

PAPER

Multiple Face Recognition Using Local Features and Swarm Intelligence

Chidambaram CHIDAMBARAM^{†a)}, *Student Member*, Hugo VIEIRA NETO^{††}, Leyza Elmeri Baldo DORINI^{††}, and Heitor Silvério LOPES^{††}, *Nonmembers*

SUMMARY Face recognition plays an important role in security applications, but in real-world conditions face images are typically subject to issues that compromise recognition performance, such as geometric transformations, occlusions and changes in illumination. Most face detection and recognition works to date deal with single face images using global features and supervised learning. Differently from that context, here we propose a multiple face recognition approach based on local features which does not rely on supervised learning. In order to deal with multiple face images under varying conditions, the extraction of invariant and discriminative local features is achieved by using the SURF (Speeded-Up Robust Features) approach, and the search for regions from which optimal features can be extracted is done by an improved ABC (Artificial Bee Colony) algorithm. Thresholds and parameters for SURF and improved ABC algorithms are determined experimentally. The approach was extensively assessed on 99 different still images – more than 400 trials were conducted using 20 target face images and still images under different acquisition conditions. Results show that our approach is promising for real-world face recognition applications concerning different acquisition conditions and transformations.

key words: local features, iterative search, face recognition

1. Introduction

Among biometric recognition systems (e.g. iris and fingerprints), face biometric data plays an important role in security applications since they can be acquired at distance without any knowledge or collaboration of individuals. Although some research on automatic recognition of faces date from the 1970's, it is still an active area that continues to receive significant attention from both public and private research communities [1]. In addition, due to the development of new processing technologies and the ever-increasing capacity of data storage devices, a huge amount of digital images and video sequences have been acquired, thus requiring the development of algorithms able to extract interest regions or features automatically from these. In several cases, such images or videos are acquired under uncontrolled environmental conditions, frequently including noise, blur, occlusions and changes in scale and illumination, for example – this makes face recognition even more complex. Such problem has received attention from different areas, including image processing, pattern recognition, computer vision, artificial intelligence, computer graphics,

neuroscience, robotics and evolutionary computing [2], [3].

The main focus of this work is to deal with still images containing multiple faces (referred to just as still images throughout this work), which typically present a complex background and different acquisition conditions. Although there are several processing stages to recognize faces in this kind of images, we consider two main steps: the extraction of local invariant features to be considered in the matching process and the search for regions in the images from which the optimal local features can be extracted.

In the face recognition context, it is essential to determine relevant features to be considered in the comparison between the target face image (also referred to as query image) and the faces present in the still images. At the same time, the extraction of invariant features to represent faces are desirable to overcome geometric transformations, occlusions, changes in illumination and other issues. Many feature extraction approaches were presented in the literature, which can be coarsely classified into global and local feature extraction methods. Global features are extracted from the entire face region, meanwhile local features are obtained from specific discriminative regions of the face [4]. Most global face recognition methods rely on subspace methods (for example, Principal Component Analysis) and are expressively used in recognition applications. However, methods based on local features are becoming more usual since they tend to be more robust against variations such as facial expressions, noise and occlusions [5].

Since 2004, object recognition research has focused on extracting local image features from interest points [6]–[8] – object recognition processes based on interest point detectors is currently considered very effective in practice. In order to identify distinctive local image features, among many interest point detectors and descriptors available in the literature, SIFT (Scale Invariant Feature Transform) [6] and SURF (Speeded-Up Robust Features) [9] are the most popular. Asbach and coworkers [10] assessed several interest point detectors such as DoG (Difference of Gaussian), LoG (Laplacian of Gaussian), Hessian and Harris for face detection. Based on their assessment, they concluded that the scale invariant Harris detector used with SURF descriptors is the most promising combination for face location. Although SIFT is the most appealing descriptor for practical uses and most widely used, SURF provides a faster detector with robust descriptors – the speed aspect does not affect recognition performance [9]. In recent work for iris recog-

Manuscript received April 5, 2013.

Manuscript revised March 5, 2014.

[†]The author is with the Santa Catarina State University (UDESC), Joinville, Brazil.

^{††}The authors are with the Federal University of Technology - Paraná (UTFPR), Curitiba, Brazil.

a) E-mail: chidambaram@udesc.br

DOI: 10.1587/transinf.E97.D.1614

dition, features are extracted using SURF to handle image issues such as occlusion, non-uniform illumination and head tilt during acquisition [11].

Changes in illumination is one of the prominent bottlenecks in face detection and recognition tasks. Many approaches have been proposed to handle these issues applying intensity normalization procedures [12], local binary patterns (LBP) [13], illumination normalization techniques which include the self-quotient image (SQI) [14], [15], pre-processing algorithms [16] and plane subtraction with histogram equalization [17]. Some other works focused specifically on recognition of single faces using image-filtering techniques [18], [19] and on holistic approaches [20]. However early approaches concentrated their effort in handling variabilities due to illumination by detecting edges, because edges normally tend to be insensitive to a range of illumination conditions [21]. Here we hypothesize that the use of local descriptors may provide additional robustness against changes in illumination when face detection and recognition are concerned.

On the classification and recognition side, traditional search algorithms are computationally expensive. However, many real-life face recognition applications require fast and efficient search and matching algorithms [22]. Metaheuristic optimization algorithms, such as those from the Swarm Intelligence area, were successfully applied to face recognition problems. Several optimization algorithms have been successfully applied to face recognition purposes, such as Particle Swarm Optimization (PSO) [23], [24] and Artificial Bee Colony (ABC) [25] algorithms. Face detection and recognition constitute problems in which optimization algorithms have great potential to improve detection or recognition accuracy. Among several research works concerning face processing tasks, many use PSO. For example, Perez and Vallejos [26] proposed a face localization method to design an improved face template using PSO and predefined face templates. Similarly, other works propose to detect faces using PSO, using the Adaboost framework [27] and linear Support Vector Machines [28], [29].

Our main contribution in this work is to construct a system that recognizes a target face image in still images acquired under different conditions and containing multiple faces. For that, we propose a robust approach using SURF [9] for interest point detection and description, and an improved ABC algorithm (iABC) [25]. SURF was chosen because of its promising performance and computational efficiency [9], [30]. Similarly, iABC was chosen because it was already applied to object recognition and has proven to be an excellent optimization algorithm [25]. In addition to the search process for an optimal solution, the iABC can also determine the optimal image parameters (horizontal and vertical coordinates, rotation angle and scale). Although there are many other ways for searching and recognizing faces using SURF in still images, a swarm intelligence approach combined with interest point detectors can result in an effective and efficient recognition process [31].

The proposed approach is entirely based on the dis-

criminative power of local features obtained from interest points and does not require supervised learning – to the best of our knowledge, most face detection and recognition approaches available in the literature are based on supervised learning. Also, the use of swarm intelligence makes it possible to recognize faces without any explicit face detection step.

The remaining of the paper is organized as follows: In Sect. 2, a detailed view of the proposed approach and its development is explained. Experiments and results are discussed in Sect. 3. In Sect. 4 conclusions are drawn and future work is outlined.

2. Proposed Approach

In previous work [31], we conducted experiments using a single still image, acquired in different conditions, and multiple target face images. In order to optimize computational effort, at the beginning of the iterative process, all of the interest points of the still image were stored in a separate matrix structure. During the matching process, instead of computing interest points for each image patch cut from the still image, its interest points were obtained directly from the static matrix structure. In the present work, in addition to the matrix scheme (designated as matrix SURF-iABC approach), we have also assessed the matching process using the traditional way (designated just as SURF-iABC approach), in which the interest points are computed for every possible face image patch cut from the still image. Experiments were conducted using 20 target face images and more than 100 still images – four important experimental analyses are detailed in the present work:

1. Extensive experiments for tuning thresholds and parameters for SURF and iABC algorithms
2. Performance comparison between matrix and traditional SURF-iABC approaches
3. Experiments grouped into 11 different image conditions in 34 still images
4. Experiments grouped into 20 target face images in 65 still images

The general view of the proposed face recognition approach is shown in Fig. 1. The input target face image is searched in the still image through an iterative optimization process and, simultaneously at each iteration, a patch with similar size of the target face image is cut from the still image and its features are extracted. This process is repeated for a given number of iterations to find out the most similar face. More generally, the entire recognition process is based on two main steps: (1) definition of an image region for feature extraction in the still image by the iABC algorithm, and feature extraction that includes interest point detection and descriptors extraction from the image patch using SURF; (2) determination of repeated interest points and therefore the similarity between the target face image and the image patch cut from the still image. The main aspects of these two steps are addressed in the following sections.

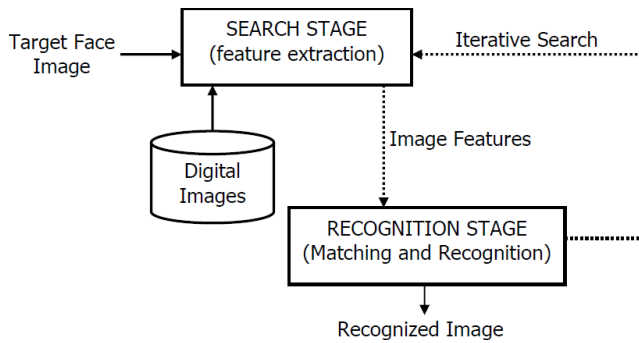


Fig. 1 General view of proposed face recognition approach.

Using a four-dimensional vector for each *individual* of the *population* of the iABC algorithm, an image patch is cut from the still image and its interest points are identified, from which the descriptor vectors are retrieved for matching between the image patch and the target face image. During the matching stage, the corresponding interest points between images are identified using distance measures of coordinates and descriptors. From the valid correspondences, *fitness* values are defined. Interest point locations and descriptors may vary and therefore, distance measures have to be computed based on coordinate distance and descriptor distance thresholds, respectively. Similarly, during the iterative process, stagnation of *individuals* affecting the search for optimal solutions (best match for the target face image in the still images) may happen. The escape from stagnation conditions can be done using an explosion mechanism.

2.1 Interest Point Detection and Description – SURF

Interest points can be characterized in several ways. They may be defined as a set of image pixels that have high level of variation in reference to a predetermined local measure [8] and can also be considered as salient regions that are highly distinctive [6].

In many recent computer vision applications, distinctive and representative regions of images are identified using interest points, which have been mainly applied for object recognition and related tasks [6], [32], [33]. Object recognition may be successful only if it is possible to find some distinctive image features among many alternative objects in cluttered real scenes [34]. Besides, it is essential to identify local image features which are invariant to image scaling, translation, rotation and illumination. Hence, interest points can be an alternative to extract distinctive and invariant face image features under different conditions [10], [35].

Among many interest point detectors, two can be mentioned as the most known popular: the Scale Invariant Feature Transform (SIFT) [6] and Speeded-Up Robust Features (SURF) [9]. Most detectors generate descriptor vectors which contain information regarding the neighborhood of every interesting point in an image. Both SIFT and SURF-based methods are used to detect interest points but the implementation of these detectors follow different schemes.

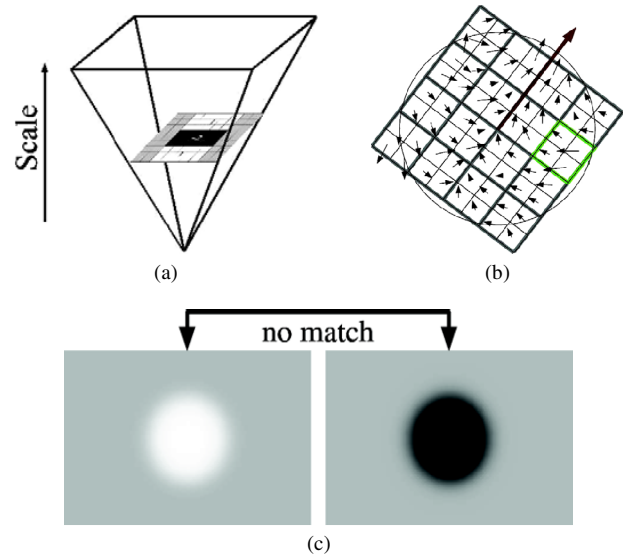


Fig. 2 (a) Analysis of the scale-space by up-scaling the filter size using a pyramid scheme; (b) Descriptor construction using an oriented quadratic grid with 4×4 square sub-regions and Haar wavelet responses ($\sum d_x$, $\sum d_y$, $\sum |d_x|$, $\sum |d_y|$) around an interest point; (c) Interest point matching considering contrast types.

SURF is a scale and rotation-invariant detector which detects interest points by selecting distinctive locations such as corners, blobs and T-junctions. Instead of reducing the image size as a pyramid [34], scale space is analyzed by up-scaling the filter size as shown in Fig. 2 (a). In this scheme, due to the use of integral images [36], the computational cost of each scale-space level becomes constant and independent of the filter size. According to Bay and colleagues [9], although SURF can be similar to SIFT in concept, the former is less sensitive to noise and outperforms the latter. Another major advantage of SURF is that it requires low computation time to detect and describe the interest points in comparison to SIFT. SURF builds a descriptor vector of 64 dimensions for each interest point, which is obtained by concatenating all 4×4 sub-regions of four-dimensional vectors of the underlying intensity structure, as shown in Fig. 2 (b). Furthermore, indexing during the matching stage is based on the sign of the Laplacian, which is demonstrated in Fig. 2 (c). This avoids erroneous matching of descriptors with different types of contrast. Hence, the way the distribution of intensity content within the interest point neighborhood is obtained reduces the time for feature computation and matching, and also increases descriptor robustness. Although SIFT is more popular, the performance of SURF is equal or better than SIFT and its computational efficiency is significantly better than SIFT [30].

The main task in this work is to find out similar features between two face images so that a high rate of correct matches can be achieved. To accomplish this goal, stability of interest points is necessary – the repeatability rate, which is the percentage of detected points that are repeated in two images, is the only measure of stability which is strongly accepted as a standard computer vision performance metric

for interest points [8]. As an assessment criterion, repeatability directly measures the quality of the features for tasks like image matching, object recognition and 3D matching. Hence, such criterion can be used for any kind of scenes or images [37]. Measurements of repeatability will quantify the number of repeated points detected under varying conditions, such as image rotation, scale changes, variations in illumination, presence of noise and changes of view point. The repeatability criterion is valid only for planar scenes in which the geometric relation between two images is completely defined [37].

An interest point is considered repeated if an interest point I_i from image Img_1 is similar to the interest point I_j of image Img_2 . This means that both interest points should lie in the same coordinates of both images and must have similar descriptors. However, exact localization becomes unlikely if one of the images went through geometric transformations or some other changes. Considering this issue, the rate of repeated interest points (RI) can be computed once they lie within a common region of Img_1 and Img_2 . To satisfy this condition, a distance error (or threshold) for coordinates has to be taken into account when the distance between the coordinates of interest points of images ($DistCoord$) Img_1 and Img_2 is computed. Therefore, to evaluate whether two interest points are repeated, they must first satisfy the following condition:

$$(I_i, I_j) \in RI, \text{ if } DistCoord((x, y)_{I_i}, (x, y)_{I_j}) < CDE_r, \quad (1)$$

where CDE_r represents the Euclidean coordinate distance threshold for interest points.

Similarly, once two interest points are within the common region validated by Eq. (1), a distance between descriptors of the same interest points also has to be considered when the comparison between two descriptors is evaluated. The repeated interest points (RI) can be defined as the number of interest points I_i and I_j that lie in a common region of images Img_1 and Img_2 and are within a certain distance between descriptors (DDE_r):

$$(I_i, I_j) \in RI, \text{ if } DistDesc((x, y)_{I_i}, (x, y)_{I_j}) < DDE_r, \quad (2)$$

where $DistDesc$ represents the Euclidean descriptor distance threshold for interest points validated by Eq. (2).

In the following sections, the repeatability rate r of interest points between images Img_1 and Img_2 is defined by the following equation:

$$r = \frac{|RI|}{\min(NI_i, NI_j)}, \quad (3)$$

where $NI_i = |I_i|$ and $NI_j = |I_j|$ represent the total number of interest points of Img_1 and Img_2 , respectively.

2.2 Iterative Search – ABC

In recent years, many metaheuristic algorithms based on

specific intelligent behaviors of swarms have been proposed and applied to several real-world problems, mainly to solve numerical and combinatorial optimization problems [38]. Based on the model proposed by Tereshko and Loengarov [39] for the foraging behavior of honey bee colonies, Karaboga [40] proposed the Artificial Bee Colony (ABC) algorithm. In this subsection, this population-based optimization algorithm is briefly detailed.

In the ABC algorithm [41], each *food source* is considered as a possible solution for an optimization problem. The *nectar amount* represents the quality (*fitness*) of the solution represented by a *food source*. At the beginning, the number of *employed bees* and *onlooker bees* must be defined, usually being the same. The quantity of *employed bees* represents the number of solutions (SN) in the *population*. The algorithm starts by associating all *employed bees* to randomly generated *food source* locations which are considered an initial population of SN . Each solution is represented by X_i , such that $i \in (1, 2 \dots SN)$. Each X_i is a d -dimensional vector, and d represents the number parameters to be optimized. Once the *employed bees* are created, the search process starts and is repeated by a predefined number of cycles, represented by $M CN$ (Maximum Cycle Number). During the search process, an *employed bee* shares the information about *food sources* with *onlooker bees* through the waggle dance. *Scout bees* will search for a new *food source* location randomly and the new random solution will replace the abandoned one. The selection of a *food source* location that should be abandoned is determined by a *Limit* variable associated to each solution. The *Limit* is defined by $d \times SN$. The whole process is repeated for a predetermined $M CN$ or until a termination criterion is reached.

2.3 ABC Algorithm for Face Recognition

A target face image is represented by a 4-tuple (x, y, s, θ) , as defined in Sect. 2.2. These four transformation parameters (horizontal and vertical coordinates, scale factor and rotation angle) should be optimized to find out the most similar patch in a still image. By considering our image context, the search space is limited by restricting the range of the parameters as follows: $x = [0, n]$ (columns), $y = [0, m]$ (rows), $s = [0.5, 1.5]$ (scale), $\theta = [-\pi/2, \pi/2]$ (rotation).

A *bee* or solution is a set of (x, y, s, θ) representing a position and orientation in the digital image. Then, each solution for the target face image can be represented by a four-dimensional vector $X_{id} = (X_{i1}, X_{i2}, X_{i3}, X_{i4})$, where $d = 4$ and $i \in 1, 2, 3 \dots SN$. A new position (*food source*) in the neighborhood of a specific solution is determined by altering the value of one randomly chosen solution parameter of X_{id} and keeping the remaining parameters unchanged. This neighborhood position can be computed using Eq. (4):

$$X_{ij} = X_{ij} + \phi(X_{ij} - X_{kj}), \quad (4)$$

where $k, i \in [1, SN]$, ϕ is a random number in the range $[-1, 1]$ and k is a random index that should be different from

i. The probability p_i of selecting a *food source* i by an *onlooker bee* can be computed by Eq. (5):

$$p_i = \frac{fit_i}{\sum_{n=1}^{SN} fit_n}, \quad (5)$$

where fit_i is the *fitness* of a solution. After the abandoned solutions are determined using *Limit*, the new random *scout bees* can be produced using Eq. (6):

$$X_{ij} = Xmin_j + \text{rand}(0, 1)(Xmax_j - Xmin_j), \quad (6)$$

where $Xmin$ and $Xmax$ represent the lower and upper bound values allowed for the four parameters (x, y, s, θ) and $\text{rand}(0, 1)$ is a random value in the range $[0, 1]$.

2.4 Improved ABC Algorithm (iABC)

According to the the basic ABC algorithm, only one parameter is perturbed at a given time when a new neighborhood solution is generated for both *employed* and *onlooker bees*. In the *scout bee* production phase, only one *scout bee* is generated when the limit counter value of a specific solution exceeds the *Limit* parameter. Otherwise, no *scout bee* will be produced. During the optimization process, the population of solutions can converge to a sub-optimal region in search space and result in stagnation of the best solution for a certain number of cycles continuously. When stagnation occurs, in order to restart the search for the best solution, an explosion procedure is applied for the object recognition problem [24]. The explosion procedure generally aids the algorithm to search in different regions of search space and to find the best solution gradually during the iterations.

Our main objective in the face recognition task is to recognize a target face image as fast as possible so that this kind of algorithm can be applied to real-world problems. In this context, the improved ABC algorithm comprises three main mechanisms that were tested in the study conducted by Chidambaram and Lopes [25]: (1) the perturbation of multiple variables; (2) generation of multiple *scout bees*; (3) explosion of stagnated solutions. Based on the three proposed mechanisms and the mechanisms which were already present in the basic ABC algorithm, such as the perturbation of a single variable and generation of a single *scout bee* using the *Limit* parameter, several experiments were done. Combination of these mechanisms resulted in eight different strategies. Finally, the best strategy was determined by assessing the results of all experiments from an object recognition problem [25]. The best strategy consisted of perturbation of all four variables, without generation of *scout bees* and with the explosion or decimation of stagnated population. Hence, in this work we have used the improved ABC algorithm to recognize faces in digital images.

3. Experiments and Results

3.1 Image Preparation and Experimental Setup

Based on real-world conditions, eleven different categories



Fig. 3 Sample still images with multiple faces under different conditions.

of still images were prepared for the experiments reported in the following subsections. Still images with multiple faces were captured under three main illumination conditions: (1) using a specific lighting system with two lights, denominated as Illum-I (Experiment I); (2) using a specific lighting system with one light plus room lights (fluorescent lamps), denominated as Illum-II (Experiment II); (3) using room lights (fluorescent lamps) only, denominated as Illum-III (Experiment III). Using these three illumination conditions, other images with head tilted (Rotation) (Experiment IV) and face occlusion (Experiment V) were acquired as shown in Fig. 3. Images with changes in scale and noise (blur and color noise), were artificially generated by an image editor in two levels for each category. For the scale, the size of the images were reduced to 95% (Scale-I, Experiment-VI) and enlarged to 105% (Scale-II, Experiment-VII). Likewise, two different noise levels were applied to the images (Blur-I, Blur-II, Color Noise-I, Color Noise-II) – experiments VIII to XI respectively refer to the images just mentioned.

All target face images used in this work were obtained under Illum-I (shown in Fig. 4). It is important to emphasize that all target face images (single faces) were obtained separately and are different from those in still images with multiple faces. The size of the images with multiple faces is 2592×1944 pixels and the target face images varies from 180 to 270 pixels in width and 240 to 340 pixels in height.

In this work, all image processing functions were implemented using OpenCV, and the improved ABC algorithm was written in C programming language. All experiments were executed on a cluster of computers with Pentium quad-core processors running Linux.

3.2 Tuning of Parameters and Thresholds

Initially, the parameters ($SN = 80$, $MCN = 100$) used in this work were defined empirically after some preliminary experiments. However, the size of bee population, maximum number of cycles and the maximum number of runs were obtained from previous work [25], [31].

These experiments include the tuning of two param-



Fig. 4 Target face images 1 to 20, from top left to bottom right.

Table 1 Recognition rates obtained for different convergence and decimation factors of iABC.

Convergence Factor (%)	Decimation Factor (%)			
	20	30	40	50
10	74.74	75.79	71.23	74.69
20	73.68	71.75	73.68	72.28
30	71.23	72.11	72.14	69.47
40	74.56	72.09	70.88	72.98

eters for the iABC algorithm and two thresholds for the SURF algorithm. The iABC parameters are: stagnation or convergence factor, and decimation factor. The convergence factor defines when the decimation (also referred to as explosion) of a part of population should be done during the optimization process. In fact, population convergence happens when the search process is conducted into a sub-optimal region without exploration of other regions of the search space, leading to the stagnation of fitness values without any improvement for many subsequent cycles. To overcome this situation, a decimation strategy is applied after the stagnation of the population – part of the population is substituted by new individuals, but, always keeping the best individual of the previous population. This kind of strategy improves the diversity of solutions and the possibility of finding even better solutions [42]. Using our set of still images, we have conducted several experiments to determine the two parameters of the iABC algorithm – according to the recognition rates shown in Table 1, 10% and 30% can be selected as the best values for the convergence and decimation factors, respectively.

Experiments were also conducted to determine the threshold values for coordinate distance and descriptor distance in SURF. Both thresholds refer to the Euclidean distance of coordinates and descriptors of two interest points independently. The experiments were done using our set of images by varying the coordinate distance from 20 to 60 in

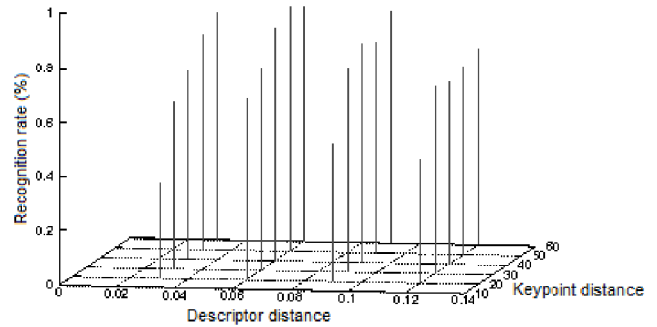


Fig. 5 Recognition rates corresponding to different threshold values for coordinate (keypoint) and descriptor distances in SURF.

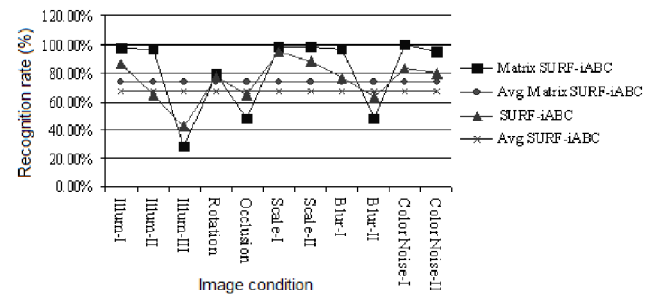


Fig. 6 Comparison of recognition rates obtained with SURF-iABC and matrix SURF-iABC approaches.

Table 2 Average recognition rate and execution time of SURF-iABC and Matrix SURF-iABC approaches.

Average Value	SURF-iABC	Matrix SURF-iABC	Gain (%)
Rec. Rate (%)	70.28	77.81	10.71
Exec. Time (s)	5.31	4.94	6.97

steps of 10 and the descriptor distance from 0.03 to 0.12 in steps of 0.03. As shown in Fig. 5, a coordinate distance of 50 and a descriptor distance of 0.06 provide the maximum recognition rate.

3.3 Preliminary Experiment Comparing SURF-iABC and Matrix SURF-iABC

The main objective of this preliminary experiment is to define which approach, SURF-iABC or matrix SURF-iABC, is the best for our face recognition approach. Therefore, we have compared both approaches based on recognition rate and execution time. This experiment was done using the parameters and thresholds defined in Sect. 3.2 and a set of images including all conditions. Results are shown in Fig. 6.

In Fig. 6, the two middle lines (Avg MatrixSURF-iABC and Avg SURF-iABC) correspond to the average recognition rate of all images in each approach. According to Table 2, the gain in the average recognition rate of matrix iABC-SURF compared to the traditional iABC-SURF approach is around 10.71% and the gain in execution time is around 6.97%. The execution time presented in Table 2

refers to the number of seconds spent per cycle during the iterative process of the iABC algorithm. From this analysis, we can conclude that the performance of the matrix SURF-iABC approach is better than the traditional SURF-iABC approach. Therefore, the matrix SURF-iABC is used in the remaining experiments of the paper.

Even though the matrix SURF-iABC yields better overall performance than the SURF-iABC approach, performances for image conditions such as Illum-III, Occlusion and Blur-II (shown in Fig. 6) call our attention. In these conditions, the traditional SURF-iABC has overcome the matrix SURF-iABC approach. For example, the average recognition rates of matrix SURF-iABC and SURF-iABC for Illum-III are 28.75% and 42.50%, respectively. This highlights the fact that the matrix SURF-iABC approach performance may not be the best for all image conditions. Therefore, it is important to emphasize that both Illum-III and Occlusion can occur in real-world conditions from which face images are commonly acquired.

3.4 Experiments using The Matrix SURF-iABC Approach in Different Conditions

The main objective of the present experiment is to study the robustness of the proposed approach and check whether it can effectively recognize faces in images obtained under different conditions. The first three experiments (I, II and III) with Illum-I, Illum-II and Illum-III were conducted using three still images and 20 target face images. All other experiments (IV to XI) were conducted with 10 target face images. A total of 140 trials were conducted using 36 different still images as stated in Sect. 3.1. The trials were grouped according to the eleven existing image conditions. Some sample images with multiple faces and target face images are presented in Figs. 4 and 3. Results are shown in Table 3.

Changes in illumination is one of the prominent bottlenecks of face processing tasks – additionally to pose variation, face images are changed in such a way that recognition performance is affected significantly and some approaches have proposed to handle this issue [18], [43]. According to the Face Recognition Vendor Test (FRVT) 2006, changes in illumination appear at the top of the list of issues that affect recognition performance [44] and still remains a problem for state-of-the-art algorithms [45]. Therefore, we have conducted several experiments using still images captured under different lighting conditions.

From the results shown in Table 3, it can be observed a strong influence of lighting conditions in the recognition rate. The recognition rate decreases drastically from Illum-I to Illum-II to Illum-III. The images for Illum-I and Illum-II were obtained using a special lighting system whereas for Illum-III they were obtained just using normal room lighting conditions (fluorescent lamps). This later condition really produces face images with non-uniform illumination and shadows according to the position of lights and their incidence on different parts of the scene. Hence, variations of light intensity associated with the spatial location of faces

Table 3 Recognition rates and execution times of matrix SURF-iABC using images under different conditions.

Exp.	Experiment Type	Recognition Rate (%)	Exec. Time per Cycle (s)
I	Illum-I	81.17	4.14
II	Illum-II	67.33	3.91
III	Illum-III	31.23	3.42
IV	Rotation	51.92	3.69
V	Occlusion	85.33	2.10
VI	Scale-I	85.67	3.31
VII	Scale-II	89.00	3.75
VIII	Blur-I	68.67	3.69
IX	Blur-II	47.67	3.82
X	Color Noise-I	85.33	3.78
XI	Color Noise-II	85.00	3.89

can be one of the factors that influences negatively the face recognition performance.

Besides different illumination conditions, images under different orientations should be tested to assess the robustness of our face recognition approach, as inclined faces (head tilted) can also appear in still images. The size of faces (scale) may also vary in still images. Furthermore, in spite of the fact that most face recognition methods are based on full face images, in real-world scenarios it is possible that occluded faces occur. In order to measure the influence of image rotation, occlusion and scaling, and to study their impact on recognition performance, experiments IV to VII were conducted (the images used in experiments VI and VII were artificially manipulated to change the scale). As shown in Table 3, the recognition of all conditions is above 85%, except for rotations.

The last part of the experiments (VIII to XI) using Matrix SURF-iABC is related to the presence of noise and blur in the still images. It is appropriate to mention here that the effects of blur and noise should be studied in unconstrained visual scenarios like face recognition applications [46]. Blur generally arises due to lens out-of-focus, atmospheric turbulence, motion of the camera or the object, and inaccurate camera settings [46], [47]. In this context, the main influence of blur will be on the transition of edges [47]. Optimal lighting conditions can not avoid issues that arise from the camera sensor – studying the effects of blur and color noise in applications such as face recognition is essential to develop robust approaches. To study the impact of these issues, we have performed two sets of experiments using artificially blurred and noisy (color) images. In the experiments, low recognition rates were obtained for blurred images, more specifically regarding the Blur-II condition.

In our proposal, in addition to the *fitness* value which is used to determine optimal solutions (face images), we have also implemented an additional verification procedure to check whether the corresponding central coordinates (X_1 , X_2) generated by the iABC algorithm are within the actual region of the identified face (ground-truth) in a still image with multiple faces. If the coordinates are inside the ground-truth region, then the recognized face is considered as a valid solution.

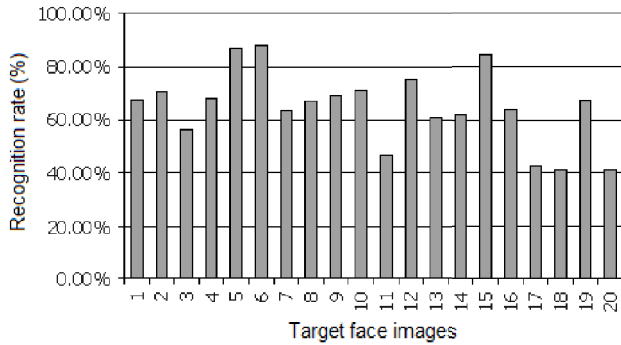


Fig. 7 Average recognition rate of target face images 1 to 20, searched in images under different conditions.

3.5 Experiment using The Matrix SURF-iABC Approach and Target Face Images

In Sect. 3.4, we assessed our algorithm on several images focusing on different image conditions (I to XI, as shown in Table 3). In this subsection, our goal is to assess the robustness of the proposed approach focusing on 20 target face images. Each face was searched in 12 different still images that were not used in the previous experiments. The still images used in this experiment cover image conditions such as Illum-I, Illum-II, Illum-III, Occlusion and Rotation. Thus, we have conducted 240 trials searching for the target face images shown in Fig. 4.

The average recognition rates obtained in this scenario are shown in Fig. 7, where the results are grouped according to each target face image. The frequency distribution of the recognition rates among all images can be separated as follows: 25% of target face images had between 41 and 60% recognition rate, 60% had between 61% and 80%, and 15% had between 81% and 100%. Hence, this results demonstrate the promising ability of our proposed approach to recognize faces in different conditions. Similarly to what was done in the previous experiment, in this experiment the central coordinates generated by the iABC algorithm for the recognized faces was compared to their ground-truth location in the still images.

3.6 Discussion of Results

The discussion here focuses on the recognition rate obtained for condition Illum-III, since it corresponds to the worse performance of our approach in comparison to other image conditions in all experiments that were conducted so far. The average recognition rates obtained from the illumination conditions Illum-I, Illum-II and Illum-III for the experiments reported are summarized in Table 4, which shows an abrupt decrease in recognition rate under Illum-III.

Figure 8 illustrates the recognition rates of illumination conditions Illum-I and Illum-III for all target face images obtained from the experiment reported in Sect. 3.5. One can clearly note from the graphs in Fig. 8, the non-

Table 4 Comparison of average recognition rates in illumination conditions Illum-I, Illum-II and Illum-III.

Experiment Type	Illum-I	Illum-II	Illum-III
Image Condition (I-XI)	80.17%	67.33%	31.23%
Target Face Image (1-20)	82.79%	70.83%	40.42%

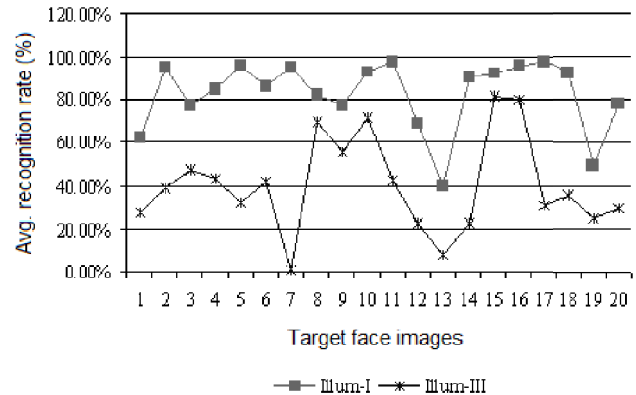


Fig. 8 Average recognition rates of Illum-I and Illum-III based on target face images (experiment from Sect. 3.5).

uniform recognition rates of target images, in some cases even close to 0%. High recognition rates in the graphs refer to face images with uniform illumination whereas low recognition rates refer to face images with partial illumination. Hence, the main drawback of the present approach is related to non-uniform illumination conditions. A possible way to overcome this issue is to use illumination compensation approaches or different feature extraction methods.

4. Conclusions

The main contribution of this work is a face recognition approach that is able to recognize the most similar face in still images with multiple faces under real-world conditions. We have developed a novel SURF-iABC approach, which does not rely on supervised learning and is based on local features. The approach was extensively tested with a large variety of image conditions. Based on the promising results shown, we can conclude that the proposed approach is robust and efficient enough to recognize faces in images with multiple faces, except when non-uniform illumination is concerned.

Because the propose approach does not rely on supervised learning, it can be applied effectively to face recognition applications without the need of any previous knowledge or training. Furthermore, the use of local features contributes to enhance recognition rates for occluded faces. Also, the use of Swarm Intelligence represented by the iABC algorithm provides more power to recognize images in different conditions, such as geometric transformations and the presence of noise.

Since there are many factors that can not be easily measured from one image to another under varying environmental conditions, the effects of changes in illumination in image recognition tasks still remain as an open subject for

research. Though there are many directions for further research, future work will focus on the problem of changes in illumination, aiming to improve the overall performance of the proposed approach.

References

- [1] X. Tan, S. Chen, Z.H. Zhou, and F. Zhang, "Face recognition from a single image per person: A survey," *Pattern Recognit.*, vol.39, pp.1725–1745, 2006.
- [2] M.H. Yang, D.J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.1, pp.34–58, 2002.
- [3] H. Zhou and G. Schaefer, "Semantic features for face recognition," *Proc. 52nd International Symposium ELMAR 2010*, pp.33–36, Sept. 2010.
- [4] W. Zhou, A. Ahrary, and S.I. Kamata, "Image description with local patterns: An application to face recognition," *IEICE Trans. Inf. & Syst.*, vol.E95-D, no.5, pp.1494–1505, May 2012.
- [5] C. Singh, E. Walia, and N. Mittal, "Robust two-stage face recognition approach using global and local features," *The Visual Computer*, vol.28, no.11, pp.1085–1098, 2012.
- [6] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol.2, no.60, pp.91–110, 2004.
- [7] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *Int. J. Comput. Vis.*, vol.60, no.1, pp.63–86, 2004.
- [8] L. Trujillo and G. Olague, "Synthesis of interest point detectors through genetic programming," *Proc. Genetic and Evolutionary Computation Conference*, pp.887–894, 2006.
- [9] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol.110, no.3, pp.346–359, 2008.
- [10] M. Asbach, P. Hosten, and M. Unger, "An evaluation of local features for face detection and localization," *Proc. 9th International Workshop on Image Analysis for Multimedia Interactive Services*, pp.32–35, May 2008.
- [11] H. Mehrotra, P.K. Sa, and B. Majhi, "Fast segmentation and adaptive SURF descriptor for iris recognition," *Mathematical and Computer Modelling*, vol.58, no.1–2, pp.132–146, 2013.
- [12] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.15, no.10, pp.1042–1052, 1993.
- [13] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on features distributions," *Pattern Recognit.*, vol.29, no.1, p.5159, 1996.
- [14] H. Wang, S. Li, and Y. Wang, "Face recognition under varying lighting conditions using self quotient image," *Proc. 6th International Conference on Face and Gesture Recognition*, pp.819–824, 2004.
- [15] O. Arandjelovic, "Making the most of the self-quotient image in face recognition," *Proc. 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, pp.1–7, 2013.
- [16] R. Gross and V. Brajovic, "An image preprocessing algorithm for illumination invariant face recognition," *Proc. 4th International Conference on Audio- and Videobased Biometric Person Authentication*, p.1018, 2003.
- [17] J.R. del Solar and J. Quinteros, "Comparing pre-processing approaches for illumination invariant face recognition," *tech. rep., UCH-DIE-VISION-2006-04*, Dept. Elect. Eng., Universidad de Chile, 2006.
- [18] O. Arandjelovic and R. Cipolla, "A methodology for rapid illumination-invariant face recognition using image processing filters," *Comput. Vis. Image Understand.*, vol.113, pp.159–171, 2009.
- [19] A. Nabatchian, E. Abdel-Raheem, and M. Ahmadi, "Illumination invariant feature extraction and mutual-information-based local matching for face recognition under illumination variation and occlusion," *Int. J. Pattern Recognit.*, pp.2576–2587, 2011.
- [20] A. Matsumoto, Y. Shirai, N. Shimada, T. Sakiyama, and J. Miura, "Robust face recognition under various illumination conditions," *IEICE Trans. Inf. & Syst.*, vol.E89-D, no.7, pp.2157–2163, July 2006.
- [21] P.N. Belhumeur and D.J. Kriegman, "What is the set of images of an object under all possible illumination conditions?," *Int. J. Comput. Vis.*, vol.28, no.3, pp.1–16, 1998.
- [22] V.N. Pawar and S.N. Talbar, "An investigation of significant object recognition techniques," *Int. J. Computer Science and Network Security*, vol.9, no.5, pp.17–29, 2009.
- [23] M. Sugisaka and X. Fan, "An effective search method for neural network based face detection using particle swarm optimization," *IEICE Trans. Inf. & Syst.*, vol.E88-D, no.2, pp.214–222, Feb. 2005.
- [24] H.A. Perlin, H.S.L. Lopes, and T.M. Centeno, "Particle swarm optimization for object recognition in computer vision," *Lect. Notes Comput. Sci.*, vol.5027, pp.11–21, 2008.
- [25] C. Chidambaram and H.S. Lopes, "An improved artificial bee colony algorithm for the object recognition problem in complex digital images using template matching," *Int. J. Natural Computing Research*, vol.1, no.2, pp.54–70, 2010.
- [26] C. Perez and J. Vallejos, "Face detection using PSO template selection," *IEEE International Conference on Systems, Man and Cybernetics*, vol.5, pp.4220–4224, Oct. 2006.
- [27] A.W. Mohammed, M. Zhang, and M. Jonhston, "Particle swarm optimization based Adaboost for face detection," *Proc. IEEE Congress on Evolutionary Computation*, pp.2494–2501, 2009.
- [28] E. Marami and A. Tefas, "Using particle swarm optimization for scaling and rotation invariant face detection," *Proc. IEEE Congress on Evolutionary Computation*, pp.1–7, July 2010.
- [29] J. Wei, Z. Jian-qi, and Z. Xiang, "Face recognition method based on support vector machine and particle swarm optimization," *Int. J. Expert Systems with Applications*, vol.38, no.4, pp.4390–4393, 2011.
- [30] V. Pimenov, "Fast image matching with visual attention and SURF descriptors," *Proc. 19th International Conference on Computer Graphics and Vision*, pp.49–56, 2009.
- [31] C. Chidambaram, M.S. Marçal, L.B. Dorini, H.V. Neto, and H.S. Lopes, "An improved ABC algorithm approach using SURF for face identification," *Lect. Notes Comput. Sci.*, vol.7435, pp.143–150, 2012.
- [32] C.B. Perez and G. Olague, "Evolutionary learning of local descriptor operators for object recognition," *Proc. 11th Annual Conference on Genetic and Evolutionary Computation*, pp.1051–1058, 2009.
- [33] P. Azad, T. Asfour, and R. Dillmann, "Combining Harris interest points and the SIFT descriptor for fast scale-invariant object recognition," *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.4275–4280, Oct. 2009.
- [34] D.G. Lowe, "Object recognition from local scale-invariant features," *Proc. International Conference on Computer Vision*, pp.1150–1157, 1999.
- [35] C. Fernandez and M. Vicente, "Face recognition using multiple interest point detectors and SIFT descriptors," *Proc. 8th IEEE International Conference on Automatic Face Gesture Recognition*, pp.1–7, Sept. 2008.
- [36] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.511–518, 2001.
- [37] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *Int. J. Comput. Vis.*, vol.37, no.2, pp.151–172, June 2000.
- [38] R.S. Parpinelli and H.S. Lopes, "New inspirations in swarm intelligence: A survey," *Int. J. Bio-Inspired Computation*, pp.1–16, 2011.
- [39] V. Tereshko and A. Loengarov, "Collective decision-making in honey bee foraging dynamics," *Computing and Information Systems*, vol.9, no.3, pp.1–7, 2005.
- [40] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," *tech. rep., Erciyes University, Engineering Faculty*,

Comp. Eng. Dept., 2005.

- [41] D. Karaboga and B. Akay, "A comparative study of artificial bee colony algorithm," *Applied Mathematics and Computation*, vol.214, no.1, pp.108–132, 2009.
- [42] C.M.V. Benítez and H.S. Lopes, "Hierarchical parallel genetic algorithm applied to the three-dimensional hp side-chain protein folding problem," *Proc. 2010 IEEE International Conference on Systems, Man and Cybernetics*, pp.2669–2676, 2010.
- [43] J.R. del Solar and J. Quinteros, "Illumination compensation and normalization in eigenspace-based face recognition: A comparative study of different pre-processing approaches," *Pattern Recognition Letters*, vol.29, no.14, pp.1966–1979, 2008.
- [44] NSTC, "Face recognition - document of subcommittee on biometrics," tech. rep., National Science and Technology Council (NSTC), 2006.
- [45] J. Beveridge, D.S. Bolme, B.A. Draper, G.H. Givens, and Y.M. Lui, "Quantifying how lighting and focus affect face recognition performance," *Proc. Computer Vision and Pattern Recognition Workshop*, pp.74–81, 2010.
- [46] R. Gopalan, S. Taheri, P. Turaga, and R. Chellappa, "A blur-robust descriptor with applications to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.34, no.6, pp.1220–1226, June 2012.
- [47] J. Van de Weijer and C. Schmid, "Blur robust and color constant image description," *IEEE International Conference on Image Processing*, pp.993–996, Oct. 2006.



Leyza Elmeri Baldo Dorini obtained her MSc and DSc in Computer Science from the University of Campinas in 2006 and 2009, respectively. She is currently with the Department of Informatics and the Graduate School on Biomedical Engineering (PPGEB) at the Federal University of Technology - Paraná (Curitiba). Her major research interests are in the fields of pattern recognition, and image processing and analysis.



Heitor Silvério Lopes obtained his BSc and MSc in Electrical Engineering from the Federal University of Technology - Paraná (UTFPR), in 1984 and 1990, respectively, and his PhD from the Federal University of Santa Catarina in 1996. He is currently with the Department of Electronics and the Graduate School on Electrical Engineering and Applied Computer Science (CPGEI) at UTFPR Curitiba, where he is the head of the Bioinformatics Laboratory. His major research interests are in the fields of evolutionary computation, bioinformatics, high-performance computing and computer vision.

evolutionary computation, bioinformatics, high-performance computing and computer vision.



Chidambaram Chidambaram obtained his BEng in Mechanical Engineering from the Madurai Kamaraj University in 1985, his MEng in Production Engineering from the Federal University of Santa Maria in 1992, and his DSc in Computer Engineering from the Federal University of Technology - Paraná (UTFPR) in 2013. He is currently with the Department of Information Systems at the Santa Catarina State University (Joinville). His major research interests are in face recognition and swarm intelligence.

intelligence.



Hugo Vieira Neto obtained his BSc in Electrical Engineering and his MSc in Applied Computer Science from the Federal University of Technology - Paraná (UTFPR) in 1995 and 1998, respectively, and his PhD in Robotics from the University of Essex in 2006. He is currently with the Department of Electronics and the Graduate School on Electrical Engineering and Applied Computer Science (CPGEI) at UTFPR Curitiba. His major research interests are in the fields of computer vision, machine learning, and mobile robotics.

learning, and mobile robotics.