Pattern Recognition Letters 31 (2010) 857–868

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Face and iris localization using templates designed by particle swarm optimization

Claudio A. Perez *, Carlos M. Aravena, Juan I. Vallejos, Pablo A. Estevez, Claudio M. Held

Department of Electrical Engineering and Advanced Mining Technology Center, Universidad de Chile, Av. Tupper 2007, Santiago, Chile

ARTICLE INFO

Article history:
Received 9 February 2009
Received in revised form 11 December 2009
Available online 4 January 2010

Communicated by J. Yang

Keywords:
Face localization
Eye localization
PSO templates
Anthropometric templates
Real-time face localization
Real-time eye localization

ABSTRACT

Face and iris localization is one of the most active research areas in image understanding for new applications in security and theft prevention, as well as in the development of human–machine interfaces. In the past, several methods for real-time face localization have been developed using face anthropometric templates which include face features such as eyes, eyebrows, nose and mouth. It has been shown that accuracy in face and iris localization is crucial to face recognition algorithms. An error of a few pixels in face or iris localization will produce significant reduction in face recognition rates. In this paper, we present a new method based on particle swarm optimization (PSO) to generate templates for frontal face localization in real time. The PSO templates were tested for face localization on the Yale B Face Database and compared to other methods based on anthropometric templates and Adaboost. Additionally, the PSO templates were compared in iris localization to a method using combined binary edge and intensity information in two subsets of the AR face database, and to a method based on SVM classifiers in a subset of the FERET database. Results show that the PSO templates exhibit better spatial selectivity for frontal faces resulting in a better performance in face localization and face size estimation. Correct face localization reached a rate of 97.4% on Yale B which was higher than 96.2% obtained with the anthropometric templates and much better than 60.5% obtained with the Adaboost face detection method. On the AR face subsets, different disparity errors were considered and for the smallest error, a 100% correct detection was reached in the AR-63 subset and 99.7% was obtained in the AR-564 subset. On the FERET subset a detection rate of 96.6% was achieved using the same criteria. In contrast to the Adaboost method, PSO templates were able to localize faces on high-contrast or poorly illuminated environments. Additionally, in comparison with the anthropometric templates, the PSO templates have fewer pixels, resulting in a 40% reduction in processing time thus making them more appropriate for real-time applications.

1. Introduction

Real-time face localization is among the most active research areas in image understanding because of new applications in human–machine interfaces, gesture recognition, face recognition, surveillance, user verification and theft prevention (Le and Satoh, 2006; Yang et al., 2002; Perez et al., 2007, 2005, 2003; Chiang et al., 2003). Also, face localization is one of the first steps in many applications involving gaze estimation to aid handicapped individuals or to control devices in complex technological environments such as airplane cockpits, hospital operating rooms and industrial control units (Park and Lim, 2001; Sirohey et al., 2002; Wang et al., 2002; Yang et al., 2002). In addition, the biometrics world market is predicted to almost triple in the next five years and face recognition products represent 11.4% of the total market (International Biometric Group, 2008). Face recognition is an important segment of the biometrics market and it has been shown that increased error in face or iris localization will yield lower face recognition rates (Campadelli et al., 2007). In (Campadelli et al., 2007) using the XM2VTS face database it is shown that precision for iris localization is critical for face recognition. Localization errors of 0.05 (maximum iris detection error divided by distance between irises) resulted in a face recognition drop of nearly 20% for several different methods. Therefore, it is very important to find better face and iris localization methods to improve face recognition rates. Accurate iris localization refers to the error minimization when detecting the eye position, i.e., error as the distance between the ground truth and the detected position of the eye. The minimization of the error in iris localization is essential in many face recognition algorithms. Fig. 1 shows the recognition rate as a function of the standard deviation of the iris localization error (noise level) in the FERET database for the four subsets (Fb, Fc, Dup1 and Dup2). The face recognition algorithm is based on Gabor jets (Zou et al., 2005), and was selected because it reaches one of the highest recognition rates in the FERET database published up to date. It can be

* Corresponding author. Address: Image Processing Laboratory, Department of Electrical Engineering, Universidad de Chile, Casilla 412-3, Av. Tupper 2007, Santiago, Chile. Tel.: +56 2 978 4196; fax: +56 2 695 3881.
E-mail address: clperez@ing.uchile.cl (C.A. Perez).

0167-8655/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2009.12.029

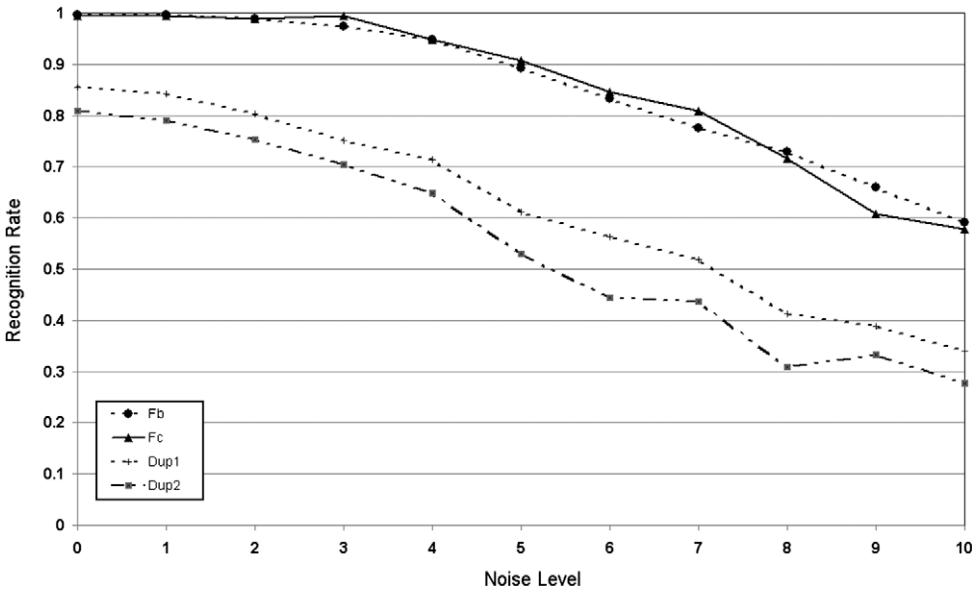

Fig. 1. Recognition rate on the FERET database as a function of iris localization error.

observed that an error of 5 pixels (STD) produces a decrease in recognition rate between 10% and 25% on the different subsets of the FERET database. These results emphasize the importance of accurate iris detection for face recognition.

Methods for face localization have been classified into knowledge based, feature invariant, template matching and appearance based (Yang et al., 2002). Template-based face localization has proven to be accurate and appropriate for real-time applications in man–machine interfaces (Maio and Maltoni, 2000; Jin et al., 2007; Kato et al., 2007; Phimoltares et al., 2007; Mohanty et al., 2006; Perez et al., 2007; Ji, 2002; Ying et al., 2001). In previous research the original method to detect faces using templates (Maio and Maltoni, 2000) was extended to detect rotated faces and irises in real time (Perez et al., 2007, 2009). A crucial part of this method computes a line integral of a face template over a directional image. A set of face templates is used to compute the line integral to take into account face size (Maio and Maltoni, 2000; Perez et al., 2007; Jin et al., 2007). A set of templates was built to consider face anthropometric features such as eyes, eyebrows, nose, mouth and lower part of the chin. These face features provide edges mainly along horizontal or vertical directions within the face. Eyes, eyebrows and mouth regions were assigned 0°, nose was assigned 90°, and chin was assigned a variable angle according to the tangent vector to the chin contour capturing the face elliptical shape. The upper part of the face was not considered in the template since the hair style frequently introduces great variance in shape. These templates are called anthropometric templates because they include the described distinctive face features (Perez et al., 2007, 2009). Although the results of this method are good, specially in real-time applications, there is still space for face and iris localization improvements. Therefore, a natural next step was to develop a method to generate new templates that could improve the present face and iris localization results and reduce computational time.

In the present paper, we present the research to design an improved template for face localization. Preliminary results were obtained using PSO to design templates in (Perez and Vallejos, 2006). The method was applied on faces of the same person to detect his/her face. Faces were not aligned and therefore an error resulted for different persons. In the present paper faces are aligned using the eyes as references. This improves significantly the PSO algorithm convergence. Due to this improvement the PSO algorithm can be now applied to any set of faces from different persons. The initial population of the PSO algorithm was increased and stop criteria was also modified based on the fitness value instead of number of iterations. In (Perez and Vallejos, 2006) the method was assessed only by visual inspection without an objective to measure face or iris localization. In the present paper, we include objective measures for face and iris localization errors. In the present paper, we added the iris localization stage which is compared to other methods. The original method was applied only to two video sequences and now we added more video sequences, faces from Yale B face database, AR face database and FERET database. We also include a comparison with three other methods including Adaboost-based method, a combined binary edges and intensity method, and a SVM-based method.

The search for these templates involves a high dimensional space and therefore, an exhaustive search is not possible. Even for the case of the smallest template in the proposed method with a fixed number of points, the number of computations involved in searching for all possible solutions is huge. Eq. (1) shows the size of the search space for the smallest template of 30×40 pixels and for a fixed number of 200 points

$$\text{Number of templates} = \binom{1200}{200} = \frac{1200!}{200!(1200-200)!}. \quad (1)$$

The fact that face templates are usually larger than 30×40 pixels and that a variable number of points are allowed in each template, justifies the choice of a heuristic algorithm to guide the search. The proposed method for template optimization uses the particle swarm optimization (PSO) algorithm because of its proven effectiveness in heuristic search and relatively small computational time required (Eberhart and Kennedy, 1995). The PSO algorithm was originally introduced by Eberhart and Kennedy (1995). Several papers have expanded the original algorithm and analyzed the convergence and stability (Clerc and Kennedy, 2002; Chatterjee and Siarry, 2006; Eberhart and Shi, 1998, 1999; Shi and Eberhart, 1998; Trelea, 2003; Zhang et al., 2005).

In this paper, we use the particle swarm optimization (PSO) algorithm to generate new templates to maximize the line integral value while using a small number of points to decrease the computational time. Face size estimation, spatial selectivity, face and iris localization rates are computed.

2. Methodology

2.1. General context for template-based face detection

The directional image contains the main gradient angles extracted from the frontal face. The template is employed to compute a line integral over the directional image. An ideal template should produce a large detection value when it is located right over the subject's face and a low value elsewhere. As the directional image changes with face gestures, face movements and illumination variations, the face template should represent the main features that remain rather constant on the face. In the previous work, anthropometric templates were created using this assumption including information about eyes, eyebrows, nose, mouth and lower part of the chin (Perez et al., 2007).

2.2. PSO template generation

The PSO algorithm is a relatively new combinatorial heuristic algorithm based on the interaction of social systems such as fish schooling and bird flocking. Since the introduction of the PSO algorithm, several improvements have been proposed introducing new parameters for intensive search as well as to avoid solution combinatorial explosion (Clerc and Kennedy, 2002; Shi and Eberhart, 1998).

The PSO algorithm simulates the individual (a particle) social behavior moving in a multidimensional space. Each solution is identified with an individual particle which has specific coordinates in the search space. Each particle has a position (X_i) and a speed (V_i) in the multidimensional space. The particle position indicates the possible solution in the multidimensional space and the speed indicates the amount of change between the current position and the previous one. The algorithm also stores the previous best position (P_i) of each particle. This information is used to adapt the particle position in the solution space for the next iteration. At initialization, the algorithm generates a random particle population and each particle is initialized with a random position and speed. The speed of each particle is initialized in the range $[-V_{\max}, V_{\max}]$, where $(V_{\max})$ is the maximum speed of the particles. After the initial population has been generated, a fitness function is evaluated for each particle position. The current result of the fitness function is compared to the best previous result for each particle. If the current result is higher, the best previous particle is replaced by the current one $(P_i = X_i)$.

Then for each particle (X_i), the index $g(i)$ of the particle with the best previous fitness value within a neighborhood around the particle is determined. Particles are modified according to:

$$V_{ij}(t+1) = w \cdot V_{ij}(t) + \varphi_1 \cdot (P_{ij} - X_{ij}(t)) + \varphi_2 \cdot (P_{g(i)j} - X_{ij}(t)), \quad (2)$$

$$X_{ij}(t+1) = X_{ij}(t) + V_{ij}(t+1), \quad (3)$$

where φ_1 and φ_2 are random numbers in the range $[0 \cdots \varphi_{\max}]$. Generally $\varphi_{\max}$, is equal to 2, but can change with different implementations. In this equation, w is the inertia parameter (Shi and Eberhart, 1998).

The PSO method is applied to frontal faces to generate a set of templates for different face sizes. One alternative for this method is to use faces from different subjects to generate a generic PSO template. Another alternative is to create personalized templates by using a set of faces from the same subject to create the PSO templates. Fig. 2 shows a block diagram of the template generation by the PSO algorithm.

2.2.1. Preprocessing

The method needs a small set of segmented frontal face images in order to apply the PSO algorithm. Therefore, the first step to generate the PSO templates is to segment the frontal face by determining the parameters (x_1, x_2, y_1, y_2, T) shown in Fig. 3. The vertical face position and size are within the interval $[y_1 - \Delta y_1, y_2 + \Delta y_2]$ and the horizontal face position and size are within the interval $[x_1, x_2]$. On the vertical axis y, y_1 corresponds to the center of the eyebrows and y_2 to the center of the mouth. Variables x_1 and x_2 are the distances to the left and right borders of the face, respectively, with the nose located at the center of the face. Δy_1 is the distance between the center of the eyebrows and the upper limit of the face. Δy_2 is the distance between the center of the mouth and the lower limit of the face. The latter two are computed by

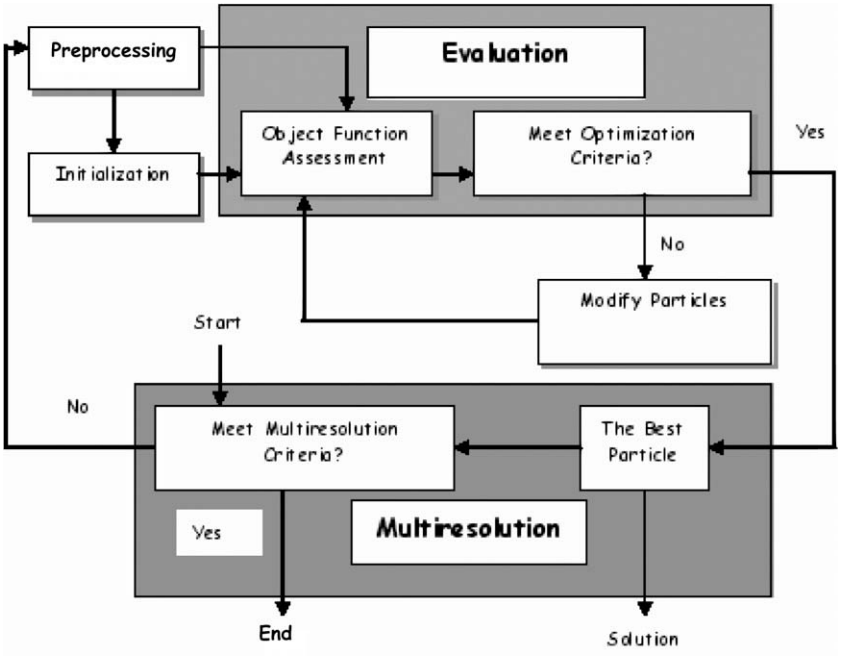

Fig. 2. Block diagram for the PSO template generation method.

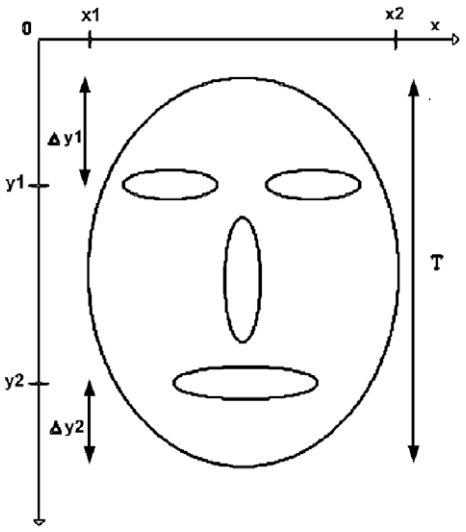

Fig. 3. Parameters for manual frontal face segmentation.

using anthropometric measurements. The segmented faces are converted to gray scale and normalized in size. Then, a directional image is computed (Maio and Maltoni, 2000; Perez et al., 2007) as shown in Fig. 4. The directional image contains the average of the tangent vectors in a 3×3 window in the normalized gray scale segmented image. A larger window could also be used to calculate the directional image to obtain a coarse template of the face, improving computation time but lowering accuracy on the final template set. Considering different face sizes, the resulting set of directional images is used by the PSO algorithm. Eye positions are used to align faces in order to make the PSO algorithm converge to a solution.

2.2.2. Template initialization

In the method proposed in this paper, each pixel in the template represents an angle between 0° and 180°, quantized in steps given by $180/w_d$, where $(w_d) = 36$ resulting in steps of 5°. Angle information is then stored for each template pixel in values ranging from 0 to 36. Therefore, pixel values above 36 do not belong to the template. This angle discretization is not a critical parameter. A larger step size in the angle discretization provides faster convergence at the expense of less precise angular fitting on the final template. For the PSO template optimization, the position of each particle is defined by a vector of size N equal to the number of pixels of the template. The range for each pixel in the template $(X_{i,l,m})$ depends on the quantization used in the directional image $(w_d) = 36$ and on a margin, $[m_R]$, that allows to define regions of the particle position that are not a part of the template. This range is given by $[0, w_d + 2 \cdot m_R]$. A component of the vector representing the particle position that falls within the allowed range will be present in the PSO template, otherwise it will not belong to the template. Only components of the vector that are part of the template are considered for the line integral computation (fitness). Fig. 5 and Eq. (4) show the range where a PSO template pixel or the vector component for the particle position will be considered to belong to the PSO template. $IM_{i,l,m}$ represents the module of point (l, m) of particle (X_i) and $MM_{i,l,m}$ is the angle at point (l, m) of particle (X_i).

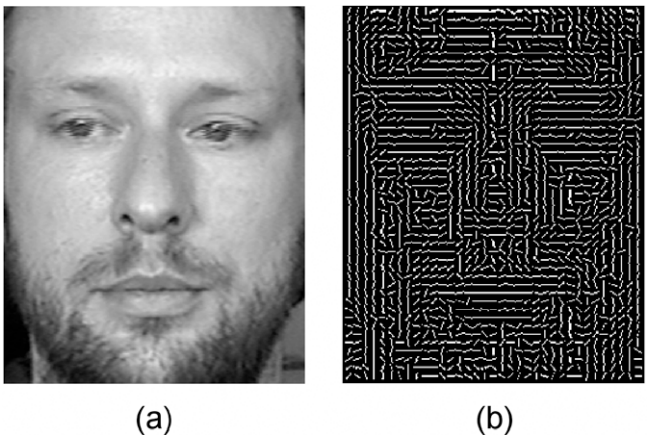

Fig. 4. (a) Segmented frontal face; (b) directional image of (a).

For each vector component the allowed or not allowed ranges are defined as follows (4):

$$(IM_{i,l,m}, MM_{i,l,m}) = \begin{cases} (1, X_{i,l,m} - m_R) & X_{i,l,m} \in [m_R, w_d + m_R], \\ (0, 0) & X_{i,l,m} \notin [m_R, w_d + m_R]. \end{cases} \quad (4)$$

Speed initialization is performed randomly in the range $[-w_d, w_d]$. Position of each particle is initialized by assigning a random number for each pixel in the interval $[0, w_d + 2 \cdot m_R]$. Speed is also initialized by assigning a random number for each pixel in the range $[-w_d, w_d]$. A modification of a particle position will change the value of each PSO template pixel, changing its angular information. This angular information may take a value within the interval $[m_R, w_d + m_R]$, or may take a value out of the allowed range, disappearing from the PSO template. The speed of each particle modifies each particle position according to Eq. (13). Particle size depends on the size of the template to be generated.

2.2.3. Evaluation

A fitness function is defined and computed for each particle. Since the line integral value provides a measure of similitude between the template and the directional image (Perez et al., 2007), the fitness function is the line integral value of the template over the face directional image. The line integral f_i^k for particle i is given by (5)

$$f_i^k = \frac{1}{N_R \cdot PM_i} \sum_{h=0}^{N_R-1} \sum_{l=0}^{N_w} \sum_{m=0}^{N_H} IM_{i,l,m} \cdot \left(I_{ref_{h,l,m}} \cdot \alpha_{i,l,h,m} - (1 - I_{ref_{h,l,m}}) \cdot 90^\circ \right). \quad (5)$$

In Eq. (5), $I_{ref_{h,l,m}}$ indicates that the component of the face directional image, h, at position (l, m) is present. N_R is the number of frontal faces, N_w is the particle width, N_H is the particle height, and PM_i is the total number of points of particle X_i. Particle X_i represents a template. The index k indicates the iteration number. Factor $\alpha_{i,l,m,h}$, called "angular similarity" (Perez et al., 2007), corresponds to the angle difference at a particular coordinate in the PSO template and the angle measured at the directional image I_α from face h. This factor is defined as,

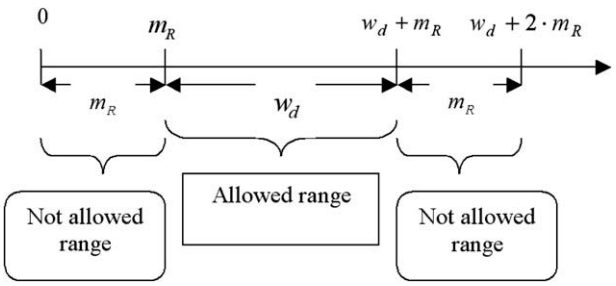

Fig. 5. Particles initialization range where PSO template pixels (or vector components) are allowed.

$$\alpha_{i,l,m,h} = 90^\circ - 2 \cdot \min\left(\left|MM_{i,l,m} - I_{\alpha_{h,l,m}}\right|,\ 180^\circ - \left|MM_{i,l,m} - I_{\alpha_{l,m,h}}\right|\right), \quad (6)$$

where $I_{\alpha_{l,m,h}}$ = angle at point (l,m) for face h.

The line integral values in each template position (l,m) for each particle are stored for the next iteration as follows,

$$f_{i,l,m}^k = \frac{\sum_{h=0}^{N_R-1}\left(I_{ref_{h,l,m}} \cdot \alpha_{i,l,h,m} - (1 - I_{ref_{h,l,m}}) \cdot 90^\circ\right)}{N_R}. \quad (7)$$

The line integral value computed for each particle is penalized by a factor that takes into account the total number of points in each particle, i.e., the total number of points in the PSO template. This factor provides a preference for those templates with fewer points and therefore the time required in the line integral computation will be lower. Computational time is very important in real-time applications.

Finally if the stop criterion is satisfied, the algorithm stops and goes to the Multiresolution stage where the next template size is optimized, otherwise it iterates until the criterion is met. When all template sizes are computed the algorithm stops. The following criteria are used:

- *Threshold value:* If the line integral value is above the threshold the algorithm stops. The maximum value for the line integral is 18 and the threshold was set at 15 which is a value above the line integral value obtained for the anthropometric templates (Perez et al., 2007).
- *Maximum number of iterations:* The algorithm stops when the maximum number of iterations has been reached.

2.2.4. Particle modification

The particle position and velocity are modified according to the following:

- *Best particle:* For each particle, if the line integral value (f_i^k) in the current iteration k is negative, the best particle is modified only if the components of the line integral $(f_{i,l,m}^k)$ are greater than the previous values, i.e.,

$$\text{if } f_{i,l,m}^k \geqslant \max_{n=0\ldots k-1}\left\{f_{i,l,m}^k\right\} \Rightarrow P_{i,l,m} = X_{i,l,m}. \quad (8)$$

  If the actual line integral value (f_i^k) is positive, the individual terms of the best particle are modified only if the components of the line integral $(f_{i,l,m}^k)$ are greater than the previous values or if these values are below the threshold T_{PI} as follows:

$$\text{if } f_{i,l,m}^k \geqslant \max_{n=0\ldots k-1}\left\{f_{i,l,m}^k\right\} \quad \text{or} \quad f_{i,l,m}^k \leqslant T_{PI} \Rightarrow P_{i,l,m} = X_{i,l,m}. \quad (9)$$

  Threshold value T_{PI}, allows for a faster convergence of the PSO template and improves template quality by filtering out irrelevant points.
- *Best particle within a neighborhood*: For each particle the line integral values are compared within a neighborhood of the particle. The index of the best particle within the neighborhood is stored as $g(i)$. T indicates the neighborhood size.

$$g(i) = \max_{l=i-T\ldots i+T}\left\{\max_{n=0\ldots k}\left\{f_l^k\right\}\right\}. \quad (10)$$

- *Inertia parameter:* It was chosen as a decreasing function of the line integral value to improve convergence. The inertia parameter as proposed in (Chatterjee and Siarry, 2006; Eberhart and Shi, 1998; Shi and Eberhart, 1998) is

$$w = \left[1 - \frac{f_i^k}{52}\right] \cdot \Gamma, \quad (11)$$

where Γ takes alternating values with a period of 15 iterations as $\Gamma = 0.5$ for local search and $\Gamma = 1.5$ for global search. A non-linear decreasing function has been used previously for Γ in (Chatterjee and Siarry, 2006; Eberhart and Shi, 1999).

The particles are changed according to

$$V_{i,l,m}(t+1) = w \cdot V_{i,l,m}(t) + \varphi_1 \cdot (P_{i,l,m} - X_{i,l,m}(t)) + \varphi_2 \cdot (P_{g(i),l,m} - X_{i,l,m}(t)), \quad (12)$$

$$X_{i,l,m}(t+1) = X_{i,l,m}(t) + V_{i,l,m}(t+1). \quad (13)$$

2.2.5. Multiresolution

In this stage, the PSO template size is increased and parameters are changed to transform the scale in order to have templates for different face sizes. The best template for each size is stored after this stage. Templates are optimized for face sizes within the range 120–240 pixels height with a step of 2 pixels. However, angular information stored in the templates is compressed in a factor given by the directional image used (Maio and Maltoni, 2000; Perez et al., 2007), resulting in 61 templates that fall within the range 30×40 pixels to 60×80 pixels. Larger template size requires longer computational time on the PSO optimization. As an example, on a computer Athlon64 3000+, 1 GB RAM, the smallest template (30×40) takes approximately 1.5 min to optimize, while the largest template (60×80) takes 15 min.

2.2.6. Iris localization

Iris localization uses the same method proposed in our previous work with anthropometric templates (Perez et al., 2007). Based on the face anthropometry the region where the eyes are most probably located is determined . An iris template is then applied in that region that searches for the iris–sclera border. The method has shown great accuracy in iris localization (Perez et al., 2007).

2.3. Metrics

2.3.1. Measuring face detection rate

The face localization method determines a rectangular box containing the face. Boxes are represented by their center location (x,y), width t and height $t/0.75$. In order to measure the correctness of face localization, the output box is compared against a manually marked box, ground truth, for each face. Two different criteria have to be accomplished in order to classify a localized face as correct (Verschae and del Solar, 2003; Verschae et al., 2008):

1. Distance between face centers has to be less than P times the sum of both widths. P was chosen equal to 0.25

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} < P \cdot (t_1 + t_2). \quad (14)$$

2. An area-based criterion establishes that the localized face width should not double that of the ground truth

$$(t_1 < t_2 \wedge t_2 < 2 \cdot t_1) \vee (t_2 < t_1 \wedge t_1 < 2 \cdot t_2). \quad (15)$$

Fig. 6 shows the ground-truth box (white) and a localized face box (black). In this case, the localized face box does not meet the distance between both centers' criteria.

2.3.2. Measuring iris localization

To measure iris localization error, we used the same metrics applied in (Song et al., 2006). The relative error is then defined by:

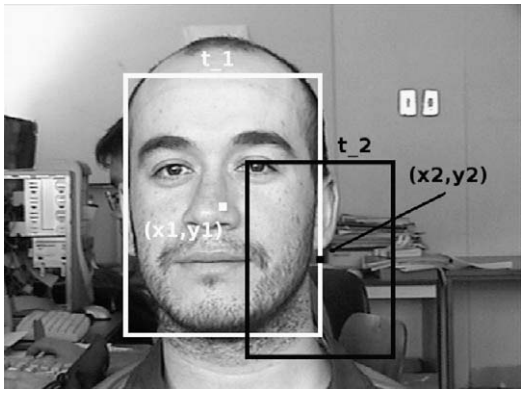

Fig. 6. Example of ground-truth box (white) and face localized box (black).

$$err = \frac{\max(d_l, d_r)}{d_{lr}}, \tag{16}$$

where d_l is the left eye disparity, i.e. the distance between the manually localized left iris position and the automatically localized position, d_r is the right eye disparity, and d_{lr} is the Euclidean distance between the manually localized left and right eyes. The detection is considered to be correct if $err < 0.25$.

2.4. Tests

2.4.1. Yale B face database

This database (Georghiades et al., 2001) contains images of 10 different subjects, each in nine different poses. For each pose, pictures were taken under 64 illumination conditions, provided by a single light source located in different angles. An image with ambient (background) light was also taken for each pose given a total of 5850 images (10 subjects $\times$ 9 poses $\times$ 65 illumination conditions).

For the test on Yale B database, a generic as well as a personalized PSO template was used. Personalized templates were created for each person, using four images per subject. These images were selected from the frontal pose dataset of each individual (pose 0). Fig. 7b shows the four images selected to create the PSO template for subject yaleB01 of Yale B database. Fig. 7a shows the images used to generate our generic PSO templates and correspond to pictures of four students from the Electrical Engineering Department.

Tests were performed on Yale B database to compare face localization performance of the PSO templates, the anthropometrics templates (Perez et al., 2007) and Adaboost (Jones and Viola, 2003) method in a database with different lighting conditions, changing pose and with some degree of face rotation. The Adaboost method implementation was trained using bootstraping and 5000 examples of faces and 5000 examples of non-faces, resulting in a cascade detector with 24 stages.

2.4.2. AR face database

The method based on the PSO templates was compared to the method that uses combined binary edges and intensity information (Song et al., 2006) for iris localization. In (Song et al., 2006) the method was applied on two subsets of the AR face database (Martinez and Benavente, 1998), AR-63 and AR-564. We applied the PSO template method to the same subsets of the AR face database and therefore, results on iris localization may be directly compared. Both the generic and personalized PSO templates for the AR subsets were employed. The personalized templates were obtained by using faces contained in the AR database for the PSO algorithm that were not part of the AR-63 or AR-564 subsets. We also tested our method against the method proposed in (Campadelli et al., 2009) that is based on Support Vector Machines trained on optimally chosen Haar wavelet coefficients. The algorithm presents very interesting results on several databases using the same metrics applied in our paper to iris localization. Full results for the AR face database are presented and also separated into six categories for different expressions (neutral, smile and anger) and light conditions (side light left, side light right and all lights).

2.4.3. FERET face database

Tests on iris localization for different relative errors were performed on a subset of the FERET database and compared to an SVM-based method (SVM2) as published in (Campadelli et al., 2007). The SVM2 method was selected because it yielded very

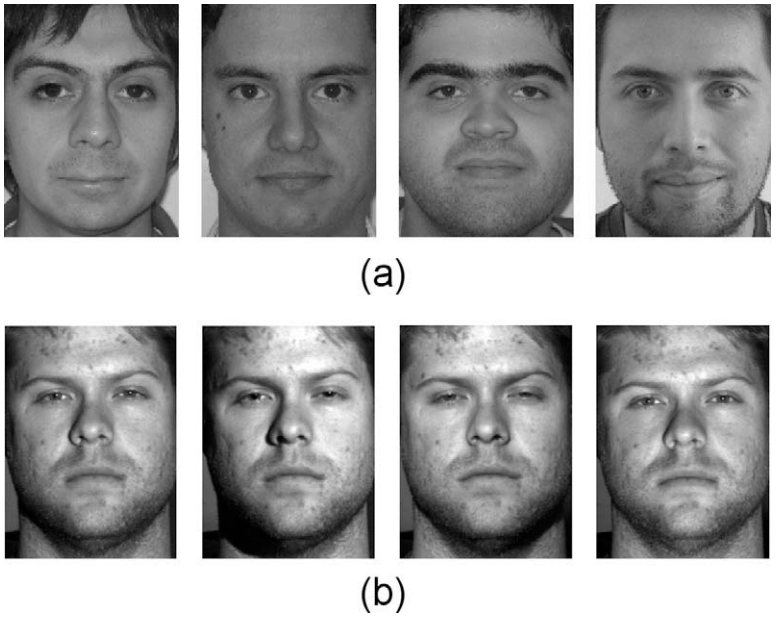

Fig. 7. (a) Four images used to create the generic PSO templates. Subjects correspond to four students at the Electrical Engineering Department, Universidad de Chile. (b) Four images of subject yaleB01 of Yale B database used to create a personalized PSO template.

good results in iris localization and because it was already compared to several other methods in (Campadelli et al., 2007). Additionally, we compared to the method proposed in (Campadelli et al., 2009) that improved the results obtained in (Campadelli et al., 2007) and also added mouth localization. Both methods use the same metrics for iris localization error that we use and therefore, a direct comparison is possible. The subset consisted of 300 images randomly selected from subsets fa and fb of the FERET database. A personalized set of templates was obtained by using eight faces from FERET that were not part of the 300 test images.

2.4.4. Video sequences

The PSO templates were compared with the anthropometric templates (Fig. 8) by computing the line integral value of the template over the directional images on a database composed of four video sequences as shown in Table 1. Each video sequence is composed of frontal faces from one individual but different video sequences correspond to different individuals. Personalized PSO templates were created for this test, one set for each person. Additionally, the spatial selectivity of PSO templates and anthropometric templates was assessed by computing the line integral within a window of 11×11 pixels around the face center. Also, the computational time employed by the line integral computation for face localization was compared for the anthropometric and PSO templates.

3. Results

3.1. PSO templates

Fig. 9 shows results of the evolution of a PSO template through 1000 iterations. Fig. 9a shows the first randomly initialized template. Fig. 9b–f shows the template for iteration number (b) 15, (c) 100, (d) 300, (e) 700 and (f) 1000. It can be observed that the number of pixels in the PSO template decreases as the number of iterations increases and therefore the number of computations involved in the line integral over the directional image will be reduced.

Fig. 10a, d and g shows three frontal faces, b, e and h shows the corresponding directional images and c, f and i shows the resulting PSO templates, respectively. By comparing the PSO template to the anthropometric template of Fig. 8, it can be observed that the first one fits better to the directional image.

The PSO templates were compared to the anthropometric templates in face localization in four video sequences. The total number of pixels in the template determines the number of computations to evaluate the line integral. A template with the smallest number of pixels will yield small computational time. A template with the largest line integral value when computed on the directional image over the face value will produce better face localization. Table 2 shows the resulting number of pixels for the personalized PSO templates, the generic PSO templates and for the anthropometric templates. As expected, personalized PSO templates had fewer points than both anthropometric and generic PSO templates.

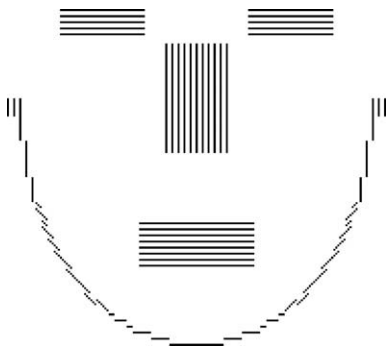

Fig. 8. Anthropometric template.

Table 1
Video-sequence database to test the PSO templates.

Video sequence	Number of frames		
	Template selection	Test	Total
1	466	230	696
2	202	175	377
3	306	523	829
4	395	214	609

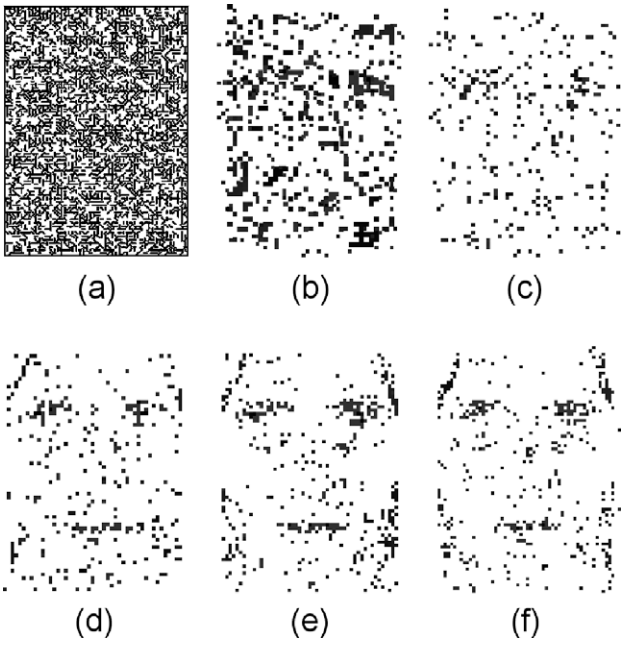

Fig. 9. Evolution of the PSO algorithm for 1000 iterations. Templates are shown for: (a) random initialization; (b) iteration number 15; (c) iteration number 100; (d) iteration number 300; (e) iteration number 700, and (f) final template at iteration 1000.

An important result of using PSO templates is that the computational time required to localize a face was reduced to 60% of that employed by an anthropometric template as shown in Table 3. The computational time was measured in video sequence 2 employing an Athlon XP 2000+, 1.67 GHz computer with 512 MB RAM. Another important advantage of the PSO templates is that they can be automatically generated from a small set of frames. In contrast, the anthropometric templates require to select manually the regions for the main facial features: eyes, eyebrows, nose, mouth and chin.

3.2. Face size estimation

The error in the face size estimation was computed as the error of the difference between the template size computed value and the ground truth for four video sequences (a–d). Error is presented relative to face size. Fig. 11a and b shows the error in the face size estimation for the anthropometric templates (black level) and the PSO templates (gray level) applied to video sequences 3 and 4 as a function of the frame number. It can be observed that the face size estimation error is always lower in the PSO template relative to the anthropometric template.

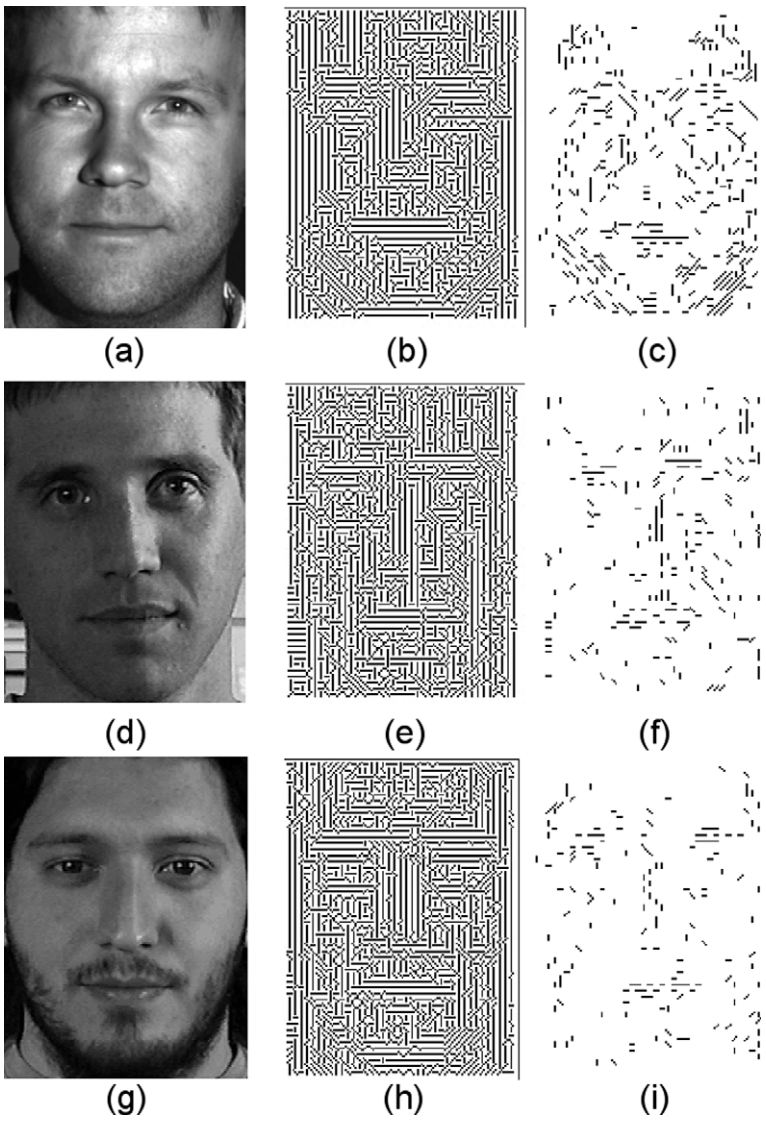

Fig. 10. Frontal faces (a, d, g), corresponding directional images (b, e, h), and corresponding final PSO templates (c, f, i).

Table 2
Total number of pixels in the resulting PSO templates and in the anthropometric templates.

Template	Number of pixels
Anthropometric	24,429
Personalized PSO templates	15,528
Generic PSO templates	18,768

Table 3
Comparison of the average computational time required to detect a face.

Template	Average time (ms)	Standard deviation
Anthropometric	28.5	0.0068
PSO	18.6	0.0053

3.3. Line integral

Fig. 11c and d shows the line integral value computed for each frame of video sequences 3 and 4 for the PSO template (gray level) and anthropometric template (black). It can be observed that for all frames the line integral value is greater in the PSO templates than in the anthropometric templates.

3.4. Spatial selectivity

The spatial selectivity of the PSO and anthropometric templates was assessed by computing the line integral value within a window of 11×11 pixels around the face center. Fig. 12a shows the original image, Fig. 12b shows the line integral value for the window of 11×11 pixels around the face center represented in gray levels for the PSO templates and Fig. 12c shows the line integral value within the 11×11 pixels window for the anthropometric templates. Darker gray levels mean lower values of line integral, with black being the lowest and white being the highest. Fig. 13a shows the line integral value along a line in the x axis crossing the face center horizontally. Fig. 13b shows the line integral value along a vertical line in the y axis across the face center. In Figs. 12 and 13, it can be observed that the PSO template provides a narrower spatial tuning around the face center providing a larger line integral value and a higher slope when the line integral is computed away from the face center. Therefore, the PSO template shows

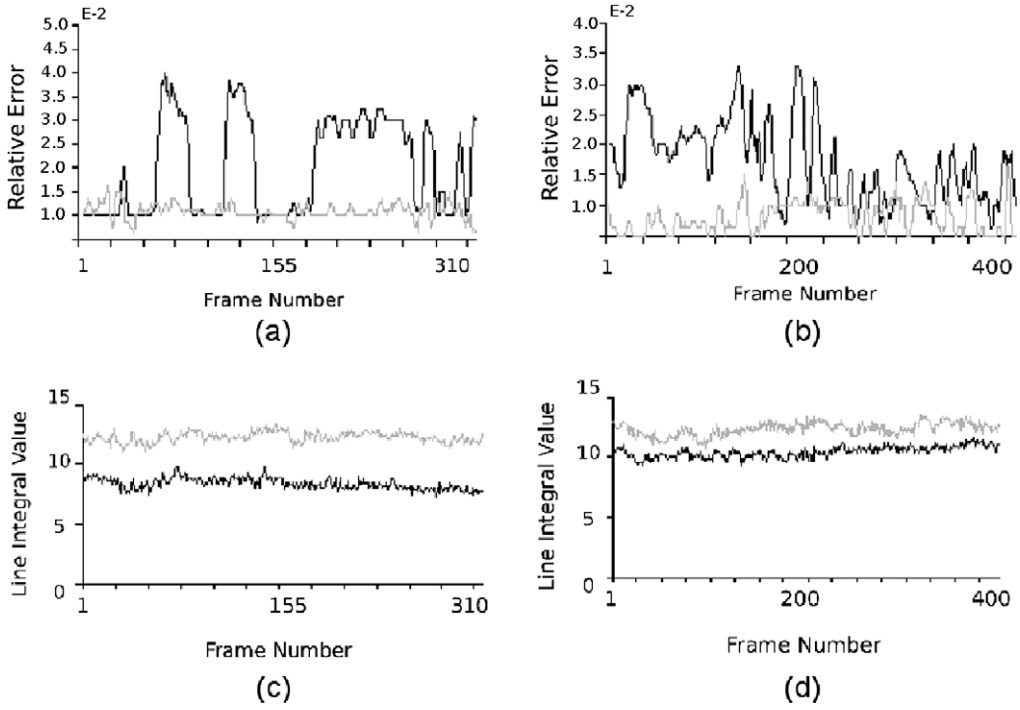

Fig. 11. (a) and (b): Frame by frame relative error in the face size estimation between the anthropometric templates (black) and the PSO template (gray level) applied to video sequences 3 and 4. Error is relative to face size; (c) and (d): line integral value for the PSO template (gray level) and the anthropometric template (black) applied to video sequences 3 and 4 as a function of the frame number.

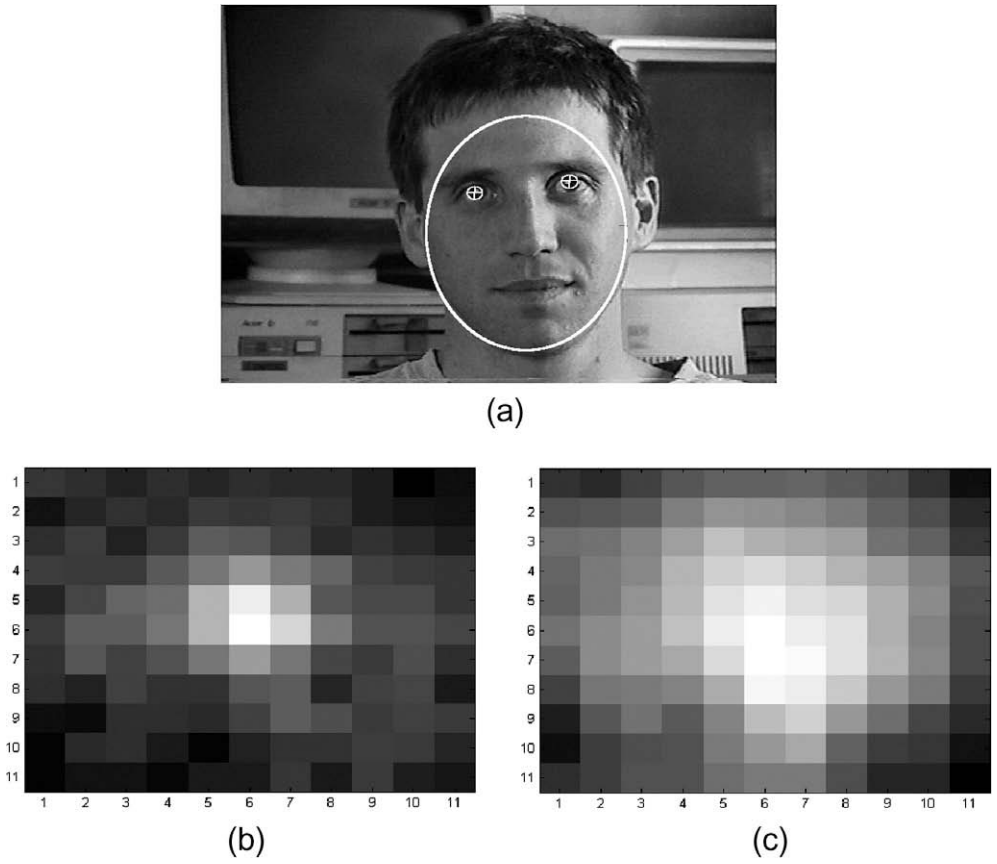

Fig. 12. (a) Original image; (b) line integral value computed with the PSO template in an 11×11 window around the face center; (c) line integral value computed with the anthropometric template in an 11×11 window around the face center.

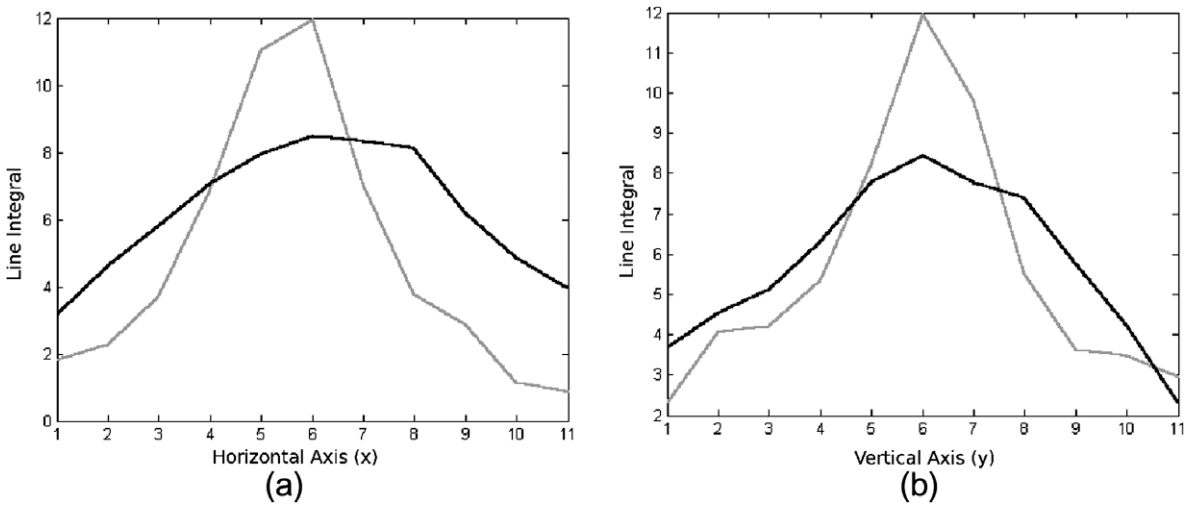

Fig. 13. (a) Line integral value computed along the *x* axis across the face center with the PSO template (gray) and anthropometric template (black). The computation was performed along 11 pixels through the face center; (b) line integral value computed along the *y* axis across the face center with the PSO template (gray) and anthropometric template (black). The computation was performed in 11 pixels through the face center.

Table 4
Yale B face localization rates.

Method	Det. rate (%)
Personalized PSO templates	82.65
Generic PSO template	81.92
Anthropometric templates	81.74
Adaboost	60.54

higher spatial tuning to localize faces when compared to anthropometric templates.

3.5. Face localization rates

3.5.1. Yale B database

Table 4 shows face localization rates for Yale B face database for PSO templates, anthropometric templates and Adaboost face localization method. It can be observed that the highest face

Fig. 14. Examples of correct face localization.

Table 5
Iris localization rates for two subsets of the AR face database.

	Relative error (*err*)							
	1	0.8	0.50	0.40	0.33	0.25	0.15	0.05
Song et al. (2006) AR-63	96.8	96.8	96.8	96.8	96.8	96.8	N/A	N/A
Song et al. (2006) AR-564	96.6	95.9	92.4	90.2	89.7	86.5	N/A	N/A
Anthropometric templates AR-63	100	100	100	100	100	100	100	74.1
Anthropometric templates AR-564	100	99.7	99.7	99.3	99.3	98.7	98.1	88.1
Generic PSO templates AR-63	100	100	100	100	100	100	100	75.6
Generic PSO templates AR-564	100	100	100	99.7	99.5	99.3	98.8	87.3
Personalized AR PSO templates AR-63	100	100	100	100	100	100	100	77.8
Personalized AR PSO templates AR-564	100	100	100	99.8	99.7	99.7	99.1	88.6

localization rate was obtained for the PSO templates. Also, it is shown that PSO as well as anthropometric templates performed better than the Adaboost face localization method. Since Yale B database has strong changes in illumination conditions, the template-based methods are more robust than Adaboost. Examples of correct face localization in the Yale B face database are shown in Fig. 14.

3.6. Iris localization

3.6.1. AR face database

Table 5 shows the results of iris localization for the anthropometric templates, the generic PSO templates and the personalized AR PSO templates in the AR-63 and AR-564 subsets of the AR face database. Results of the method proposed by Song et al. (2006) in iris localization is also shown for comparison. It can be observed that anthropometric templates reached better iris localization than the method of Song et al. by approximately 10% for the most restrictive criterion. Iris localization with the PSO templates, both generic and personalized, resulted in even a lower error reaching 100% precision for normalized error $err = 0.25$. Even for normalized errors of 0.15 and 0.05 PSO templates reached the highest iris localization rates.

In Table 6, it can be observed that PSO templates reached better results for iris localization than those of the best method published until today on the AR face database. Table 7 shows the performance of our algorithm across different partitions of the AR face database for different face expression and illumination conditions. Performance is lower for the smile expression. The performance of PSO templates is similar to the other methods for different illumination conditions.

Table 6
Comparison of iris localization as a function of the relative error (*err*) for Campadelli et al. (2009) and PSO templates on the AR face database.

	Relative error (*err*)				
	0.25	0.20	0.15	0.10	0.05
Campadelli et al. (2009)	99.3	98.0	96.7	93.7	76.8
PSO templates	99.7	99.4	99.0	95.9	87.9

Table 7
Iris localization performance on the subsets of the AR face database with different expressions and illumination.

	Neutral	Smile	Anger	Side lights	All lights
$err = 0.10$	98.5	89.3	95.7	97.4	97.9
$err = 0.25$	100	98.9	100	100	100
AUC	88.4	80.5	84.8	86.1	85.4

Table 8
Comparison of iris localization as a function of the relative error (err) for (Campadelli et al., 2009), SVM2 (Campadelli et al., 2007), anthropometric templates and PSO templates on FERET database.

	Relative error (*err*)				
	0.25	0.20	0.15	0.10	0.05
Campadelli et al. (2009)	99.7	99.3	99.3	97.3	67.7
SVM2	95.9	95.0	91.4	89.5	67.7
Anthropometric templates	92.1	91.9	88.7	84.6	74.1
PSO templates	96.6	95.0	92.3	89.0	76.0

3.6.2. FERET face database

Table 8 shows the results of iris localization for the anthropometric templates and PSO templates for a subset of the FERET face database as used in (Campadelli et al., 2007). The table also shows the results reported in (Campadelli et al., 2007) for the method called SVM2 in a similar subset of the FERET database. This allows the comparison of our method to a previous method that obtained very good results in iris localization in the FERET database. The method SVM2 was advantageously compared to several other methods in iris localization in (Campadelli et al., 2007) and therefore our proposed PSO template also compares advantageously to those other methods. In Table 8 it can be observed that PSO templates reached the highest iris localization for relative errors of 0.05, 0.15, 0.2 and 0.25. It is important to highlight the significant improvement for the most demanding error criteria used ($err = 0.05$). In Table 8 it can be observed that the method proposed in (Campadelli et al., 2009) performed better than all other methods for all errors of 0.05, 0.15, 0.2 and 0.25. One of the advantages of our proposed method is fast computation that takes 150 ms to achieve iris localization (including coarse face detection, face localization and iris localization) on a Pentium 4, 3.0 GHz, 1 GB RAM computer. The module that performs iris localization takes only 5 ms which permits a real-time implementation in a face recognition access control which we have in operation.

4. Conclusion

Precision of face and iris localization methods has proven to be essential in many applications including face recognition. Previous anthropometric template-based methods for face localization use face features commonly recognized by the human visual system such as eyes, eyebrows, nose, mouth and chin as the key features in the template. In this paper, a new method has been developed to generate PSO templates based on the PSO algorithm. The method was applied to the frontal face detection problem to four video sequences, the Yale B face database, the AR face database and FERET database. Results show the advantages of the PSO templates over the anthropometric templates. PSO templates maximize the response of the line integral performed over the directional image to localize the face. Results show that the line integral values are

larger and the face size estimation is better with the PSO templates compared to the anthropometric templates. The method also allows control of the maximum number of points in the template thus reducing the computational time, allowing real-time processing. Test on Yale B face database shows the robustness of template-based methods in extreme lighting conditions and in faces with small rotations. The PSO templates were compared to the Adaboost method for face localization showing improved results in the Yale B face database. We also tested the PSO templates in the iris localization problem and compared the results to those proposed recently in (Song et al., 2006; Campadelli et al., 2007; Campadelli et al., 2009). An improved performance in iris localization in comparison with Song et al. (2006) and Campadelli et al. (2007) but lower performance than Campadelli et al. (2009) on the FERET database was shown. On the AR database our iris localization method performed better than the other published methods. An important aspect of our methodology is that it can be applied in real time. Another important result is that PSO templates can be personalized for a particular user or situation, improving its selectivity and computational time even more. This is a common case in human–machine interfaces for handicapped individuals. PSO templates also showed improved spatial selectivity relative to the anthropometric templates. The method could be extended by using templates for partially occluded faces, weighting more the part of the template/face that is available. In the case of Yale B dataset where part of the face is not available due to non-homogeneous illumination, methods for illumination compensation could be used as a preprocessing algorithm. There are several methods for illumination compensation available based on discrete Cosine Transform (DCT) (Perez and Castillo, 2008), local normalization (LN) (Perez and Castillo, 2008) and self-quotient image (SQI) (Perez and Castillo, 2009). In a recent paper (Perez et al., 2009) we extended the anthropometric template-based face and iris detection to rotated faces. In (Perez et al., 2009) templates were created for rotated faces in angles $(-40^\circ, 40^\circ)$ in coronal axis and $(-45^\circ, 45^\circ)$ in transversal axis. The PSO templates could be extended in the future to rotations as demonstrated for anthropometric templates. In this case PSO templates would have to replace anthropometric templates and we expect to find similar results as in the present paper for frontal faces, i.e., increased selectivity and reduced computational cost.

Acknowledgments

This research was funded by FONDECYT through Project No. 1080593 and by the Department of Electrical Engineering, Universidad de Chile.

References

Campadelli, P., Lanzarotti, R., Lipori, G., 2007. Eye localization: A survey. The fundamentals of verbal and nonverbal communication and their biometrical issues. NATO Sci. Series 18, 234–245.

Campadelli, P., Lanzarotti, R., Lipori, G., 2009. Precise eye and mouth localization. Internat. J. Pattern Recognition Artificial Intelligence 23 (3), 359–377.

Chatterjee, A., Siarry, P., 2006. Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization. Comput. Oper. Res. 33, 859–871.

Chiang, C., Tai, W., Yang, M., Huang, Y., Huang, C., 2003. A novel method for detecting lips, eyes and faces in real time. Real-Time Imaging 9 (4), 277–287.

Clerc, M., Kennedy, J., 2002. The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Trans. Evol. Comput. 6 (1), 58–73.

Eberhart, R., Kennedy, J., 1995. Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Networks, Perth, Australia, vol. 4, pp. 1942–1948.

Eberhart, R., Shi, Y., Mar. 1998. Parameter selection in particle swarm optimization. In: Evolutionary Programming VII: Proceedings of the Seventh Annual Conference on Evolutionary Programming, San Diego, California, pp. 591–600.

Eberhart, R., Shi, Y., July 1999. Empirical study of particle swarm optimization. In: Congress of Evolutionary Computation, Washington, DC, USA, pp. 1945–1950.

Georghiades, A.S., Belhumeur, P.N., Kriegman, D.J., 2001. From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE Trans. Pattern Anal. Machine Intell. 23 (6), 643–660.

International Biometric Group, October 2008. Biometrics market and industry report 2009–2014. <http://www.biometricgroup.com/reports/public/market_rep ort.html>.

Ji, Q., 2002. 3d face pose estimation and tracking from a monocular camera. Image Vision Comput. 20 (7), 499–511.

Jin, Z., Lou, Z., Yang, J., Sun, Q., 2007. Face detection using template matching and skin-color information. Neurocomputing 70 (4-6), 794–800.

Jones, M., Viola, P., July 2003. Fast multiview face detection. Mitsubishi Elect, Res. Lab., Tech Rep. (TR2003-96).

Kato, Y., Hirano, T., Nakamura, O., 2007. Fast template matching algorithm for contour images based on its chain coded description applied for human face identification. Pattern Recognition 40 (6), 1646–1659.

Le, D.D., Satoh, S., 2006. A multi-stage approach to fast face detection. IEICE – Trans. Information Systems E89-D (7), 2275–2285.

Maio, D., Maltoni, D., 2000. Real time face location on gray scale static images. Pattern Recognition 33 (9), 1525–1539.

Martinez, A., Benavente, R., 1998. The AR face database. Tech. Rep. 24, CVC Tech. Rep.

Mohanty, P., Sarkar, S., Kasturi, R., 2006. A non-iterative approach to reconstruct face templates from match scores. 18th Internat. Conf. Pattern Recognition, vol. 4. IEEE Computer Society, Los Alamitos, CA, USA, pp. 598–601.

Park, K.S., Lim, C.J., 2001. A simple vision-based head tracking method for eye-controlled human/computer interface. Internat. J. Human–Comput. Studies 54 (3), 319–332.

Perez, C.A., Castillo, L.E., 2008. Genetic improvements in illumination compensation by the discrete cosine transform and local normalization for face recognition. In: SPIE Internat. Symp. Optomechatronic Technologies, ISOT 2008, 17–19 November, San Diego, CA, USA, vol. 7266, pp. 72661B-1–72661B-8.

Perez, C.A., Castillo, L.E., 2009. Illumination compensation for face recognition by genetic optimization of the self-quotient image method. In: ISOT 2009 Internat. Symp. Optomechatronic Technologies, 21–23 September, Istanbul, Turkey, pp. 322–327.

Perez, C.A., Vallejos, J.I., 2006. Face detection using PSO template selection. IEEE Internat. Conf. on Systems Man Cybernet., 4220–4224.

Perez, C., Salinas, C., Estevez, P., Valenzuela, P., 2003. Genetic design of biologically inspired receptive fields for neural pattern recognition. IEEE Trans. Systems Man Cybernet. B 33 (2), 258–270.

Perez, C., Gonzalez, G., Medina, L., Galdames, F., 2005. Linear versus nonlinear neural modeling for 2-d pattern recognition. IEEE Trans. Systems Man Cybernet. A 35 (6), 955–964.

Perez, C., Lazcano, V., Estevez, P., 2007. Real-time iris detection on coronal axis rotated faces. IEEE Trans. Systems Man Cybernet. C 37 (5), 971–978.

Perez, C., Lazcano, V., Estevez, P., Held, C., 2009. Template based face and iris detection on rotated faces. Internat. J. Optomechatron. 3, 1–14.

Phimoltares, S., Lursinsap, C., Chamnongthai, K., 2007. Face detection and facial feature localization without considering the appearance of image context. Image Vision Comput. 25 (5), 741–753.

Shi, Y., Eberhart, R., 1998. A modified particle swarm optimizer. In: IEEE Internat. Conf. Evol. Comput., pp. 69–73.

Sirohey, S., Duric, Z., Rosenfeld, A., 2002. A method of detecting and tracking irises and eyelids in video. Pattern Recognition 35, 1389–1401.

Song, J., Chi, Z., Liu, J., 2006. A robust eye detection method using combined binary edge and intensity information. Pattern Recognition 39 (6), 1110–1125.

Trelea, I., 2003. The particle swarm optimization algorithm: Convergence analysis and parameter selection. Information Process. Lett. 85, 317–325.

Verschae, R., del Solar, J.R., 2003. A hybrid face detector based on an asymmetrical adaboost cascade detector and a wavelet-Bayesian-detector. Lect. Notes Comput. Sci. 2686, 742–749.

Verschae, R., del Solar, J.R., Correa, M., 2008. A unified learning framework for object detection and classification using nested cascades of boosted classifiers. Machine Vision Appl. 19 (2), 85–103.

Wang, J.G., Sung, E., 2002. Study on eye gaze estimation. IEEE Trans. Systems Man Cybernet. B 32 (3), 332–350.

Yang, M.H., Kriegman, D.J., Ahuja, N., 2002. Detecting faces in images: A survey. IEEE Trans. Pattern Anal. Machine Intell. 24 (1), 34–58.

Ying, L., Xiang-lin, Q., Yun-jiu, W., 2001. Eye detection by using fuzzy template matching and feature-parameter-based judgement. Pattern Recognition Lett. 22 (10), 1111–1124.

Zhang, L.P., Huan-Jun, Shang-Xu, 2005. Optimal choice of parameters for particle swarm optimization. J. Zhejiang Univ. Sci. 6A, 528–534.

Zou, J., Ji, Q., Nagy, G., 2005. A comparative study of local matching approach for face recognition. IEEE Trans. Image Process. 16 (10), 2617–2628.