Precision Range Image Registration Using a Robust Surface Interpenetration Measure and Enhanced Genetic Algorithms.

Article *in* IEEE Transactions on Pattern Analysis and Machine Intelligence · June 2005

DOI: 10.1109/TPAMI.2005.108 · Source: PubMed

3 authors:

Luciano Silva
Universidade Federal do Paraná
134 PUBLICATIONS **1,853** CITATIONS

SEE PROFILE

Olga Regina Pereira Bellon
Universidade Federal do Paraná
124 PUBLICATIONS **1,778** CITATIONS

SEE PROFILE

Kim L. Boyer
University at Albany, The State University of New York
168 PUBLICATIONS **4,652** CITATIONS

SEE PROFILE

3D Face Recognition using Simulated Annealing and the Surface Interpenetration Measure

Chauã C. Queirolo, *Student Member, IEEE*, Luciano Silva *Member, IEEE*,
Olga R. P. Bellon, *Member, IEEE*, and Maurício Pamplona Segundo

Abstract—This paper presents a novel automatic framework to perform 3D face recognition. The proposed method uses a Simulated Annealing-based approach (SA) for range image registration with the Surface Interpenetration Measure (SIM), as similarity measure, in order to match two face images. The authentication score is obtained by combining the SIM values corresponding to the matching of four different face regions: circular and elliptical areas around the nose, forehead and the entire face region. Then, a modified SA approach is proposed taking advantage of invariant face regions to better handle facial expressions. Comprehensive experiments were performed on the FRGC v2 database, the largest available database of 3D face images composed of 4,007 images with different facial expressions. The experiments simulated both verification and identification systems and the results compared to those reported by state-of-the-art works. By using all the images in the database, a verification rate of 96.5% was achieved at a False Acceptance Rate (FAR) of 0.1%. In the identification scenario, a rank-one accuracy of 98.4% was achieved. To the best of our knowledge, this is the highest rank-one score ever achieved for the FRGC v2 database when compared to results published in the literature.

Index Terms—3D Face Recognition, Surface Interpenetration Measure (SIM), Range Image Registration.

1 Introduction

Face recognition is a very challenging subject. So far, studies in 2D face recognition have reached significant development, but still bear limitations mostly due to pose variation, illumination, and facial expression [38]. One way to overcome such limitations is the use of face depth information. In the 90's, 3D face recognition stood out due to advances of 3D imaging sensors. However, 3D images also have limitations, such as the presence of noise and difficult image acquisition [7].

A common approach for 3D face recognition is the use of registration techniques to perform range image matching. The Iterative Closest Point (ICP) [5] algorithm, or one of its variants, is usually sought to accomplish this task. The Mean Squared Error (MSE), minimized during the convergence process, is then used to compute the similarity between two face images [8], [9], [14], [24]. This approach can also be employed with deformation techniques to model facial expressions, minimizing its effects on face recognition [17], [19], [23].

The performance of both 2D and 3D face recognition systems can be verified on the Face Recognition Grand Challenge (FRGC) [29]. The FRGC is an international benchmarking composed of six challenging problems with its expected performance guidelines. The third experiment, *i.e.* ROC III, regards the 3D face recognition problem using a challenging database composed of 4,007 images from 466 subjects displaying different facial expressions. For this experiment, the goal is to achieve a verification rate of 98.0% at a False Acceptance Rate (FAR) of 0.1% [29].

We present a complete framework for face recognition using only 3D information (range images) as input. The framework can be employed in both verification and identification systems and has four main stages (see Fig. 1): (1) image acquisition, (2) preprocessing, (3) matching, and (4) authentication.

Initially, the 3D face image can be acquired by different techniques, such as laser and structured light scanners [15]. In the preprocessing stage, the input image is smoothed and the face region is automatically detected and segmented. Then, some facial feature points are detected to be used during the matching process [33]. Each segmented face region is matched with its corresponding one in the database. A Simulated Annealing-based approach (SA) for range image registration runs the process by using two robust measures to assess alignment precision: (1) M-estimator Sample Consensus (MSAC) [37], and (2) Surface Interpenetration Measure (SIM) [36]. In previous works, the SIM showed to be a discriminatory measure when working on face images [4], [30].

A modified SA approach is also employed to handle facial expression effects during face matching. At last, an authentication score is computed based on the SIM values obtained from the alignment of each region. This score is useful for: (1) verifying whether two face images belong or not to the same person, and (2) retrieving the identity of a subject.

One major contribution of our work is to propose the SIM as similarity score between two 3D face images. Differently from previous works in the literature which focus on range image matching using ICP (guided only by MSE), we rely on the fact that in [36] it was proven

• *C. C. Queirolo, L. Silva, O. R. P. Bellon and M. Pamplona Segundo are with the IMAGO Research Group, Universidade Federal do Paraná, Caixa Postal 19092, Curitiba, PR, Brazil, 81531-980.*
E-mail: chaua, luciano, olga, mauricio@inf.ufpr.br.

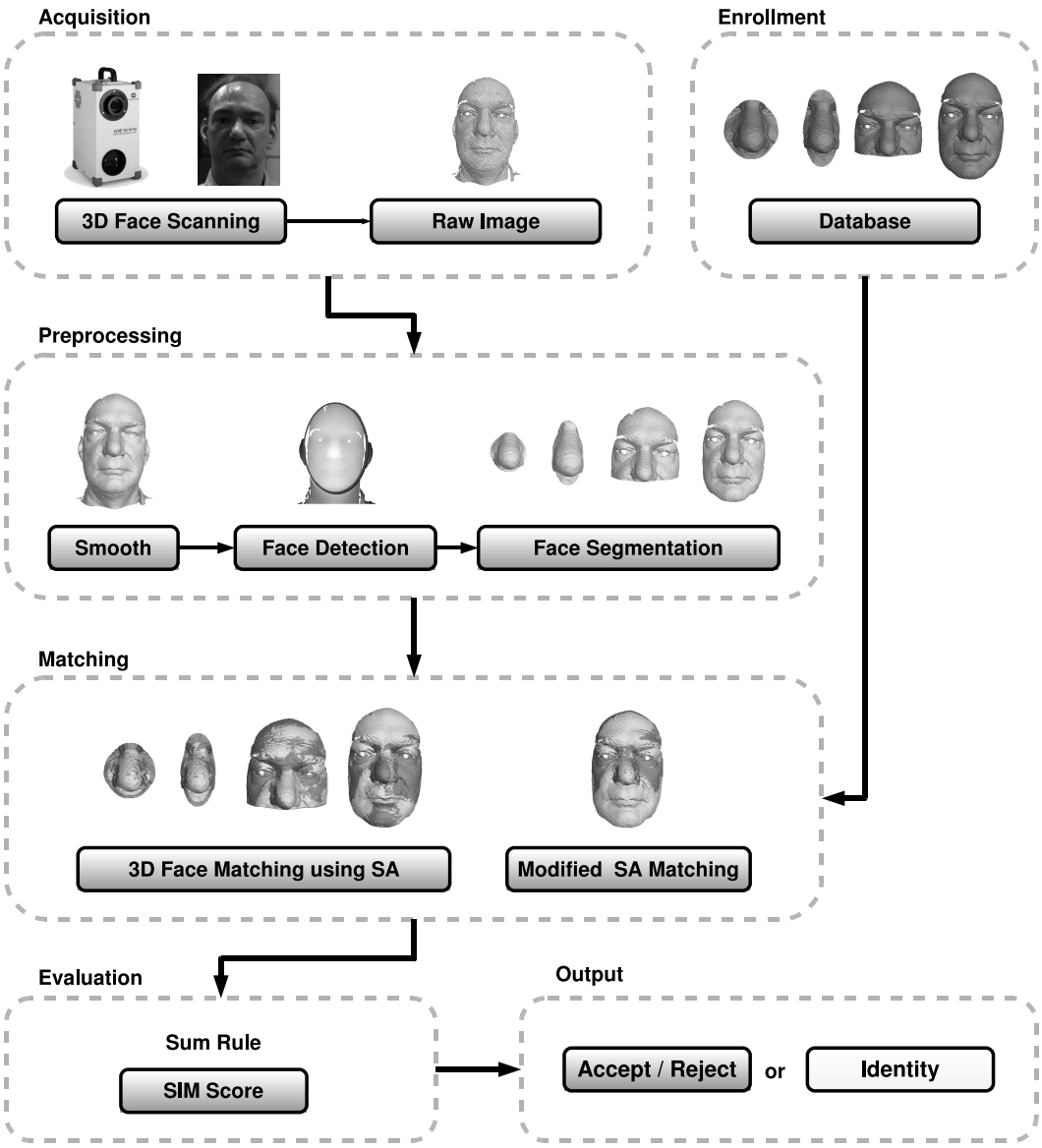


Fig. 1. The main stages of the framework for 3D face authentication.

that MSE could allow imprecise local convergence for range image registration, even when improved ICP-based approaches are used [16], [32]. Moreover, in [36] it is suggested that MSE is a good measure to start the image registration process, but the SIM is more suitable to be used at "the end of the game" to assess the quality of the registration.

We performed experiments on the FRGC v2 database, including experiments in controlled datasets and the benchmarking experiments. The verification experiments include an analysis of the results based on 0% FAR. We also present a complete analysis of the results with comparisons against the most relevant related works [10], [14], [19], [22], [27], [28]. This analysis is one of the most comprehensive reported in literature regarding 3D face recognition on the FRGC v2 database, from our best knowledge. Faltemier *et al.* [14] also present a complete set of experiments. Details of our experiments can be verified at the IMAGO Research Group homepage [1].

This paper is organized as follows. Section 2 discusses the main related works. Section 3 presents details about the images available in the FRGC v2 database. Section 4 describes the preprocessing stage. Section 5 regards the proposed method for face image matching and Section 6 regards the 3D face authentication process. Section 7 presents the experimental results and Section 8 concludes with the final remarks.

2 Related Works

This section reviews the most relevant works regarding 3D face recognition which use the FRGC v2 database. They compute the similarity score between two face images mostly based on the MSE minimized during the 3D image matching, sometimes combined with other measures.

Chang *et al.* [8] describe a method called Adaptive Rigid Multi-Region Selection (ARMS). ARMS uses three different overlapping regions around the nose: (a) circle, (b) ellipse, and (c) the nose itself. The nose region was designated because it is invariant in the presence of facial expression. Each region is matched with its corresponding one using ICP and the resulting Root Mean Square Error (RMSE) of each alignment is combined using the product rule. This value is then used as similarity measure between face images. Experiments were performed using the super set of images from the FRGC, composed of 4,485 3D face images. A rank-one accuracy of 97.1% was achieved when using a dataset composed of neutral images; and a rank-one accuracy of 87.1% when using a dataset containing faces with facial expression.

Maurer *et al.* [27] introduce a method that combines 3D and 2D face images. The 2D face recognition is performed using commercial software, although authors suggest that it can be replaced by any other approach. The 3D shape recognition is based on the ICP algorithm and uses distance map as similarity measure between two faces. The scores, computed from texture and shape are combined using a weighted sum rule. They used the FRGC v2 database and achieved a verification rate of 87% at 0.1% FAR, when using only shape information. The authors classified the main error sources affecting their method in three categories: (1) distorted shape images, (2) 2D and 3D images with different facial expressions, and (3) others, *e.g.* hair covering the face.

Husken *et al.* [18] propose a fusion between 2D and 3D Hierarchical Graph Matching (HGM) to perform face recognition. This approach uses an elastic graph that carries texture information and positions of facial landmarks. The graph is automatically adapted to a face using these landmarks, *e.g.*, nose, eyes and mouth. The distances between two graphs are used as the similarity score to distinguish two face images. Experiments were performed in the FRGC v2 database and a verification rate of 86.9% at 0.1% FAR was achieved when using only 3D images.

Lin *et al.* [22] describe a method that fuses the summation invariant features extracted from different regions for face classification. Ten rectangular areas are selected from the input range image and then a subset of the summation invariant features is extracted. The matching score from each region is combined using the sum rule or a weighted sum rule optimized by the Linear Discriminant Analysis (LDA). Using LDA to define weights allowed the achievement of higher recognition rates. Experiments were executed on the FRGC v2 database, and for the ROC III experiment they achieved 90.0% of verification rate, at 0.1% FAR.

Mian *et al.* [28] describe a multimodal 2D+3D face recognition system. Initially, the 3D face image pose is automatically corrected and normalized along with its texture image. For 2D face images the Scale-Invariant Feature Transform (SIFT) is applied to detect local features in faces. This method shows to be more robust regarding facial expression and illumination when compared to the Principal Component Analysis (PCA) baseline performance. The similarity between 2D face images is determined by the Euclidean distance of facial feature descriptors. For 3D face images, first the face is segmented into two regions: (1) the eyes and forehead region, and (2) the nose region. Then, these regions are matched with their corresponding ones using the ICP algorithm. The final similarity score is obtained by combining scores from each matching approach. Considering only results for the 3D face images, a verification rate of 86.6% was achieved at 0.1% FAR. By using a neutral gallery and all remaining images as probes, this verification rate is increased to 98.5% at 0.1% FAR.

Cook *et al.* [10] employ Log-Gabor filter to execute 3D face recognition. After applying the filter, the face image is segmented into 49 squared regions and further decomposed by three scale filters totalizing 147 subregions. The PCA is applied to each filter response from each subregion to represent the face with 147 features. The Mahalanobis-Cossine distance is employed to compute the similarity between two feature vectors. In the experimental results, the best performance was achieved by using 75 central subregions. On the FRGC v2 database they reported a rank-one accuracy of 96.2% and a verification rate of 92.3%, at 0.1% FAR.

Kakadiaris *et al.* [19] present a fully automated framework for 3D face recognition using the Annotated Face Model (AFM) to overcome the facial expression problems. An AFM is previously created to be used during alignment and deformed in the matching stage, allowing the creation of a metadata for authentication. In earlier steps of the framework, the input image is processed for spike removal and hole filling. A subsample of the data is created to improve the method efficiency. Before the fitting procedure starts, the raw data is aligned to the AFM model using a multistage matching method which combines three matching procedures: (1) Spin Images, (2) ICP and (3) SA on Z-buffers. A deformed model is obtained after the fitting process and it is converted into a geometric model and a normal map. Each image is then treated separately using two wavelet transformations, Pyramid and Haar. Authentication is computed in wavelet domain by assessing a different distance metric for each wavelet type. In the ROC III experiment of the FRGC v2, they reported a verification rate of 97.0%, at 0.1% FAR. Also, they reported a rank-one performance of 97.0%.

Al-Osaimi *et al.* [3] present an approach for 3D face recognition using deformable models. A PCA subspace is created based on the facial expression deformation. Initially, the training stage requires neutral and non-neutral images from each subject. Each pair is cropped and fine aligned using the ICP to verify the shape residues. A PCA subspace is built using the shape residues, which consequently can model a generic facial expression deformation. Then, the similarity is computed by evaluating the morphed 3D faces. The experiments

were conducted in the FRGC v2 database using different datasets. They report a verification rate of 98.35% and 97.7%, at 0.1%FAR, when using neutral and non-neutral datasets, respectively. In the ROC III experiment, they achieved 94.1% of verification rate, at 0.1% FAR.

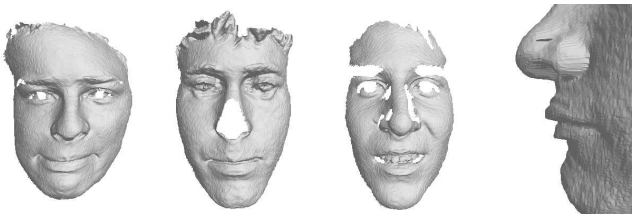
Faltemier *et al.* [14] develop a system that consists in the fusion of 28 spherical regions extracted from the face. These regions are then combined using the ICP algorithm. The distance obtained from the matching of each region was combined using different approaches, and the best results were reported using Consensus Voting and Borda Counting. To make the system more robust for small variations, which may occur in some face regions, different small regions are used for matching. Experiments on the FRGC v2 database achieved a verification rate of 93.2% at 0.1% FAR and a rank-one performance of 97.2%.

For a survey of works related to 3D and multimodal 2D+3D face recognition the reader should refer to [2], [7]. As observed in this section, most works successfully employ the ICP algorithm for 3D face recognition [8], [14], [27], [28]. Then, the MSE, sometimes combined with other measures, is used as the similarity score. In this paper we present a novel approach, applying the SIM to obtain more precise results.

3 3D Face Image Database

In our experiments we use the FRGC v2 database, the largest available database of 3D face images, composed of 4,007 images from 466 different subjects [29]. Each person has a varying number of images, up to 22 images, and among these subjects 56 have only one image. All images have resolution of 640x480 and they were acquired by a Minolta Vivid 910 laser scanner. The face images have frontal pose and several types of facial expression. The most common facial expressions available in the database are: neutral, happy, sad, disgusting, surprised and puffy cheek. Moreover, some images present artifacts which we have categorized as follows: (1) stretched or distorted images, (2) nose absence, (3) holes around nose, or (4) "waves" around mouth. Examples of these images are presented in Fig. 2. Figure captions represent the filename of each image in the database.

For instance, the disturbance presented in (1) was caused because the subject moved during the acquisition process. The second problem regards the distance between the person and the acquiring system. If a person is too close, closer regions such as nose may not be properly scanned because of focal range limitations. Besides, the laser sensor cannot deal well with transparent regions and regions with high specularity, which causes holes in the scanned image, *e.g.*, in the eyes and around the nose. Bowyer *et al.* [7] addressed this issue and identified several situations that affect range image quality, such as environment illumination. Some of these peculiarities from the FRGC v2 database have also been reported by Maurer *et al.* [27].



(a) 04746d44 (b) 04505d222 (c) 04531d293 (d) 04514d324

Fig. 2. Example of images with artifacts: (a) stretched/deformed images, (b) face images without nose, (c) holes around nose and (d) waves around the mouth.

TABLE 1
FRGC v2 Face Image Classification.

Image sets	Number of images
Small expression with artifacts	247
Small expression without artifacts	455
Neutral without artifacts	933
Neutral with/without artifacts	2,182
All	4,007

Aiming the verification of the robustness of our approach for 3D face recognition, we manually classified the FRGC v2 database according to facial expression and noise level (regarding the artifacts). We performed a careful facial expression classification because the one provided by FRGC has some misclassified images. This probably occurred because the classification was made according to the prompting request, whereas an outside observer may assign a different label to the image. Fig. 3 shows examples of faces classified as neutral expression by FRGC, but those actually have facial expression. In this paper, the database is segmented into different sets of images as presented in Table 1.

(a) 04476d230 (b) 04587d114

Fig. 3. Example of images classified as neutral in FRGC v2 database.

We classified the FRGC v2 database into three datasets according to holes in the eye region: (1) without holes, in 51 images, (2) small holes, in 1,300 images, and (3) large holes, in 2,656 images. This occurs because the eye is translucent and the light stripe is refracted instead of being reflected. Due to this, the laser scanner cannot properly capture a reflection and then no information is

produced in this region. We noticed that almost all the images from the database presented artifacts. Because of this, we do not consider the eye holes as a disturbance. Further experiments to verify its influence in the recognition process are presented in Section 7.2.1.

We also noticed two subjects that, besides having different identifications in the FRGC v2 database, are actually the same person, as confirmed by the FRGC organization. The "impostor" identification number in the database is 04643, and the other subject identification is 04783. This label was rectified for our experiments.

Fig. 4. Feature points detection: (a) six feature points; (b) wrong nose points, and (c) wrong eye corners detection.

4 3D Face Image Preprocessing

In this stage, the region of interest (*i.e.* face region) is extracted from the input image. The input images have some areas that may interfere in the 3D face matching, such as hair, neck and ears. Initially, the face image is smoothed with a median filter. The segmentation process uses our own approach based on the depth of range maps and it is basically composed of two main stages: (1) locating homogeneous regions in the input image by combining region clustering and edge information, and (2) identifying candidate regions that belong to the face region by an ellipse detection method based on the Hough Transform. This fully automatic approach can correctly segment 99% of faces from FRGC v2 database (3,991 of 4,007). The segmented faces that were not classified as correct are those with some remaining non-face regions (usually hair, neck or head accessories). Yet we use the segmented images without any post-processing in the experiments. More details about the image segmentation approach are described in [33].

After the face region is segmented, six feature points are detected to extract rigid regions of the face and improve the matching process; the inner right and left eye corners, the right and left nose corners, the nose tip and base, as shown in Fig. 4a. To perform this task we use our own approach that is able to detect automatically the eye corners and the nose features (*i.e.* tip, corners and base) in 99.92% and 99.6% of the images from FRGC v2, respectively. The nose features were not correctly detected due to nose absence only in three images (see Fig. 4b). The eye corners could not be correctly detected in 16 images due to the head pose rotation (see Fig 4c). In our experiments we used the points as they were detected. However, these images had minor negative impact in our results due to our hierarchical approach that uses different regions (see Section 6.1). Further details are presented in [33].

In this work, four regions of the face are considered (see Fig. 5): (1) the circular and (2) the elliptical areas around nose, (3) the upper head, including eyes, nose and forehead, and (4) the entire face region. Regions (1) and (2) are used because the nose area suffers less influence from facial expression as compared to other face regions [8]. Both regions have the same shape as the ones proposed by Chang *et al.* [8] and they can be extracted by applying geometry functions in feature coordinates. Figs. 5b and 5c show an example of these regions.

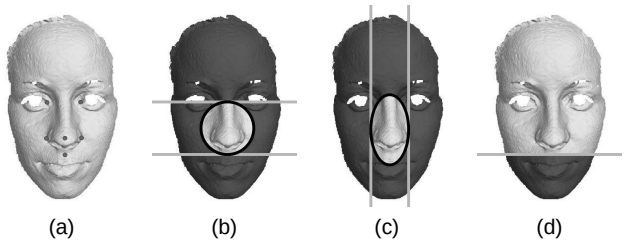


Fig. 5. Segmented regions from a same face: (a) entire face and detected feature points, (b) circular nose area, (c) elliptical nose area and (d) upper head.

In [4] it is stated that, when using the SIM, the nose region alone is not enough discriminatory for face recognition because it represents only a small fraction of the face; therefore, regions (3) and (4) were also used because they have more information about the face. Region (3) is obtained by selecting all face points above the nose base point, as shown in Fig. 5d. Therefore, the four segmented regions are used to perform face recognition.

5 3D Face Matching

In [34] the authors present an extensive comparison between the ICP guided by the MSE and a stochastic approach based on Genetic Algorithm (GA) guided by the SIM. They present results showing that the ICP tends to converge to local minimum solutions due to the MSE. On the other hand, when using the SIM combined with a stochastic approach, it is possible to converge to a global minimum. Although GA obtains more precise results for range image registration, it is extremely time-consuming. Later works [35], [36] showed that the SIM is more discriminating than the MSE when assessing the alignment of two range images.

In [4] the Genetic Algorithm is replaced by the Simulated Annealing, which is much faster than the GA, but still slower than the ICP. This work presents a comparison between the ICP and SA and show that SA produces better results. Moreover, it is noticed that

for 3D face recognition, the SIM works well even with alignments produced by the ICP.

In the present work we use the SA to perform the range image matching using two robust measures: (1) MSAC [37], and (2) SIM [35], [36]. The SA starts the registration process using the MSAC and the final tuning is achieved by using the SIM. The resulting SIM value is then used to assess the final 3D face matching. The following sections present a brief explanation regarding the SIM and SA.

5.1 Surface Interpenetration Measure (SIM)

The SIM was developed by analyzing visual results of two aligned surfaces, each one rendered in a different color, crossing over each other (*i.e.* like an interweaving) repeatedly in the overlapping area. The interpenetration effect results from the very nature of real range data, which presents slightly rough surfaces with small local distortions caused by limitations of the acquiring system [11], [36].

Through the quantification of interpenetration, one can more precisely assess the registration results and provide a highly robust control. Registrations of two range images presenting good interpenetration have high SIM values, and erroneous alignments produce low SIM values and those small differences in MSE can yield significant differences in SIM. Furthermore, alignments with high SIM present a very low interpoint distance between the two surfaces. That is, SIM is a far more sensitive indicator of alignment quality when comparing "reasonable" alignments.

For more details of this measure, the reader should refer to [35], [36]. The SIM pseudocode is presented in Algorithm 1. As suggested in [35], some constraints were applied to SIM to avoid incorrect corresponding points and obtain more precise alignments. A constraint $m = 5$ was defined as the maximum angle allowed between normal vectors at c and p, $\vec{n}_c$ and $\vec{n}_p$, respectively. Another constraint was defined to eliminate the corresponding points on surfaces boundaries. In this case, $p \in C$ if $c \notin D$, with D the set of boundary points in B, with boundary thickness defined as $b = 1$. The window centered in p was defined as $n = 5$, as suggested in [36].

Algorithm 1 Surface Interpenetration Measure – SIM.

Require: Two range images A and B
for $p \in A$ **do**
  Define a $n \times n$ neighborhood N_p centered in p;
  Search the corresponding point c of p in B, with $c \notin D$;
  Compute angle θ between normal vectors $\vec{n_p}$ and $\vec{n_c}$, regarding points p and c;
  if $\theta < m$ **then**
    for $q_i, q_j \in N_p$, with $q_i \neq q_j$ **do**
      if $[(q_i - c) \cdot \vec{n_c}] \cdot [(q_j - c) \cdot \vec{n_c}] < 0$ **then**
        $C_{(A,B)} \Leftarrow C_{(A,B)} \bigcup p$;
      end if
    end for
  end if
end for
return $\frac{C_{(A,B)}}{|A|}$

5.2 SA approach for Range Image Registration

The Simulated Annealing (SA) [20] is a stochastic algorithm for local search. From an initial candidate solution, SA iteratively searches a better neighbor solution to the problem as compared to the current one. The main difference between SA and other local search algorithms, *e.g.* Hill Climbing, is that SA can accept a worse solution than the current candidate during the iterative process; therefore, SA does not remain "tied" to local minima and because of this it has higher chances to reach its goal, which is a solution close enough to the global one.

The SA method requires the modeling of the problem in such a way that it is possible to move from a candidate solution to any neighbor one. For registration of two range images, six parameters (regarding rotation and translation on X, Y and Z axis, respectively) are defined as a "transformation vector". When applying this vector to one image, it is possible to align it with another one.

Our SA-based approach basically has three main stages: (1) initial solution, (2) coarse alignment, and (3) fine alignment, which are illustrated in Fig. 6. First, an initial solution is obtained by aligning two face images by their center of mass. This procedure is a rigid transformation which aligns both images to the same coordinate system and consequently provides a faster convergence process. A coarse alignment is performed using a SA-based searching procedure to minimize a robust assessment measure based on MSAC estimator [37], which is combined with resulting MSE of corresponding points between the two face images. MSAC defines a threshold that classifies corresponding points either as inliers or outliers. The error associated to outliers is a fixed penalty and to inliers is the error measure itself. By reducing the error associated to outliers, their influence is minimized on registration process and, therefore, a better alignment can be obtained. In the last step, a fine alignment is obtained by a SA-based searching procedure using SIM as assessment measure, where the goal is to maximize interpenetrating points between the two faces.

The "*temperature*" parameter t is reduced very slowly and two iterations are performed for each allowed "*temperature*" until the final one is achieved, otherwise system becomes frozen [25]. The initial t_0 must be "hot" enough to accept almost every neighbor solution from the candidate one. If it is set to a low value, then final solution might be close to the initial one; otherwise, the search procedure will run randomly at least in initial stages of SA. Initial "*temperature*" was empirically defined as $t_0 = 0.002$ and $t_0 = 0.07$ for stages (2) and (3),

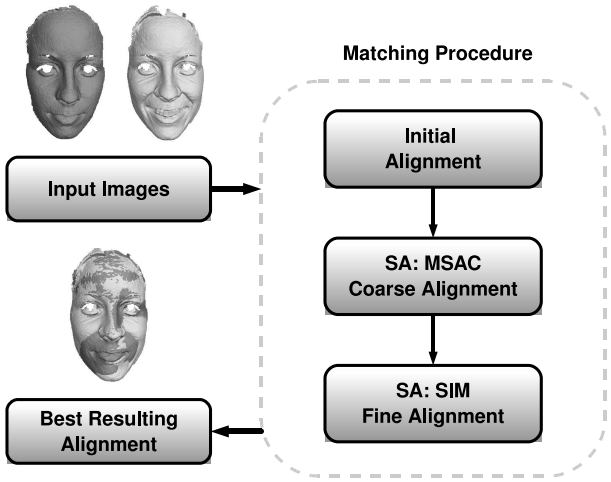


Fig. 6. SA main stages: (1) Initial solution, (2) Coarse alignment using MSAC, and (3) Fine alignment using SIM.

respectively. With these values for t_0, approximately 60% of worst solutions are accepted. According to Rayward-Smith *et al.* [31], this is a criterion that should be met for initial "*temperature*". The stop criterion can be attested if best solution does not change within k iterations (*i.e.* the system is frozen).

The search procedure is performed by adding small random values within $[-1, 1]$ to each element of transformation vector in an attempt to reach better neighbor solutions. In our experiments we observed that not all valid points are required to obtain a good alignment; instead, a sampling rate s of valid points equally spaced is used. Consequently, it is possible to boost algorithm speed and still achieve good alignments. For our experiments we defined $s = 20$ and $s = 30$, for stages (2) and (3) respectively, and we defined $k = 200$. Although we only use a sample of points, the final SIM value that assesses alignment is computed taking into account all valid points from both surfaces. The MSAC threshold was defined as $T = 3.0$, which represents the inlier boundary distance for corresponding points. These threshold values were empirically defined using a set of images from FRGC v1 database, and then validated on a larger dataset.

5.3 Modified SA approach for 3D Face Matching

We observed that the matching between a neutral face and another with expression may lead to imprecise alignments. These alignments produce low SIM values that consequently affect the authentication process. Consequently, and also because happy expression is the most common when people take pictures, we present a modified SA procedure to improve results in these situations. By observing the faces with happy expression one can see that the most affected regions are the cheek and mouth areas.

The main idea of this approach is to guide the matching of a neutral face with the other one with expression to areas that have more expression invariant (*e.g.* the nose and forehead regions). Other authors have employed similar approaches. Martinez [26] proposes a method for 2D face recognition to deal with the occlusion problem. The main idea is to segment the face image into several regions, and analyze each one of them individually. Faltemier *et al.* [14] segment the face into 28 regions, which are aligned with their corresponding ones. Then the regions that presented fewer disturbances are combined to produce the authentication score.

During the SA registration process, the matching is guided to be stronger on those invariant areas. This approach is performed by dividing a face in nine sectors, based on feature points detected by the end of preprocessing stage, as presented in Section 4. Fig. 7 shows the nine sectors and invariant regions used during matching.

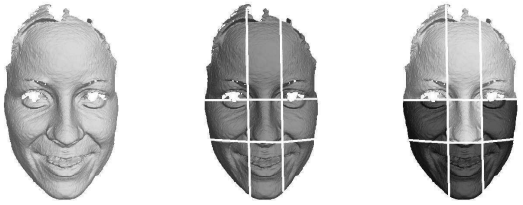


Fig. 7. Sectors used for the modified SA approach: (a) original approach, (b) face sectors, and (c) brighter sectors used for matching.

In the original matching process, the SIM is computed based on the number of interpenetrating points from a surface A in respect to B. In this modified approach, each interpenetrating point q that belongs to one of the invariant regions receives a high weight (*i.e.* $w = 500$), otherwise, it receives a penalty, $w = 0$. The final SIM score is still achieved by setting to all interpenetrating points with $w = 1$.

The modified SA approach for face matching is more suitable when we have to compare a neutral face image with others having facial expressions. One possibility is to apply this technique when the normal face matching fails during recognition process. This can be performed by using a hierarchical model [12]. Results using this approach can be verified in Section 6.

6 3D Face Authentication

To accomplish 3D face authentication we used the SIM scores computed by matching segmented facial regions. If two face images are from the same person, the matching produces a high SIM value; otherwise, it computes a low SIM value. We performed one first experiment to verify which region produced the best result for face recognition. Each image from the neutral noiseless dataset was matched to all the other neutral noiseless images in the FRGC v2 database, and the verification rate was computed at 0% FAR. Table 2 shows verification

results using only each region. The first column describes the face region used, followed by the verification rate. The last column presents the number of false rejections, *i.e.*, number of combinations from the same subjects that were rejected.

TABLE 2
Verification rate for each face region, at 0% FAR.

Regions	Verification Rate	False Rej.
Nose circle (C)	87.4%	627
Nose ellipse (E)	89.6%	519
Upper head (U)	85.0%	749
Face region (F)	89.6%	516
Face using modified SA (M)	87.3%	631

By observing these results, the upper head region presented the highest verification rate mainly because the mouth, which is largely affected by expressions, was removed. Results show that the modified SA procedure presented inferior results when compared to the original SA approach. However, in [30] we stated that modified SA was more suitable when comparing neutral face images to others with expression; and an improvement of 15% was reported over original SA approach. Then, we performed a second experiment, in which we combined scores for each region using the sum rule. The value obtained from the sum is then used as similarity score for two faces. Results are presented in Table 3.

The best performance was achieved combining all regions. Some other works also employ the sum rule to combine different measures [8], [14], [24]. In fact, Kittler *et al.* [21] prove that the sum rule outperforms other combination classifiers. By including the SIM computed for the modified SA matching into the sum rule, the number of false rejections was reduced from 56 to 39 cases; therefore, we decided to use the metric $C + E + U + F + M$ as the similarity score of our method. This score is also suitable for both identification and verification systems. In fact, other approaches can be used to combine classifiers, *e.g.*, product rule, which is employed by [8].

6.1 Hierarchical Evaluation Model

We have developed a hierarchical model to be applied in verification systems using 0% FAR [12]. In this approach, one region is analyzed only if the matching score of the previous one was not sufficient to determine whether images belong to the same subject or not.

The matching classification is based on two thresholds: (1) recognition threshold, and (2) rejection threshold. If the matching score is higher than the recognition threshold, both images are assumed to belong to the same person. If lower than a rejection threshold; images are labeled as being from different subjects. In case the matching score lies between the thresholds, no affirmation can be made and the next region of hierarchy is then used for a new classification attempt. We employed the same strategy proposed by Drovetto *et al.* [12] to define rejection and acceptance thresholds, where these values are automatically defined to avoid false rejection and false acceptance, respectively. As a result, the hierarchical evaluation model is more appropriated when dealing with systems that require a 0% FAR.

TABLE 3
Verification rate combining all face regions, at 0% FAR.

Comb. Rules	Verification Rate	False Rej.
C+E	92.6%	369
C+U	95.6%	220
C+F	95.5%	225
C+M	96.4%	178
E+U	97.7%	113
E+F	97.7%	112
E+M	97.3%	136
U+F	88.9%	553
U+M	86.7%	654
F+M	89.8%	508
C+E+U	97.9%	105
C+E+F	97.1%	144
C+E+M	97.4%	131
C+U+F	98.5%	76
C+U+M	98.2%	88
C+F+M	98.3%	86
E+U+F	99.1%	44
E+U+M	99.0%	48
E+F+M	98.9%	56
U+F+M	89.0%	547
C+E+U+F	98.9%	56
C+E+U+F+M	**99.2%**	**39**

This method aims to achieve high verification rates and keep execution time short. The order in which each region is processed plays an important role, since smaller ones (*e.g.* nose) can be computed faster than larger ones. Consequently, the average execution time tends to be near the necessary to match smaller regions (usually around 4 seconds). In the last hierarchy level, the matching score of all regions are combined using the sum rule. This approach can boost verification rate because sometimes one single region can lead to the correct result while the combination of all regions cannot. This particular situation can be observed on images that have hair occlusion, noise or facial expression [12].

6.2 Extended Hierarchical Evaluation Model

We also propose an extended version of the hierarchy approach. In each step of the hierarchy, instead of assessing only the matching score obtained at that level, the sum of all levels computed to the moment are also used. The matching hierarchy was defined as the following levels:

(1) circle, (2) elliptical, (3) upper head, (4) face region, and (5) face region using the modified SA. At the end of the hierarchy, the sum of all regions is used to verify whether images belong or not to the same subject.

Fig. 8 illustrates how this approach is organized. In total, 28 measures are computed and the sum of the five SIM values is the last one to be assessed. The computed measures for each level are shown in Table 4.

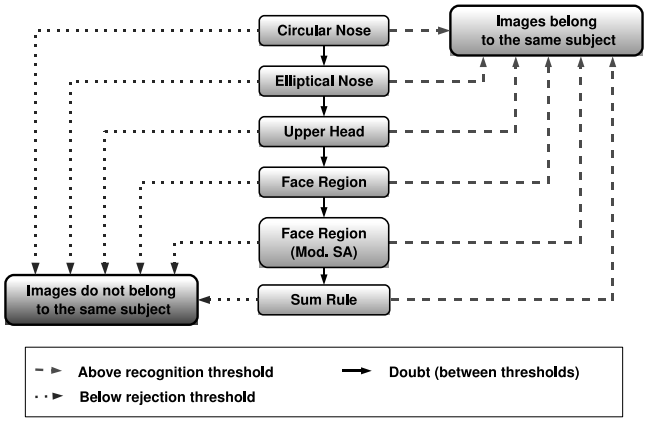


Fig. 8. Hierarchical Evaluation Model Diagram [12].

TABLE 4
Scores computed at each level of the hierarchy.

Hierarchy	Evaluation Metrics
C	C
E	E, C+E
U	U, C+U, E+U, C+E+U
F	F, C+F, E+F, U+F, C+E+F, C+U+F, E+U+F, C+E+U+F
M	M, C+M, E+M, U+M, F+M, C+E+M, C+U+M, C+F+M, E+U+M, E+F+M, U+F+M,C+E+U+M, C+E+U+F+M

7 Experimental Results

Experiments were performed intending to reproduce a genuine face recognition system, where matching is performed between an acquired face image with others stored in a database. The experiments assessed two types of authentication systems: verification and identification. In verification systems, the objective is to answer if a person is who he/she claims to be. The performance is measured assessing the verification rate achieved at a predefined FAR. When using a 0% FAR, it is possible to ensure that any non-authorized person is going to be accepted in the authentication system. Identification systems retrieve the identity of a person comparing the input image with the entire database. The system outputs a rank of the most similar faces based on similarity score. Performance is measured using rank-one accuracy, which indicates when the best matching score is actually from the same subjects [6].

For experiments, the FRGC v2 database was segmented into several controlled datasets, each one with an increasing level of difficulty, *i.e.* addition of images with expression and artifacts. Table 5 shows the datasets followed by their respective description, based on classification described in Table 1. We performed a comparison of each image with all other remaining ones in the database, totaling more than forty million combinations. Then, all the matching results were divided into several datasets for a detailed analysis. All combinations were split into several groups and executed in parallel. For these experiments the label of subject 04643 was changed to 04783. The FRGC organizers confirmed that they actually are the same person (see Section 3).

TABLE 5
FRGC v2 Datasets Classification.

Dataset	Description	No. of images
Level 0	Neutral and noiseless face images.	933
Level 1	All face images with neutral expression.	2,182
Level 2	Includes neutral faces and images with small expression without artifacts.	2,637
Level 3	Includes all images with neutral expression and small expression with and without artifacts.	2,884
All	Includes all images in the database.	4,007

7.1 SA-based approach performance

The time constraint is one of the main worries when using SA to perform 3D face matching in an authentication system. We performed an experiment using a controlled environment for time analysis. A set of 200 random images were selected and combined with each other, totaling 19,900 combinations. We used a computer with the following configuration: Linux O.S., Intel Pentium D 3.4GHz, cache of 2MB and 1GB of memory. Table 6 presents the average execution time of the SA approach regarding the matching of two face regions.

TABLE 6
Average time performance using SA for face matching.

Regions	Average time
Nose circle	1.5s
Nose ellipse	1.2s
Upper head	2.0s
Face region	3.2s
Face using modified SA	3.1s

7.2 Experiment I: Verification

In this experiment we defined a controlled dataset as gallery and the other ones as probe. The number of images of each dataset is presented in Table 5. Each image

from the probe dataset was compared with the ones presented in gallery. Results are presented in Table 7. The first two columns are the gallery and probe datasets used in experiment, respectively. The third column is the verification rate at 0% FAR, followed by the number of false rejections FR. Last two columns show the verification rate at 0.1% FAR and number of false rejections FR, respectively.

TABLE 7
Experiment I: Verification rate using 0%and 0.1% FAR.

Gallery	Probe	FAR 0%	FR	FAR 0.1%	FR
Level 0	*Level 0*	99.2%	39	100.0%	0
Level 0	*Level 1*	98.9%	71	99.9%	3
Level 0	*Level 2*	97.2%	216	99.8%	18
Level 0	*Level 3*	96.5%	294	99.7%	29
Level 0	*All*	90.7%	1,067	98.5%	175
Level 1	*All*	83.5%	4,312	98.2%	475
All	*All*	70.7%	13,736	96.5%	1,648

The experimental results analysis show that when faces with expression and artifacts are added to the datasets, the verification rate is considerably affected on both 0% FAR and 0.1% FAR. This behavior is expected since we perform matching assuming that the face is a rigid object. Results from Table 7 differ from the ones presented in [30] because we changed the SA parameters aiming to reduce its computational time.

We also performed the *"All vs. All"* experiment, where every image of FRGC v2 is matched with all remaining others. It resulted 16,052,042 combinations, from which 46,920 were comparisons from same subjects. In this experiment a 95.6% authentication rate was achieved at 0.1% FAR. Fig. 9 presents the corresponding FAR curve for this experiment.

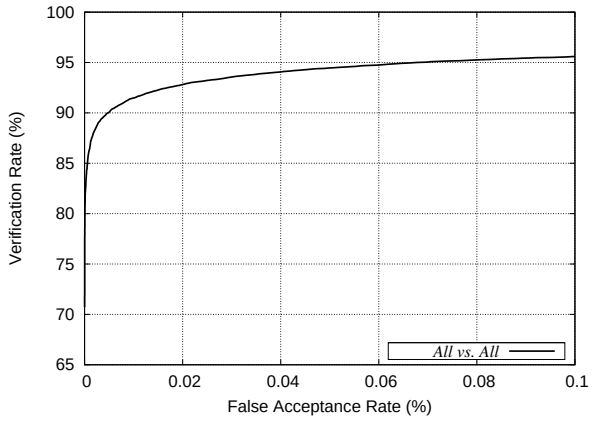


Fig. 9. System performance for "All *vs.* All" experiment.

7.2.1 Eye Holes Disturbances

We have performed experiments to verify if the presence of eye holes would affect the system performance. When the eyes are open, it happens to create holes due to this region specularity, and when the eyes are closed, the eyelashes may not be correctly scanned. We used only the *Level 0* dataset in this experiment to verify if the eye holes really affect the matching procedure. This dataset was segmented in: (1) without holes, (2) small holes, and (3) large holes. Totally 7,400 comparisons were computed by comparing all images without holes against those ones with holes.

For the upper head and face regions, a verification rate of 97.4% and 94.7% were computed, respectively. Both circular and elliptical nose regions achieved a verification rate of 100.0%. However, when using the sum rule, a verification rate of 100% was achieved, using 0% FAR. In fact, the eye holes may disturb the upper head and face matching procedure. But since we use the nose region, the eye holes disturbances can be avoided. One way to avoid it is to use the hierarchical model, described in Section 6.1.

7.3 Experiment II: Identification

For the identification experiment, we defined probe and gallery datasets. In this experiment, we defined four datasets as gallery: (1) *Level 0*, with 933 images, (2) *Level 1*, with the $2,182$ images, (3) *First* with 465 images corresponding to the first image from each subject that appears into the database, and (4) *All*, with $4,007$ images. The probe dataset for each gallery includes all the remaining images from the database which are not in the gallery dataset. From the probe dataset, only the subjects with a corresponding image in the gallery were evaluated. For the experiment using *All* dataset, each subject from probe was compared with the other $4,006$ remaining ones stored in the database. The Cumulative Match Curve (CMC) is presented in Fig. 10 and results achieved for rank-one are presented in Table 8.

Comparing a 3D face image against a huge database may not be very practical due to time constraints using the SA or even the ICP. Faltemier *et al.* [14] considered that time delay would be feasible in verification scenarios, *e.g.*, airports check-in. Besides, there are options to deal with the time-consuming issue. In our experiments, we include some improvements to speed up SA, such as adding a pre-alignment stage based in the center of mass and using only a sample of the image points to perform the alignment. Mian *et al.* [28] propose a rejection classifier, which eliminates 97% of images from the database. Then, the 3D face comparison are applied only to small subsets.

By observing the experimental results, the proposed approach achieved a rank-one recognition higher than 98%. The experiment *All vs. All* presented the highest recognition rate, because one subject may have more than one image stored into the gallery dataset. In fact, Faltemier *et al.* [13] show that by using a gallery composed of multiple images it is possible to obtain higher performance.

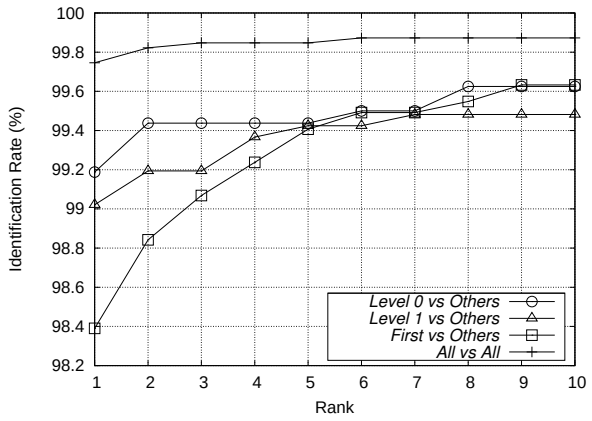


Fig. 10. Plot of CMC curves: (a) Level 0 *vs.* Others, (b) Level 1 *vs.* Others, (c) First *vs.* Others, and (d) All *vs.* All.

TABLE 8
Experiment II: Rank-one recognition rate.

Gallery	Probe	Rank-one	False Rej.
Level 0	*Others*	99.2%	13
Level 1	*Others*	99.0%	17
First	*Others*	98.4%	57
All	*All*	99.7%	10

7.4 Experiment III: FRGC ROC III

This experiment is suggested by the FRGC program, as described in [29]. The gallery dataset was defined with images from fall 2003 and the probe dataset containing images taken during spring 2004. The time interval between both datasets makes its feasibility more difficult. For this experiment, we performed both verification and identification experiments and results are shown in Table 9. In the verification experiment, when using 0 % FAR an authentication score of 77.2% was achieved. For the identification experiment, a rank-one accuracy of 99.6% was achieved. Fig. 11 presents FAR and CMC curves for this experiment.

TABLE 9
Experiment III: Results for FRGC ROC III.

Experiment	Recognition Rate
Verification (0% FAR)	77.2%
Verification (0.1% FAR)	96.6%
Identification	99.6%

7.5 Experiment IV: Hierarchical Evaluation Model

Results for verification experiments at 0% FAR can be improved by employing the hierarchical-based approach for range image matching. Table 10 shows a comparison between performance when using the sum rule as assessment measure, the hierarchical evaluation model, H, as proposed in [12], and the extended hierarchical approach, EH, described in Section 6.1.

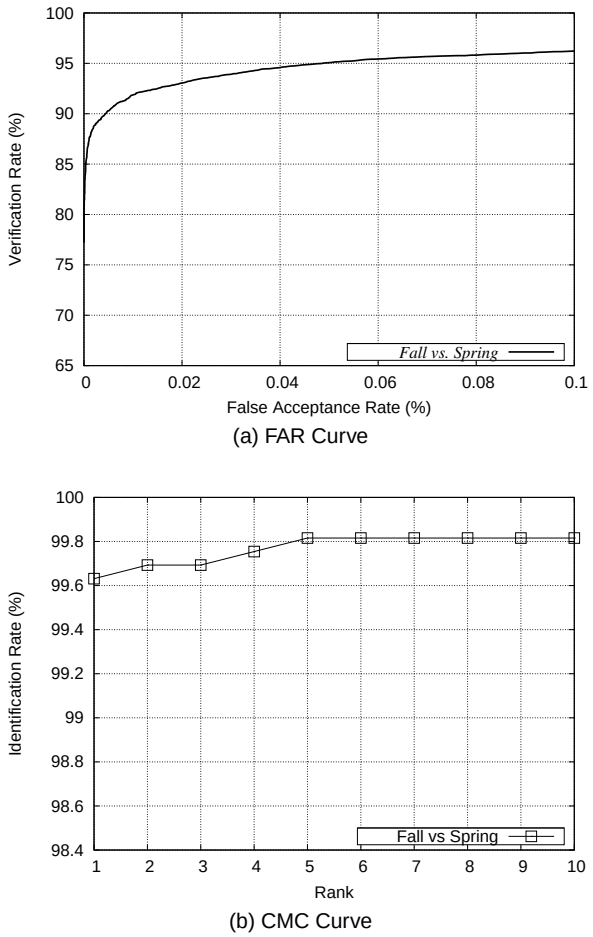


Fig. 11. Performance on experiment FRGC ROC III: (a) FAR Curve, and (b) CMC Curve.

TABLE 10
Experiment IV: Verification rate at 0% FAR.

Datasets	Sum	H [12]	EH
Level 0 vs. Level 0	99.2%	99.3%	99.4%
Level 0 vs. Level 1	98.9%	99.0%	99.2%
Level 0 vs. Level 2	97.2%	97.6%	98.0%
Level 0 vs. Level 3	96.5%	96.9%	97.3%
Level 0 vs. All	90.7%	91.6%	92.4%
Level 1 vs. All	83.5%	86.1%	88.0%
Fall2003 vs. Spring2004	77.2%	78.9%	82.4%
All vs. All	70.7%	77.0%	78.5%

Based on these experimental results one can see that the use of the hierarchical evaluation model makes it possible to boost recognition performance. Moreover, our extended approach produced better results when compared with the other two.

7.6 Comparison with other methods

Some authors have published their results using the FRGC v2 database. We reproduced the same experiments using all the images, including those with artifacts, to provide a comparison with the performance of other methods. Besides, we noticed that the recognition results presented in this section were the same with or without this adjustment. Table 11 shows a verification experiment using *All vs. All* dataset. Table 12 presents rank-one results using *First vs. Others* dataset, and results from FRGC ROC III are available in Table 13.

TABLE 11
Verification results using All *vs.* All dataset, at 0.1% FAR.

Method	Verification Rate
Mian *et al.* [28]	86.6%
Maurer *et al.* [27]	87.0%
Cook *et al.* [10]	92.3%
Faltemier *et al.* [14]	93.2%
Our approach	**96.5%**

TABLE 12
Identification results using First *vs.* Others dataset.

Method	Rank-one
Cook *et al.* [10]	92.0%
Kakadiaris *et al.* [19]	97.0%
Faltemier *et al.* [14]	97.2%
Our approach	**98.4%**

TABLE 13
Verification results for FRGC ROC III, at 0.1% FAR.

Method	Verification Rate
Husken *et al.* [18]	86.9%
Lin *et al.* [22]	90.0%
Al-Osaimi *et al.* [3]	94.1%
Faltemier *et al.* [14]	94.8%
Kakadiaris *et al.* [19]	97.0%
Our approach	**96.6%**

For the FRGC ROC III, we performed the same experiment described in [19], which uses two datasets, one with only neutral images, and the other with only those images with facial expression. These datasets were obtained using the classification provided by the FRGC v2 database. Results are presented in Table 14.

The results support that our method produces better performance for the first two experiments when compared with the other works. In FRGC ROC III experiment, our method achieved a score very close to the one presented in [19]. When using datasets classified by expressions, our method has the best performance in the neutral dataset, but slightly lower performance compared to the non-neutral dataset. One of the main reasons is that we perform the matching procedure assuming that the face is a rigid object. Other authors who employ similar approaches to perform matching also report this limitation, and to reduce the expression impact, they focus on small and invariant face areas [8], [14]. Kakadiaris *et al.* [19] can deal better with the expression effects because their framework includes deformable face models in the matching process, which improves the efficiency of their method.

TABLE 14
Verification results for FRGC ROC III experiment in different datasets, at 0.1% FAR.

Dataset	Kakadiaris *et al.* [19]	Our Approach
Neutral Expression	98.5%	99.5%
Non-Neutral Expression	95.6%	94.8%

7.7 Experiment V: Recognition rate improvement

One of the constrains applied to the SA was limiting the number of iterations k (see Section 5.2) to keep the execution time feasible. We performed one experiment to verify whether the recognition rate could be improved, when we allow the SA larger number of iterations. In this experiment the SIM was recomputed by using SA with $k = 2000$ on those matches that were false accepted and false rejected. Since our objective is to improve the SIM matching score, this set of combinations is the one which will probably achieve a higher score.

Initially, this approach was applied in the set of the combinations between different subjects that would be accepted with 0.5% FAR in both *All vs. All* and *Fall vs. Spring* experiments. After, the set of false rejection was defined and the matching using more iteration was performed. Then, the table with all the results was updated and the verification and rank-one recognition rate were computed. Table 15 presents the achieved results.

As observed, when we allow the SA a larger number of iterations, we can outperform the recognition rate in both verification and authentication modes. It can be stated that the SIM presents a high potential when employed as similarity measure for 3D face recognition.

8 Final Remarks

In this paper, we presented an automatic framework for 3D face recognition. Our method uses a Simulated Annealing-based approach for range image registration with the Surface Interpenetration Measure. Initially, the face is extracted from the input image and segmented into four regions. Each region is matched with its corresponding one from an image that was previously enrolled in a database. When performing matching of the entire face region, a modified SA approach is employed

TABLE 15
Experiment V: Recognition rate.

	Dataset	Recognition rate
Verification (0.1% FAR)	*Level 0 vs. Level 0*	100.0%
	Level 0 vs. Level 1	99.9%
	Level 0 vs. Level 2	99.9%
	Level 0 vs. Level 3	99.8%
	Level 0 vs. All	98.8%
	Level 1 vs. All	98.3%
	Fall vs. Spring	97.4%
	All vs. All	97.1%
Identification (rank-one)	*Level 0 vs. Others*	99.5%
	Level 1 vs. Others	99.6%
	First vs. Others	99.6%
	Fall vs. Spring	99.8%
	All vs. All	99.9%

to minimize facial expression effects. The similarity score is attained by combining all the computed SIM values using the sum rule.

We performed extensive experiments on the FRGC v2 database to evaluate the performance of our method. The database was manually segmented into different datasets according to the images facial expression intensity and noise level. Then we made an accurate analysis regarding our method behavior. We also detected a subject with two different identities in the database with this analysis. By using this novel approach one can distinguish whether two face images with neutral expression belong to the same subject or not with a verification rate of 99%, at 0% FAR. From the experimental results we observed that, when comparing a neutral face with other faces with expressions, the performance of our method slightly decreases. When using all images in the database, in the "All *vs.* All" experiment, faces still can be recognized with 96.5% accuracy, at 0.1% FAR.

We also developed an extended approach of the hierarchical evaluation method to perform a verification experiment at 0% FAR [12]. By including partial sum rules during hierarchy analysis, we improved the verification rate from 70.7% to 77.5%, in the "All *vs.* All" experiment. In the FRGC ROC III experiment, we achieved a verification rate of 96.6%, at 0.1% FAR, and a rank-one accuracy of 98.3%. Although our method is affected when images contain facial expression and artifacts, we obtained results very close to the one reported by Kakadiaris *et al.* [19], which employs deformable face models to deal with the facial expressions. Moreover, we performed an experiment to verify whether the recognition rate could be improved if the time constraint is avoided. The SA matching was performed using a larger number of iterations, and the results improved the recognition rate. With this, we show the potential of the SIM as similarity score for 3D face recognition.

In an identification scenario, our method achieves a 98.4% recognition rate when using the "First *vs.* Others" dataset. Also, for all the experiments performed in the identification mode, our method achieved rank-one accuracy greater than 98%. To the best of our knowledge, these are the best results for this experiment reported in the literature.

ACKNOWLEDGMENTS

The authors gratefully acknowledge CNPq, CAPES and FINEP for financial support. Also, the authors would like to thank Dr. Jonathon Phillips, Dr. Kevin Bowyer and Dr. Patrick Flynn for allowing us to use the images.

REFERENCES

[1] "IMAGO Research Group: 3D face recognition homepage," http://www.imago.ufpr.br/3D_Face_Recognition.
[2] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: A survey," *Pattern Recognition Letters*, vol. 28, no. 14, pp. 1885–1906, 2007.
[3] F. Al-Osaimi, M. Bennamoun, and A. Mian, "An expression deformation approach to non-rigid 3D face recognition," *Int'l Journal of Computer Vision*, p. Online first, September, 2008.
[4] O. R. P. Bellon, L. Silva, C. Queirolo, S. Drovetto Jr., and M. P. Segundo, "3D face image registration for face matching guided by the surface interpenetration measure," in *Proc. IEEE Int'l Conf. Image Processing*, 2006, pp. 2661–2664.
[5] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, 1992.
[6] R. Bolle, J. Connell, S. Pankanti, N. Ratha, and A. Senior, *Guide to Biometrics*. SpringerVerlag, 2003.
[7] K. W. Bowyer, K. Chang, and P. J. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition," *Computer Vision and Image Understanding*, vol. 101, pp. 1–15, 2006.
[8] K. I. Chang, K. W. Bowyer, and P. J. Flynn, "Multiple nose region matching for 3D face recognition under varying facial expression," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1695–1700, 2006.
[9] J. Cook, V. Chandran, S. Sridharan, and C. Fookes, "Face recognition from 3d data using iterative closest point algorithm and gaussian mixture models," in *Proc. Int'l Symposium 3D Data Processing, Visualization and Transmission*, 2004, pp. 502–509.
[10] J. Cook, C. McCool, V. Chandran, and S. Sridharan, "Combined 2D/3D face recognition using log-gabor templates," in *Proc. IEEE Int'l Conf. Video and Signal Based Surveillance*, 2006, p. 83.
[11] G. Dalley and P. J. Flynn, "Range image registration: A software platform and empirical evaluation," in *Proc. Int'l Conf. 3-D Digital Imaging and Modeling*, 2001, pp. 246–253.
[12] S. Drovetto Jr., L. Silva, and O. R. P. Bellon, "Hierarchical evaluation model for 3D face recognition," in *Proc. Int'l Conf. Computer Vision Theory and Applications*, vol. 2, 2008, pp. 67–74.
[13] T. Faltemier, K. Bowyer, and P. Flynn, "Using a multi-instance enrollment representation to improve 3d face recognition," in *Proc. Int'l Conf. Biometrics: Theory, Applications, and Systems*, vol. 1, 2007, pp. 1–6.
[14] T. Faltemier, K. W. Bowyer, and P. J. Flynn, "A region ensemble for 3d face recognition," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 62–73, 2008.
[15] P. J. Flynn, T. Faltemier, and K. W. Bowyer, "3D face recognition," in *Handbook of Biometrics*, A. K. Jain, A. Ross, and P. J. Flynn, Eds. Springer, 2008, ch. 11, pp. 211–229.
[16] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy, "Geometrically stable sampling for the ICP algorithm," in *Proc. Int'l Conf. 3-D Digital Imaging and Modeling*, 2003, pp. 260–267.
[17] J. Huang, B. Heisele, and V. Blanz, "Component-based face recognition with 3d morphable models," in *Proc. Int'l Conf. Audio and Video-based Biometric Person Authentication*, 2003, pp. 27–34.

[18] M. Husken, M. Brauckmann, S. Gehlen, and C. V. der Malsburg, "Strategies and benefits of fusion of 2D and 3D face recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*. IEEE Computer Society, 2005, pp. 174–174.
[19] I. Kakadiaris, G. Passalis, G. Toderici, N. Murtuza, and T. Theoharis, "Three-dimensional face recognition in the presence of facial expression: An annotated deformable model approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 640–649, 2007.
[20] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
[21] J. Kittler, M. Hatef, R. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 226–239, 1998.
[22] W.-Y. Lin, K.-C. Wong, N. Boston, and Y. H. Hu, "3d face recognition under expression variations using similarity metrics fusion," in *Proc. IEEE Int'l Conf. Multimedia and Expo*, 2007, pp. 727–730.
[23] X. Lu and A. K. Jain, "Deformation modeling for robust 3D face matching," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, 2006, pp. 1377–1383.
[24] X. Lu, A. K. Jain, and D. Colbry, "Matching 2.5D face scans to 3D models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 1, pp. 31–43, 2006.
[25] M. Lundy and A. Mees, "Convergence of an annealing algorithm," *Mathematical Programming: Series A and B*, vol. 34, no. 1, pp. 111–124, 1986.
[26] A. Martinez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 6, pp. 748–763, Jun 2002.
[27] T. Maurer, D. Guigonis, I. Maslov, B. Pesenti, A. Tsaregorodtsev, D. West, and G. Medioni, "Performance of geometrix activeid 3D face recognition engine on the FRGC data," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*. IEEE Computer Society, 2005, pp. 154–154.
[28] A. Mian, M. Bennamoun, and R. Owens, "An efficient multimodal 2D-3D hybrid approach to automatic face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 11, pp. 1927–1943, 2007.
[29] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2005, pp. 947–954.
[30] C. Queirolo, M. P. Segundo, O. R. P. Bellon, and L. Silva, "Noise versus facial expression on 3d face recognition," in *Proc. Int'l Conf. Image Analysis and Processing*, 2007, pp. 171–176.
[31] V. J. Rayward-Smith, I. H. Osman, C. R. Reeves, and G. D. Smith, *Modern Heuristic Search Methods*. John Wiley & Sons Ltd., 1996.
[32] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *Proc. Int'l Conf. 3-D Digital Imaging and Modeling*, 2001, pp. 145–152.
[33] M. P. Segundo, C. Queirolo, O. R. P. Bellon, and L. Silva, "Automatic 3d facial segmentation and landmark detection," in *Proc. Int'l Conf. Image Analysis and Processing*, 2007, pp. 431–436.
[34] L. Silva, O. R. P. Bellon, P. F. U. Gotardo, and K. L. Boyer, "Range image registration using enhanced genetic algorithms," in *Proc. IEEE Int'l Conf. Image Processing*, vol. 2, 2003, pp. 711–714.
[35] L. Silva, O. R. P. Bellon, and K. Boyer, *Robust Range Image Registration Using Genetic Algorithms and the Surface Interpenetration Measure*, ser. Machine Perception and Artificial Intelligence. World Scientific Publishing, 2005, vol. 60.
[36] ——, "Robust range image registration using the surface interpenetration measure and enhanced genetic algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, pp. 762–776, 2005.
[37] P. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, pp. 138–156, 2000.
[38] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.

Chauã C. Queirolo joined the IMAGO Research Group in 2004. He received the BS degree in Computer Science from Universidade Federal do Parana (UFPR), Brazil, in 2007, and his undergraduate research work received awards from the Brazilian Computer Society. Since 2008, he is a MSc student at IMAGO. He is a student member of IEEE Computer Society and Brazilian Computer Society. His research interests include range imaging, computer vision, computer games and dinosaurs.

Luciano Silva joined the IMAGO Research Group in 1996. He received the BS and MSc degrees in Informatics from Universidade Federal do Parana (UFPR), Brazil, in 1998 and 2000, respectively, and the PhD degree in Electrical Engineering and Industrial Informatics from Universidade Tecnologica Federal do Parana (UTFPR), Brazil, in 2003. From 2002-2003, he did doctoral studies in the Signal Analysis and Machine Perception Laboratory at Ohio State University, USA. Since 2004, he has been with the Informatics Department of UFPR, where he is currently an Associate Professor. His undergraduate work and master dissertation received awards from the Brazilian Computer Society. In 2004, he received the 1st Place Award in the Doctoral Dissertation Contest of the XVII Brazilian Symposium on Computer and Image Processing. His research interests include computer vision, 2D/3D image analysis and assistive technologies. He is a member of IEEE and Brazilian Computer Society.

Olga R. P. Bellon received the BS degree in Electrical Engineering, in 1987, from Universidade Federal do Espirito Santo (UFES), Brazil, and the MSc and PhD degrees from Universidade Estadual de Campinas (UNICAMP), Brazil, in Electrical Engineering, in 1990 and 1997, respectively. In 1995, she founded the IMAGO Research Group at Universidade Federal do Parana (UFPR), where she is a professor. From 1997-2000, she was the head of the Graduate Program of Informatics. From 2000-2002, she was the General Research Coordinator of UFPR. From 2002-2003 she spent a sabbatical year in the Signal Analysis and Machine Perception Laboratory at Ohio State University, USA. In 2004, she was the general chair of the XVII Brazilian Symposium on Computer and Image Processing joint with II Ibero-American Symposium on Computer Graphics. In 2007, Dr. Bellon was elected President of the Special Committee in Computer Graphics and Image Processing of the Brazilian Computer Society. Her main research interests lie in biometrics, pattern recognition, range image analysis, digital preservation of natural and cultural assets and virtual museums, image mining and medical imaging. She is a member of IEEE Computer Society, Brazilian Computer Society and Brazilian Health Informatics Society.

Mauricio Pamplona Segundo joined the IMAGO Research Group in 2005. He received the BS degree in Computer Science from Universidade Federal do Parana (UFPR), Brazil, in 2007, and his undergraduate research work received awards from the Brazilian Computer Society. Since 2008, he is a MSc student at IMAGO. His research interests include research interests include biometrics, automatic features extraction, pattern recognition, image segmentation, image registration and security.