

# FACE AUTHENTICATION AND NON-STANDARD IMAGE GEOMETRIES: ISSUES, RESULTS AND PROSPECTS

*Massimo Tistarelli, Enrico Grosso, and Andrea Lagorio*

University of Sassari  
Computer Vision Laboratory  
Palazzo del Pousalid, piazza Duomo 6, 07041 Alghero (SS) – Italy  
e-mail: tista@uniss.it – grosso@uniss.it  
web: www.uniss.it

## ABSTRACT

The analysis of biological features for the recognition of humans attracted the interest of many researchers in the last decades. Despite the considerable efforts, in many cases, the dimensionality of the data and the complexity of the matching problem made it difficult to devise fully operational authentication systems. This is probably also the reason of the greatest diffusion of other biometric modalities such as fingerprint and iris scanning.

As in other applications, the adoption of non-standard image sampling geometries may both reduce the dimensionality of the data and the complexity of the recognition problem. In this paper we will try to demonstrate that this is true also in the case of face image analysis where non-standard imaging geometries can be successfully applied.

## 1. INTRODUCTION

Many efforts have been devoted to the study of computer systems for automated verification of person's identity (either by means of recognition from a database of known individuals or as authentication of one's identity). This relatively new technology has an indubitable potential: surveillance, secure access control and e-commerce are just few of the many envisaged applications.

In principle (particularly for social acceptability) the analysis of face images seems to be the best way to accomplish the task of determining the personal identity. Many difficulties arise from the enormous dimensionality of the search space when dealing with natural images (both for the number of elements in a typical data set and for the number of samples for each data). These and other issues related to the definition of optimal similarity measurements, make face recognition and visual authentication a still open and challenging problem in computer vision [1-8].

On the other hand, the mutual recognition of individuals is a natural and fundamental capability of most living creatures. There are many lessons we may learn from natural systems. Among them the minimal energy consumption is a natural principle of paramount importance. Many natural mechanisms strongly rely on this principle. This paper highlights some basic principles related to living systems to gather few insights on sensory data acquisition and

processing for recognition [9,10]. Towards this end, the motivation and advantages of space-variant geometries are described, as well as current implementations of space-variant image generation. A system for personal recognition, based on space-variant face image analysis is described and some recent testing results are also reported.

## 2. NEUROPHYSIOLOGY AND INFORMATION PROCESSING

Neural systems that mediate face recognition appear to exist very early in life. In normal infancy, the face holds particular significance and provides nonverbal information important for communication and survival [15]. Face recognition ability is present during the first 6 months of life, while a visual preference for faces and the capacity for very rapid face recognition are present at birth [16,17]. By 4 months, infants recognize upright faces better than upside down faces, and at 6 months, infants show differential event-related brain potentials to familiar versus unfamiliar faces [18,19].

Much is known about the neural systems that subserve face recognition in adult humans and primates. Face-selective neurons have been found in the inferior temporal areas (TEa and TEm), the superior temporal sensory area, the amygdala, the ventral striatum (which receives input from the amygdala) and the inferior convexity [20].

Using functional magnetic resonance imaging (fMRI), an area in the fusiform gyrus was found significantly activated when the subjects viewed faces [21-23]. Within this "general face activation area" specific regions of interest have been reported responding significantly more strongly to passive viewing of face-specific stimuli. A recent fMRI study on individuals with autism and Asperger syndrome showed a failure to activate the fusiform face area during face processing. Damage to fusiform gyrus and to amygdala results in impaired face recognition [24,25].

Parts of the inferior and medial temporal cortex may work together to process faces. For example, the anterior inferior temporal cortex and the superior temporal sulcus project to the lateral nucleus of the amygdala, with the amygdala responsible for assigning affective significance to faces, and thus affecting both attention and mnemonic aspects of face processing [26,27].

Behavioral studies suggest that the most salient parts for face recognition are, in order of importance, eyes, mouth, and nose [28]. Eye-scanning studies in humans and monkeys show that eyes and hair/forehead are scanned more frequently than the nose [29,30], while human infants focus on the eyes rather than the mouth [31]. Using eye-tracking technology to measure visual fixations, Klin [32] recently reported that adults with autism show abnormal patterns of attention when viewing naturalistic social scenes. These patterns include reduced attention to the eyes and increased attention to mouths, bodies, and objects.

The high specialization of specific brain areas for face analysis and recognition motivates the relevance of faces for social relations. On the other hand, this suggests that face understanding is not a low level process but involves higher level functional areas in the brain. These, in turn, must rely on a rich series of low level processes applied to enhance and extract face- specific features:

- **Face detection and tracking.** This process involves the analysis of dynamic as well as geometric and photometric data on the retinal projection of the face.
- **Facial features extraction.** Facial features are not simply distinctive points on the segmented face, but rather a collection of image features representing specific (and anatomically stable) areas of the face such as the eyes, eyebrows, ears, mouth, nostrils etc. Other, subject-specific, features are also included, such as the most famous Marilyn Monroe's naevus.
- **Face image registration and warping.** Humans can easily recognize faces which are rotated and distorted up to a limited extent. The increase in time reported for recognition of rotated and distorted faces implies: the expectation on the geometric arrangement of facial features, and a specific process to organize the features (analogous to image registration and warping) before the actual recognition process can take place.
- **Feature matching.** This process involves the comparison between the extracted set of facial features and the same set stored in the brain. The two process of feature extraction and matching (or memory recall) are not completely separated and sequential. From the eye scan paths recorded during face recognition experiments, it seems that, after moving the eyes over few general facial features, the gaze is directed toward subject-specific features, probably to enforce the expected identity.

From these processes higher level reasoning is possible, not only to determine the subject's identity, but also to understand more abstract elements (even uncorrelated to the subject's identity) which characterize the observed person (age, race, gender, emotion etc.). These, in turn, also require the intervention of task-specific processes, such as motion

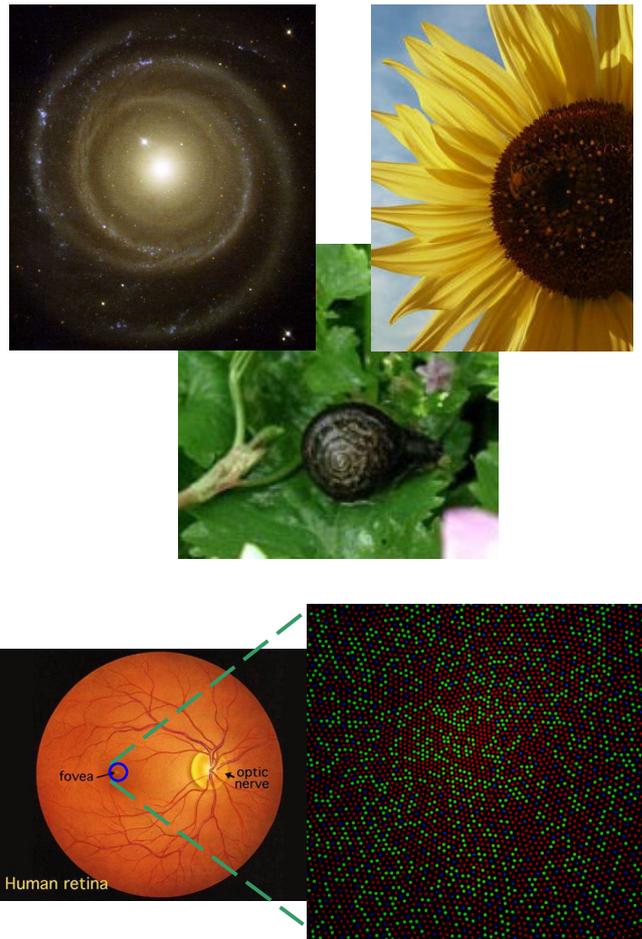


Figure 1: Examples of spiral, space-variant, topologies for the spatial organization of elements in natural systems.

analysis and facial features tracking for understanding emotion-specific patterns [4, 6, 7, 33-35].

As it is beyond the scope of this paper to trace all face-specific information processing, we will concentrate on face recognition and authentication, which not only are among the most studied aspects related to visual processing human faces, but it is probably the most representative of the tasks involved in face image analysis.

### 3. SENSOR TOPOLOGIES AND SPIRAL GRIDS

To achieve any visual task, including face recognition, humans are able to purposively control the flow of input data limiting the amount of information gathered from the sensory system [36-38]. This is needed to reduce the space and computation time required to process the incoming information.

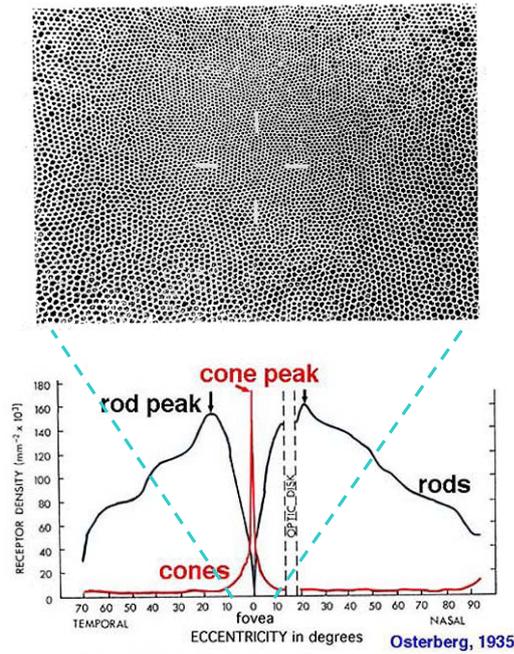


Figure 2: Spatial organization of sensory elements in the human retina. (Top) Picture of the cones in the fovea centralis. (bottom) Diagram of the spatial distribution of the photoreceptors in the retina.

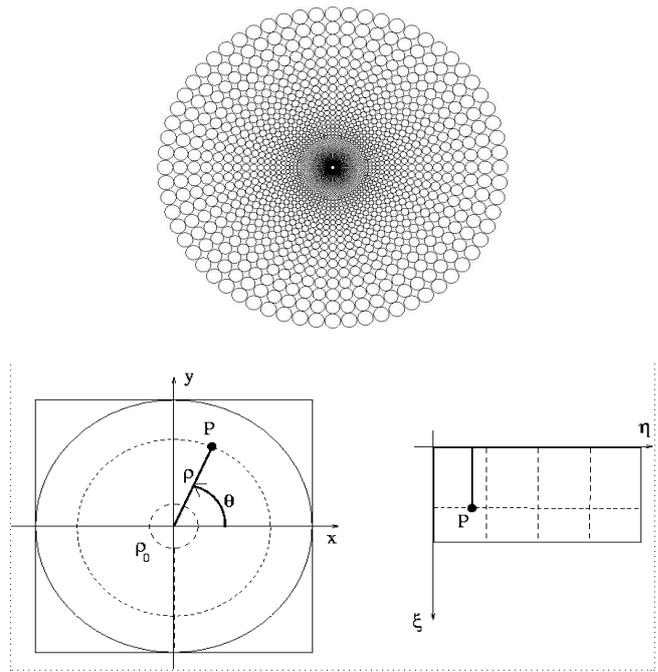


Figure 4: Design of a space-variant imaging topology for image re-mapping. (Top) Actual layout of the log-polar sampling applied to images. (Bottom) Schematic representation of the log-polar transformation and the associated coordinate systems.

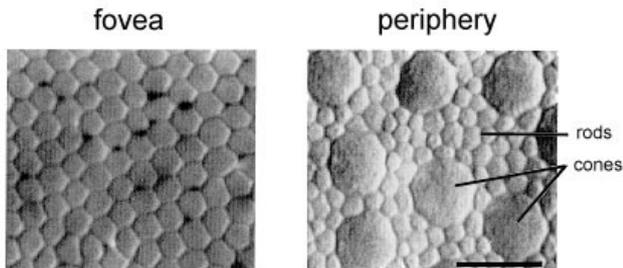


Figure 3: Comparison of the spatial arrangement and size of the cones in the fovea and in the periphery.

It seems that every natural system is designed to minimize the amount of energy spent to achieve its goals. The minimal energy policy not only influences the processes and activities of living systems, more often related to survival, but also the mechanisms determining the geometrical structure of natural systems themselves. There are many examples in nature where the geometry of the system is specifically designed to limit the amount of space required to store/organize its elements. In many cases it seems that spiral topologies are the best suited geometric arrangements to allow an optimal (space saving) distribution of the elements in space. Some examples are depicted in figure 1.

The anatomy of the early stages of the human visual system is a clear example: despite the formidable acuity in the fovea centralis (1 minute of arc) and the wide field of view (about 140x200 degrees of solid angle), the optic nerve is composed of only  $10^6$  nerve fibres. The space-variant distribution of the ganglion cells in the retina allows a

formidable data flow reduction. In fact, the same resolution would result in a space-invariant sensor of about  $6 \times 10^8$  pixels, thus resulting in a compression ratio of 1:600 [39].

The probability density of the spatial distribution of the ganglion cells, which convey the signal from the retinal layers to the optic nerve and is responsible for the data compression, follows a logarithmic-polar law. The number of cells decreases from the centre of the retina toward the periphery, with the maximal resolution in the fovea [40]. The same data compression can be obtained on electronic images, either by using a specially designed space-variant sensor [41], or re-sampling a standard image according to the log-polar transform [38,39].

The analytical formulation of the log-polar mapping describes the mapping that occurs between the retina (retinal plane  $(\rho, \theta)$ ) and the visual cortex (log-polar or cortical plane  $(\xi, \eta)$ ). The derived logarithmic-polar law, taking into account the linear increment in size of the receptive fields, from the central region (fovea) towards the periphery, is given by:

$$\begin{cases} x = \rho \sin \theta \\ y = \rho \cos \theta \end{cases} \quad \begin{cases} \xi = \log_a (\rho / \rho_0) \\ \eta = q\theta \end{cases} \quad (1)$$

where  $a$  defines the amount of overlap among neighboring receptive fields,  $\rho_0$  is the radius of the innermost circle,  $1/q$  is the minimum angular resolution of the log-polar layout, and  $(\rho, \theta)$  are the polar coordinates.

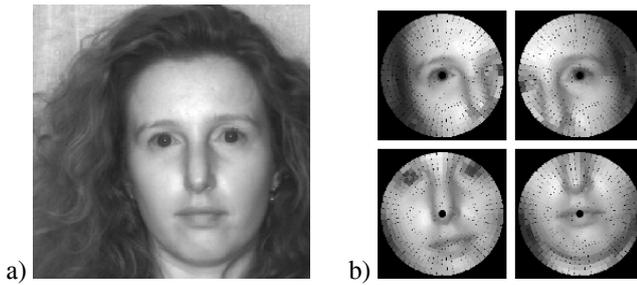


Figure 5: Extraction of space-variant fixations. a) Original face image. b) Log-polar fixations extracted from (a).

Other models for space-variant image geometries have been proposed, like the truncated pyramid [42], the reciprocal wedge transform (RWT) [43] and the complex logarithmic mapping (CLM) [44].

Several implementations of space-variant imaging have been developed: space-variant sensors [41], custom designed image re-sampling hardware [45], and special software routines [38,46]. Given the high processing power of current computing hardware, image re-mapping can be performed at frame rate without the need of special computing hardware, and also allows the use of conventional, low cost, cameras.

Space-variant image geometries have been applied to a variety of tasks, such as: visual tracking, form and object recognition, computation of stereo disparity and optical flow, depth perception and robot navigation [14]. More recently space-variant imaging has been applied to facilitate recognition with different biometric modalities [46,47]. A remarkable example is the Iris Scan iris recognition system patented by John Daugman [49]. The iris is firstly filtered by a bank of Gabor filters, then a set of features are extracted along a circular pattern following the rings of the iris. The extracted set of features are bitmapped to form a in-dimensional string, the *IrisCode*, which is the template used to perform the matching. Bigun [50] proposed a space-variant sampling of face images to build a representation based on Gabor filtering for classification purposes. Log-polar fixations have been also used to perform face image matching [14]. More recently, Jain [51] applied a bank of Gabor filters, along a spiral grip, to capture both local and global details in a fingerprint as a compact fixed length FingerCode. The fingerprint matching is based on the Euclidean distance between the two corresponding FingerCodes.

#### 4. SPACE-VARIANT FACE RECOGNITION

An important perceptual mechanism in humans is visual attention. Again, as not all visually available data is relevant for every given task, the human perceptual system is capable of making a selection of the input signal in various dimensions: “signal space” (low or high frequency data), depth (image areas corresponding to objects close or far from the observer), motion (static or moving objects) etc. The selection is controlled by a proper attention mechanism through ad-hoc band-limiting or focusing processes, which

determine the areas of interest in the scene to which direct the gaze [48].

Given the abundance of data in a face image, both space-variant image re-sampling and the adoption of a selective attention mechanism can greatly improve the performance of recognition/authentication algorithms. The algorithms described for face authentication exploit these principles by extracting few pre-determined areas within the face and applying the matching to those image areas only. Processing face images these mechanisms must be tuned to the most salient features of the face itself.

##### 4.1 Facial features detection

A technique applied to detect the facial features relies on the application of morphological operators to enhance structured iconic data such as contours, valleys and peaks in the gray levels. This information is gathered to make hypotheses for the presence of specific facial features. For example, the visible part of the sclera of the eye corresponds to a peak in the gray levels while the nostrils correspond to valleys. The position of the facial features is determined in two steps:

- by first computing the cumulative values of the filtered image along the rows. The eyes and mouth correspond to the areas with higher cumulative values; the mouth is assumed to lie in the area below the eyes;
- the same process is performed along the columns in the area corresponding to the eyes, determined at the previous step. The two maxima correspond to the horizontal position of the two eyes.

In order to avoid false matches a geometrical constraint is enforced to the position of the eyes and mouth, which is to lie at the vertexes of a triangle. The values assumed by the angles of the triangle are bounded by values determined experimentally ( $44^\circ < \alpha_i < 84^\circ$ ).

The exact position of the features is determined by computing the cross-correlation between the image and a feature template, within a 10x10 pixels window centered on the previously determined position. The template is obtained by cutting the eyes and mouth out of a sample image of the same subject. From an extensive test this choice demonstrated to give more accurate results than computing an average template. This is due to the fact that the averaging process deforms considerably the feature’s shape degrading the matching results.

The three correlation values stemming from the eyes and mouth are averaged to obtain a score between -1 and 1. If the geometric constraint is satisfied and the matching score is higher than a given threshold the fixations are considered as valid ones. In order to determine the discriminant value for the correlation score, a validation test has been performed on a set of 2019 images completely uncorrelated from the recognition database. These images have been divided into two classes: all images (1609) where the facial features are partially occluded or not visible, plus all the images where the mean difference between the estimated and the manually

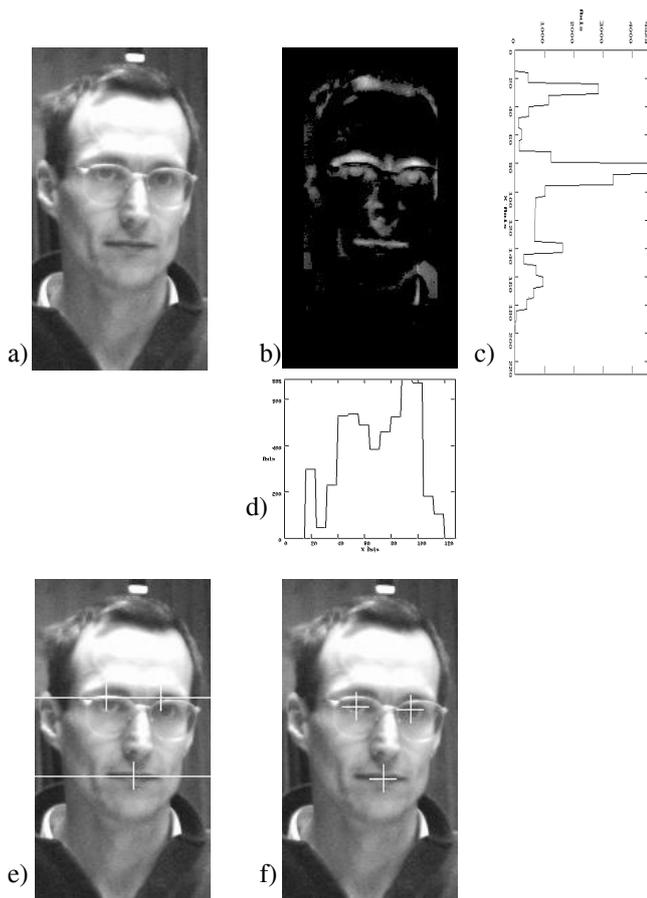


Figure 6. Facial features detection. a) Original image. b) Valley image after morphological filtering. c) Cumulative values computed along the rows of (b). d) Cumulative values computed along the columns of the window identified by the maxima extracted from (c). e) First guess for the eyes and mouth position. f) Final position of the facial features computed by template matching around the estimation in (e).

determined feature positions is greater than a given threshold<sup>1</sup>; all remaining images in the set (410).

The FAR and FRR test values were computed from the feature correlation scores of the two image sets. These statistical measures represent the capability of separating the two classes, valid and wrong features. The score value corresponding to equal FAR and FRR determines the reliability of the estimated features.

#### 4.2 Matching face representations

The face matching is performed by computing the correlation between the representation of the reference subject and the acquired face image. The algorithm is based on the following steps:

- Given the position of selected facial features (the eyes and the mouth), three space-variant fixations are

<sup>1</sup> This threshold is determined statistically by computing the probability of locating the facial features correctly in more than 50% in a given image ensemble.

extracted by re-sampling the acquired image of the subject with a log-polar grid centered on the position of the facial features.

- The log-polar images are warped to simulate views which are as close as possible to the pose and orientation of the reference subject's face (assumed to be roughly parallel to the image plane).
- Corresponding fixations are compared by computing the sum of the absolute value of gray level differences<sup>2</sup> and the normalized correlation. Two matching scores are obtained from each fixation independently:

$$C(S) = \frac{\sum_x \sum_y (M(x,y) - \bar{M})(S(x,y) - \bar{S})}{\sqrt{\sum_x \sum_y (M(x,y) - \bar{M})^2 \sum_x \sum_y (S(x,y) - \bar{S})^2}} \quad (2)$$

$$D(S) = \sum_x \sum_y \left| \frac{M(x,y)}{\sum_x \sum_y M(x,y)} - \frac{S(x,y)}{\sum_x \sum_y S(x,y)} \right| \quad (3)$$

- The scores obtained by the log-polar fixations are combined to form a 6 components vector representing the similarity between the subject and the model.

The feature extraction and matching process may be iterated over other subject-independent (like the eyebrows) or subject-dependent features, thus increasing the reliability of the recognition process.

#### 4.3 Evaluation of matching techniques

The described matching algorithm has been tested on several databases, including the FERET database [6], the Olivetti Research Laboratory database [54] and other databases acquired in our laboratory [14,38].

More recently, the face matching algorithm has been tested against very unfavourable conditions, which allowed us to determine both the lower limit in the performances and also to understand the most critical factors in iconic-based face matching. Two similar databases have been recorded. The former is an "industrial" database: the images have been acquired from a camera placed at the entrance of an industry working area. The subjects were completely unaware of the camera and, as a consequence, the faces within the pictures are quite different in size, facial expression, orientation and pose. The entire set is composed of 488 gray level images (8 bits per pixel) from 75 subjects, the image size is 384x288 pixels and the head size is always smaller than 80x120 pixels. The total number of comparison is 238144. Most of the pictures from the same subject were acquired at least days or even weeks apart. The latter face database has been acquired in our laboratory from aware subjects but with a limited cooperation. The database is composed of 124 images, with a resolution of 512x512 pixels, acquired from 45 subjects. The total number of comparison is 15376. Also

<sup>2</sup> To compute the difference, the gray levels of each log-polar fixation are first normalized to the range of intensity values of the corresponding facial feature of the reference subject.



Figure 7: Sample images from the industrial database (top) and from the database acquired in our laboratory (bottom).

in this case the pictures from the same subject have been acquired days or weeks a part.

The evaluation procedure involves a complete matching over all the images in the data set (all subjects versus all images) Given  $M$  images for each of  $N$  subjects, the results obtained are divided into two classes<sup>3</sup> (assuming that more than a single score is available for each image comparison):

- matching scores obtained comparing all different images of the same subject, equal to  $N \times M \times (M - 1)$  comparisons (*client tests*);
- matching scores obtained comparing all different images of different subjects, equal to  $N \times M^2 - N \times (N + M - 1)$  comparisons (*impostor tests*);

A covariance matrix is defined describing each of the two classes:

$$\Sigma_i = \frac{1}{N_i} \sum_{j=1}^{N_i} (x_j^i - m_i)(x_j^i - m_i)^T \quad (4)$$

where  $N_i$  represents the number of elements of class “ $i$ ” and  $m_i$  is the mean value of the same class. Given the entire ensemble of matching scores for the two classes (each score can be regarded as a vector within the class), the discrimination power can be defined through six statistical indexes: the intraset and interset distances (class separability indexes); the Bayesian error probability; the false acceptance, false rejection and the equal error rate (FAR, FRR, EER).

The first two indexes define the distances among the elements of the same class and between the two classes. By comparing the two it is possible to define the separability

<sup>3</sup> As a consequence of the experimental procedure the training set and the test set are disjoint, except for the case where the image used to build the representation of one subject is also used for an impostor test.

between the two classes, e.g. to discriminate the set of clients from all the impostors [52-54].

From the reported experiments it turns out that the system’s performances are greatly influenced by the accuracy in the estimation of the position of facial features. Two different methods were tested for feature detection: the former based on matching a generic template of the facial features, the latter applying a specific template extracted from the image of the model face, which correspond to the actual facial features of the subject to be recognized. The second approach effectively maximizes the probability of correct feature localization for the subject to be recognized.

In order to determine the classification power of the matching algorithm, a simple recognition experiment has been attempted. Given a model image for each subject (or client), all other images matched with the five highest correlation scores (the same model image has not been matched with itself) were ranked. By applying a simple voting procedure, where the subject corresponding to the highest number of instances within the selected five is chosen, a remarkable 98% recognition rate was obtained. The remaining 2% was composed of subjects were all five selected images were from different individuals but still included at least one image of the client. This result confirms the need for a two-step matching strategy where an iconic matcher is firstly used to perform a rough separation between the two classes (clients and impostors) and selecting a set of possible candidates. The following stage can be an independent measurement, for example introducing more fixations, but performed only on the image set already selected.

	Subject-based template	Generic template	Commercial system based on LFA
$P_e$	30.5%	36.5%	31.5%
FAR	15%	25%	19.68%
FRR	17%	21.3%	20.34%

**Table 1.** Performances of the face matching system on the “industrial database”. The two columns in the middle report the computed error probability and the false acceptance/rejection with two different methods for the extraction of the facial features. As a reference, the last column reports the performances of a highly reputed commercial system.

## 5. CONCLUSION

The analysis of faces is one of the most important visual processes in humans. It has been shown that several areas in the brain are devoted to the selection and extraction of face-related features for recognition as well as other reasoning processes. In particular, parts of the inferior and medial temporal cortex may work together to process faces. The model of face analysis, as learned from neurophysiology, can be applied as the basic scheme for an information system to process face images.

An iconic matching algorithm, based on the analysis of space-variant fixations, has been considered and stressed to the worst possible working conditions by using ad-hoc databases. From the error indices reported in this, as well as

previous papers [52,53], it is possible to make the following considerations:

- The face registration and/or warping step is quite critical for the entire recognition process.
- The performances of an identity verification system can not be assessed with a trial on a data set only, but multiple and different data set are necessary to understand the real failing modes.
- There is an intrinsic limit, also statistically demonstrated, in using a matching technique alone to discriminate a set of given clients from the class of all possible impostors.

The clear limitations of a single matching engine in making a discrimination among thousands of images, enforces the need for either a multi-level or a multi-algorithmic process, where several (at least two) cooperating “experts” are applied to the same authentication process.

Several aspects have been addressed explicitly but many are still under investigation. In this paper the analysis has been restricted to a simple image representation, but further work can be done by matching multiple features either at the first level or at the second processing level.

#### REFERENCES

- [1] R. Chellappa, C.L. Wilson, and S. Sirohey. “Human and machine recognition of faces: A survey”. Proceedings of the IEEE, Vol. 83, pp 705-740, 1995.
- [2] T. Kanade. Computer recognition of human faces. Birkhauser, Basel and Stuttgart, 1977.
- [3] P. Sinha and T. Poggio. “I think I know that face...”. Nature, Vol. 384, pp. 404, 1996.
- [4] L. Wiskott, J.M. Fellous, N. Kruger and C. von der Malsburg. “Face recognition and gender determination”. In Proceedings Int.I Workshop on Automatic Face and Gesture Recognition, pp. 92-97, Zurich, Switzerland, 1995.
- [5] Anil K. Jain, Ruud Bolle, Sharath Pankanti. Biometrics, Personal Identification in Networked Society, Kluwer Academic Publishers, 1999
- [6] H. Wechsler, P. Phillips, V. Bruce, F. Soulie, and T. Huang (Eds.), Face Recognition. From Theory to Applications, NATO ASI Series F, Vol. 163, Springer-Verlag, Berlin Heidelberg.
- [7] G. Cottrell and J. Metcalfe. “Face, gender and emotion recognition using holons”. In D. Touretzky, editor, Advances in Neural Information Processing Systems, Vol. 3, 564-571, San Mateo, CA, 1991. Morgan Kaufmann.
- [8] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman. “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection”. IEEE Trans. on PAMI, vol. 19, no. 7, 711-20, 1997.
- [9] G.J. Edwards, T.F. Cootes and C.J. Taylor. “Face Recognition Using Active Appearance Models”. In Proc. of 5th European Conference on Computer Vision Springer Verlag,, 581-95, 1998.
- [10] R. Brunelli and T. Poggio. “Face recognition through geometrical features”. In Proc. of 2nd European Conference on Computer Vision, S. Margherita Ligure (Italy), Springer Verlag, 792-800, 1992.
- [11] B. Mogbaddam, T. Jebara and A. Pentland. “Bayesian face recognition”. Pattern Recognition, vol. 33, no. 11, 1771-82, 2000.
- [12] G.C. Feng, P.C. Yuen, and D.Q. Dai. “Human face recognition using pca on wavelet subband”. Journal of Electronic Imaging, vol. 9, no. 2, 226-33, 2000.
- [13] C. Liu and H. Wechsler. “Evolutionary pursuit and its application to face recognition”. IEEE Transactions on PAMI, vol. 22, no. 6, 570-582, 2000.
- [14] M. Tistarelli and E. Grosso. “Active vision-based face authentication”. Image and Vision Computing: Special issue on Facial Image Analysis, M. Tistarelli ed., vol. 18, no. 4, 299-314, 2000.
- [15] C. Darwin. The expression of the emotions in man and animals. London, U.K.: John Murray, 1965. (Original work published 1872)
- [16] C. Goren, M. Sarty and P. Wu. “Visual following and pattern discrimination of face-like stimuli by newborn infants”. Pediatrics, vol. 56, 544–549, 1975.
- [17] G. E. Walton and T. G. R. Bower. “Newborns form “prototypes” in less than 1 minute”. Psychological Science, vol. 4, 203–205, 1993.
- [18] J. Fagan. “Infants’ recognition memory for face”. Journal of Experimental Child Psychology, vol. 14, 453–476, 1972.
- [19] M. de Haan and C. A. Nelson. “Recognition of the mother’s face by 6-month-old infants: A neurobehavioral study”. Child Development, vol. 68, 187–210, 1997.
- [20] C. M. Leonard, E. T. Rolls, F. A. W. Wilson and G. C. Baylis. “Neurons in the amygdala of the monkey with responses selective for faces”. Behavioral Brain Research, vol. 15, 159–176, 1985.
- [21] Gauthier, M. J. Tarr, A. W. Anderson, P. Skudlarski and J. C. Gore. “Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects”. Nature Neuroscience, vol. 2, 568–573, 1999.
- [22] N. Kanwisher, J. McDermott and M. M. Chun. “The fusiform face area: A module in human extrastriate cortex specialized for face perception”. Journal of Neuroscience, vol. 17, 4302–4311, 1997.
- [23] G. McCarthy, A. Puce, J. C. Gore and T. Allison. “Facespecific processing in the human fusiform gyrus”. Journal of Cognitive Neuroscience, vol. 8, 605–610, 1997.
- [24] R. T. Schultz, I. Gauthier, A. Klin, R. K. Fulbright, A. W. Anderson, F. R. Volkmar, P. Skudlarski, C. Lacadie, D. J. Cohen and J. C. Gore. “Abnormal ventral temporal cortical activity during face discrimination among

- individuals with autism and Asperger syndrome". *Archives of General Psychiatry*, vol. 57, 331–340, 2000.
- [25] A. R. Damasio, J. Damasio and G. W. Van Hoesen. "Prosopagnosia: Anatomic basis and behavioral mechanisms". *Neurology*, vol. 32, 331–341, 1982.
- [26] C. A. Nelson. "The development and neural bases of face recognition". *Infant and Child Development*, vol. 10, 3-18, 2001.
- [27] J. P. Aggleton, M. J. Burton and R. E. Passingham. "Cortical and subcortical afferents to the amygdala of the rhesus monkey (*Macaca mulatta*)". *Brain Research*, vol. 190, 347–368, 1980.
- [28] J. Shepherd. "Social factors in face recognition". In G. Davies, H. Ellis & J. Shepherd (Eds.), *Perceiving and remembering face*, London: Academic Press, 55–79, 1981.
- [29] L. Yarbus. *Eye movements and vision*. New York: Plenum Press, 1967.
- [30] F. K. D. Nahm, A. Perret, D. Amaral, and T. D. Albright. "How do monkeys look at faces?" *Journal of Cognitive Neuroscience*, 9, 611–623, 1997.
- [31] M. M. Haith, T. Bergman, and M. J. Moore. "Eye contact and face scanning in early infancy". *Science*, vol. 198, 853–854, 1979.
- [32] A. Klin. "Eye-tracking of social stimuli in adults with autism". NICHD Collaborative Program of Excellence in Autism. Yale University, New Haven, CT, May 2001.
- [33] B. Braathen and M. S. Bartlett and G. Littlewort and J. R. Movellan. "First Steps Towards Automatic Recognition of Spontaneous Facial Action Units" *ACM Workshop on Perceptive User Interfaces*, Orlando (FL), Nov. 15-16 2001.
- [34] R.W. Picard. "Toward computers that recognize and respond to user emotion". *IBM System*, (39), 3/4, 2000.
- [35] R.W. Picard. "Building HAL: Computers that sense, recognize, and respond to human emotion". MIT Media-Lab TR-532. Also in *Society of Photo-Optical Instrumentation Engineers. Human Vision and Electronic Imaging VI*, part of IS&T/SPIE9s Photonics West, 2001.
- [36] D.H. Ballard. "Animate vision". *Artificial Intelligence*, vol. 48, pp. 57-86, 1991.
- [37] Y. Aloimonos (Ed.). "Purposive, qualitative, active vision". *CVGIP: Image Understanding*, 56(special issue on qualitative, active vision), Vol. 56, July 1992.
- [38] M. Tistarelli. "Active/space-variant object recognition". *Image and Vision Computing*, vol. 13, no. 3, 215-226, 1995.
- [39] E. L. Schwartz, D. N. Greve, and G. Bonmassar. "Space-variant active vision: definition, overview and examples". *Neural Networks*, vol. 8, no. 7/8, pp. 1297-1308, 1995.
- [40] C. A. Curcio, K. R. Sloan, R. E. Kalina, A. E. Hendrickson. "Human photoreceptor topography". *J Comp Neurol.*, vol. 292, no. 4, 497-523, 1990.
- [41] G. Sandini, G. Metta. "Retina-like sensors: motivations, technology and applications". In *Sensors and Sensing in Biology and Engineering*. T.W. Secomb, F. Barth, and P. Humphrey (Eds). Springer-Verlag. 2002
- [42] P. J. Burt, *Smart sensing in machine vision*". In *Machine Vision: Algorithms, Architectures, and Systems*, Academic Press, 1988.
- [43] F. Tong, Ze-Nian Li. "Reciprocal-wedge transform for space-variant sensing". *IEEE Transactions on PAMI*, Vol. 17, 500-511, 1995
- [44] E. L. Schwartz. "Spatial mapping in the primate sensory projection: Analytic structure and relevance to perception", *Biological Cybernetics*, vol. 25, 181-194, 1977.
- [45] T.E. Fisher and R.D. Juday. "A programmable video image remapper", In *Proc. of SPIE*, vol. 938, 122-128, 1988
- [46] E. Grosso, M. Tistarelli. "Log-polar Stereo for Anthropomorphic Robots". In *Proc. of 6th European Conference on Computer Vision, LNCS 1842*, Springer Verlag, 299-313, 2000
- [47] G. Sandini and M. Tistarelli. "Vision and space-variant sensing". In H. Wechsler (Ed), *Neural Networks for Perception: Human and Machine Perception*, Academic Press, 1991
- [48] Y. Yeshurun and E.L. Schwartz. Shape description with a space-variant sensor: Algorithms for scan-path, fusion and convergence over multiple scans". *IEEE Transactions on PAMI*, vol. 11, 1217-1222, Nov. 1989
- [49] J. Daugman. "High confidence visual recognition of persons by a test of statistical independence". *IEEE Transactions on PAMI*, vol. 15, no. 11, 1148-1161, 1993
- [50] J. Bigun. "Retinal vision applied to facial features detection and face authentication". *Pattern Recognition Letters*, vol. 23, no. 4, 463-475, 1997
- [51] A. K. Jain, S. Prabhakar, L. Hong, S. Pankanti. "Filterbank-based fingerprint matching". *IEEE Transactions on IP*, vol. 9, no. 5, 846-859, 2000
- [52] M. Tistarelli, A. Lagorio, and E. Grosso. "Understanding Iconic Image-Based Face Biometrics". *Proc. of Int.l Workshop on Biometric Authentication*, Copenhagen, Denmark, Springer Verlag, LNCS 2359, 19-29, 2002
- [53] M. Tistarelli, A. Lagorio, and E. Grosso. "Image-Based techniques for biometric authentication". *Proc. of Int.l Conference on Multimedia and Expo, special session on Biometrics: New Challenges for User Authentication*, Lausanne, Switzerland, August 22-26, 2002
- [54] Grother P.J. "Cross validation comparison of NIST ocr databases". In *Proceedings of the SPIE*, volume 1906, 296-307, 1993
- [55] Samaria, F., Harter, A.: Parameterisation of a Stochastic Model for Human Face Identification. In *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota, FL. IEEE Computer Society Press, 1994