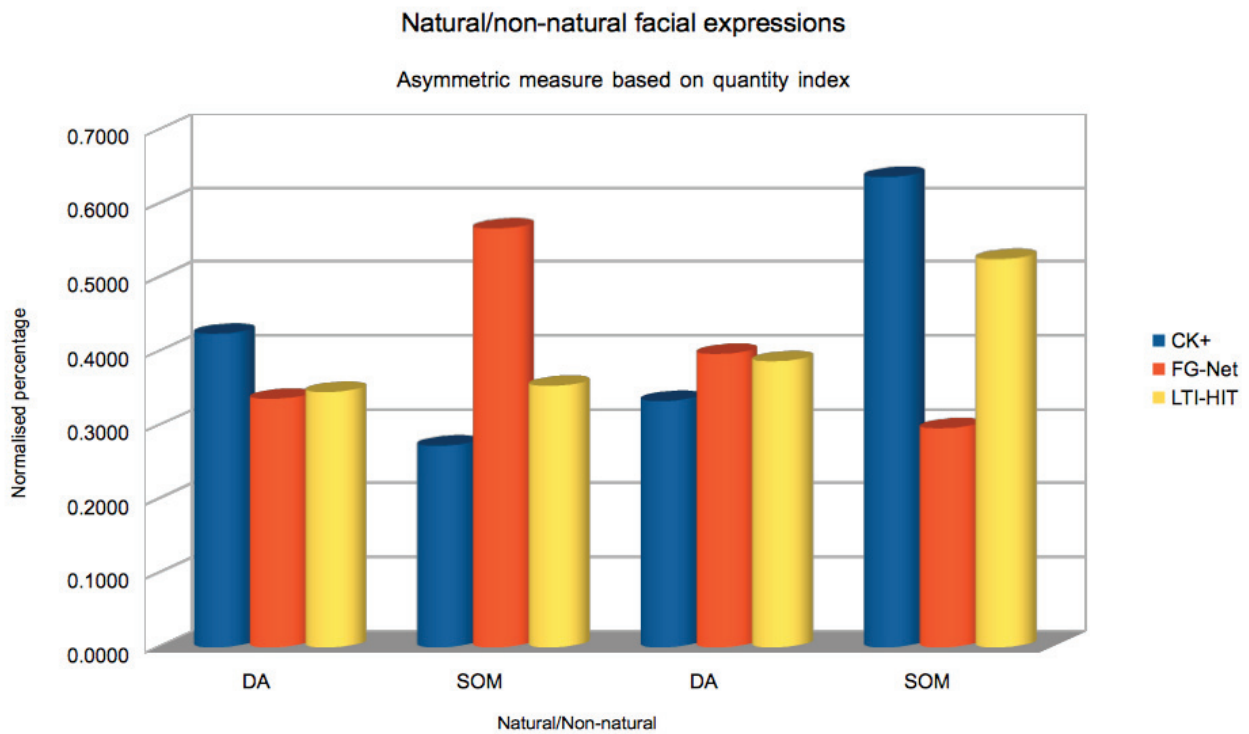
Figure 5. Asymmetry based on the quantity index. The first DA and SOM blocks correspond to Natural facial expressions. The last two DA and SOM correspond to Non-Natural facial expression.

on average and standard deviation, while the SOM application shows the centroids guided by the density and dispersion.

5. Conclusion

We have shown a bio-inspired approach to processing facial gesture for clustering into natural and non-natural facial expressions. It takes advantage of early primary visual areas in the brains of human beings for feature-extraction. Futhermore, it simulates the integration and discrimination of the superior visual areas considering the symmetrical and asymmetrical measures within the respective time span. This process ends in a bio-inspired neural network for clustering into natural and non-natural facial expressions.

The asymmetry between symmetrical regions of the face analysed during two seconds allows us to classify in: natural or non-natural facial expression. The quantity of asymmetries and its relations with different symmetrical regions is applied as a new biometric measure [24] (the precision of measurement is necessary). But we apply it for classification of natural or non natural expressions (the precision is not necessary).

In our bio-inspired proposed architecture, after face detection, an anthropomorphical grid is proposed. For each region, we obtain the ratio of quantity of active neurons (QAN) obtained from the modelled simple and complex neurons by our Gabor-like filters.

This proposed model can work for any database. Furthermore, the database FG-Net and Cohn-Kanade were tested. More than 63% of CK+ video sequences contain asymmetries. This database was built with simulation expressions (non natural expressions) and we confirm

that our results are modelled correctly regarding the capture, illumination and personal conditions.

The different conditions of each database do not allow a simple characterization. We proposed an experimental index to model these conditions (e.g. personal conditions, head motion and luminance-capture), κ. In future, we will determine this index automatically for each database and, more precisely, for each sequence of images into database.

Using the index κ we have shown that our approach is independent of each database. This independence was tested in our experiments.

The preliminary steps of our approach show the feasibility to detect suspected persons in nervous or altered situations. Here is the beginning of our methodology to detect deceitful clues in facial expressions.

Author details

Claudio Castellanos Sánchez
Laboratory of Information Technology, Cinvestav-Tamaulipas and División de Estudios de Posgrado e Investigación, Instituto Tecnológico de Ciudad Victoria, Mexico

Manuel Hernández Hernández
Departamento de Ingeniería en Sistemas Computacionales, Instituto Tecnológico Superior de Tantoyuca, Mexico

Pedro Luis Sánchez Orellana
División de Estudios de Posgrado e Investigación, Instituto Tecnológico de Ciudad Victoria, Mexico

6. References

[1] Adolphs, R., Tranel, D. & Damasio, A. R. [Sep-1998]. The human amygdala in social judgment., *Nature* 393(6684): 470–474.

[2] Adolphs, R., Tranel, D., Damasio, H. & Damasio, A. R. [1995]. Fear and the human amygdala, *The Journal of Neuroscience* 15(9): 5879–5891.

[3] Belhumeur, P. N., Hespanha, J. & Kriegman, D. J. [1997]. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19(7): 711–720.

[4] Castellanos-Sánchez, C. [2005]. *Neuromimetic connectionist model for embedded visual perception of motion*, PhD thesis, Université Henri Poincaré (Nancy I), Nancy, France. Bibliothèque des Sciences et Techniques.

[5] Clatworthy, P., Chirimuuta, M., Lauritzen, J. & Tolhurst, D. [2003]. Coding of the contrasts in natural images by populations of neurons in primary visual cortex (v1), *Vision Research* 43(18): 1983 – 2001.

[6] del Solar, J. R. & Quinteros, J. [2008]. Illumination compensation and normalization in eigenspace-based face recognition: A comparative study of different pre-processing approaches, *Pattern Recognition Letters* 29(14): 1966 – 1979.

[7] Devi, B. J., Veeranjaneyulu, N. & Kishore, K. [2010]. A novel face recognition system based on combining eigenfaces with fisher faces using wavelets, *Procedia Computer Science* 2(0): 44 – 51.

[8] Dornaika, F., Lazkano, E. & Sierra, B. [2011]. Improving dynamic facial expression recognition with feature subset selection, *Pattern Recognition Letters* 32(5): 740 – 748.

[9] Ekman, P. [1980]. Asymmetry in facial expression, *Science* 209(4458): 833–834.

[10] Essa, I. A. & Pentland, A. P. [1997]. Coding, analysis, interpretation, and recognition of facial expressions, *IEEE Trans. Pattern Anal. Mach. Intell.* 19: 757–763.

[11] Fasel, B. & Luettin, J. [2003]. Automatic facial expression analysis: a survey, *Pattern Recognition* 36(1): 259 – 275.

[12] Fellenz, W., Taylor, J., Tsapatsoulis, N. & Kollias, S. [1999]. Comparing template-based, feature-based and supervised classification of facial expressions from static images, *Proceedings of Circuits, Systems, Communications and Computers* pp. 5331–5336.

[13] Gosselin, F., Spezio, M. L., Tranel, D. & Adolphs, R. [2011]. Asymmetrical use of eye information from faces following unilateral amygdala damage, *Social Cognitive and Affective Neuroscience* 6(3): 330–337.

[14] Gu, W., Xiang, C., Venkatesh, Y., Huang, D. & Lin, H. [2012]. Facial expression recognition using radial encoding of local gabor features and classifier synthesis, *Pattern Recognition* 45(1): 80 – 91.

[15] Hasselmo, M. E., Rolls, E. T. & Baylis, G. C. [1989]. The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey, *Behavioural Brain Research* 32(3): 203 – 218.

[16] Haxby, J. V., Hoffman, E. A. & Gobbini, M. I. [2000]. The distributed human neural system for face perception., *Trends in cognitive sciences* 4(6): 223–233.

[17] Hung, Y., Smith, M. L., Bayle, D. J., Mills, T., Cheyne, D. & Taylor, M. J. [2010]. Unattended emotional faces elicit early lateralized amygdala-frontal and fusiform activations, *NeuroImage* 50(2): 727 – 733.

[18] Kim, H.-C., Kim, D., Bang, S.-Y. & Lee, S.-Y. [2004]. Face recognition using the second-order mixture-of-eigenfaces method, *Pattern Recognition* 37(2): 337 – 349.

[19] Lanitis, A., Taylor, C. J. & Cootes, T. F. [1997]. Automatic interpretation and coding of face images using flexible models, *IEEE Trans. Pattern Anal. Mach. Intell.* 19: 743–756.

[20] Lucey, P., Cohn, J., Kanade, T., Saragih, J., Ambadar, Z. & Matthews, I. [2010]. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression, *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pp. 94–101.

[21] Maalej, A., Amor, B. B., Daoudi, M., Srivastava, A. & Berretti, S. [2011]. Shape analysis of local facial patches for 3d facial expression recognition, *Pattern Recognition* 44(8): 1581 – 1589.

[22] Marcelja, S. [1980]. Mathematical description of the reponses of simple cortical cells, *Journal of the Optical Society of America A* 70(11): 1297–1300.

[23] Mitra, S. & Acharya, T. [2007]. Gesture recognition: A survey, *IEEE Transactions on Systems, Man and Cybernetics - Part C* 37(3): 311–324.

[24] Mitra, S. & Savvides, M. [2005]. Analyzing asymmetry biometric in the frequency domain for face recognition, *IEEE ICASSP* II: 953–956.

[25] Nair, B. M., Foytik, J., Tompkins, R., Diskin, Y., Aspiras, T. & Asari, V. [2011]. Multi-pose face recognition and tracking system, *Procedia Computer Science* 6(0): 381 – 386.

[26] Park, S. & Kim, D. [2009]. Subtle facial expression recognition using motion magnification, *Pattern Recognition Letters* 30(7): 708 – 716.

[27] Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrews, C., Calder, A. J., Bullmore, E. T., Perrett, D. I., Rowland, D., Williams, S. C. R., Gray, J. A. & David, A. S.

[1997].   Aspecific neural substrate for perceiving facial expressions of disgust, *Nature* 389(6650): 174 – 194.

[28]   Rowley, H. A., Baluja, S. & Kanade, T. [1998]. Neural network-based face detection, *IEEE Trans. Pattern Anal. Mach. Intell.* 20: 23–38.

[29]   Sánchez, A., Ruiz, J., Moreno, A., Montemayor, A., Hernández, J. & Pantrigo, J. [2011].   Differential optical flow applied to automatic facial expression recognition, *Neurocomputing* 74(8): 1272 – 1282.   Selected Papers from the 3rd International Work-Conference on the Interplay between Natural and Artificial Computation (IWINAC 2009).

[30]   Schmidt, K. L., Liu, Y. & Coh, J. F. [2006]. The role of structural facial asymmetry in in asymmetry peak facial expressions, *Literality* 11(6): 540–561.

[31]   Venkatesh, Y., Kassim, A. A. & Murthy, O. R. [2009]. A novel approach to classification of facial expressions from 3d-mesh datasets using modified pca, *Pattern Recognition Letters* 30(12): 1128 – 1137.

[32]   Vuilleumier, P. & Pourtois, G. [2007].   Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging, *Neuropsychologia* 45(1): 174 – 194.

[33]   Wallhoff, F. [2006]. Facial expressions and emotion database.
URL: *http://www.mmk.ei.tum.de/~waf/fgnet/feedtum.html*

[34]   Wang, T.-H. & Lien, J.-J. J. [2009]. Facial expression recognition system based on rigid and non-rigid motion separation and 3d pose estimation, *Pattern Recognition* 42(5): 962 – 977.

[35]   Weymar, M., Law, A., Ohman, A. & Hamm, A. O. [2011]. The face is more than its parts: Brain dynamics of enhanced spatial attention to schematic threat, *NeuroImage* 58(3): 946 – 954.

[36]   Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B. & Jenike, M. A. [January-1998]. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge., *The Journal of Neuroscience* 18(1): 411–418.

[37]   Williams, M. A., McGlone, F., Abbott, D. F. & Mattingley, J. B. [2005].   Differential amygdala responses to happy and fearful facial expressions depend on selective attention, *NeuroImage* 24(2): 417 – 425.

[38]   Willmore, B. D., Bulstrode, H. & Tolhurst, D. J. [2012]. Contrast normalization contributes to a biologically-plausible model of receptive-field development in primary visual cortex (v1), *Vision Research* 54(0): 49 – 60.

[39]   Xiang, T., Leung, M. & Cho, S. [2008].   Expression recognition using fuzzy spatio-temporal modeling, *Pattern Recognition* 41(1): 204 – 216.