3D FACE IMAGE REGISTRATION FOR FACE MATCHING GUIDED BY THE SURFACE INTERPENETRATION MEASURE

Olga Bellon, Luciano Silva, Chauã Queirolo, Sídnei Drovetto, Maurício Pamplona

IMAGO Research Group - Universidade Federal do Paraná
Caixa Postal 19081, 81.531-980, Curitiba - PR - Brazil
{olga,luciano,ccqs03,sidnei,mps04}@inf.ufpr.br

ABSTRACT

The Surface Interpenetration Measure (SIM) was recently proposed as a promising measure for 3D face matching, although using two limited, small range image databases. In this paper we present novel, more extensive experiments using the SIM in a well-known 3D face database available on the Biometric Experimentation Environment (BEE) to confirm qualitatively that the SIM is a effective, discriminatory measure. The experiments were performed based on range image registration by using two different methods: Iterative Closest Point (ICP) and Simulated Annealing (SA). By computing the SIM after the registration of two 3D face images one can identify if those images come from the same subject or not. With our SA-based approach we obtained high verification rate scores, which is indeed one of the main goals of the Face Recognition Grand Challenge 2006.

Index Terms— Face recognition, Simulated annealing

1. INTRODUCTION

Two-dimensional (2D) face recognition has been the subject of extensive researches for many years. More recently, three-dimensional (3D) face recognition has gained growing attention. The advances in 3D imaging technology (see Intl. Conf. on 3D Digital Imaging and Modeling) have played an important role in this scenario. Also, there are some limitations in processing 2D face images that can be solved by 3D face images, although the last ones have their own limitations too. For a survey on works presented in the literature about 3D face processing the reader should refer to [1].

Recently, some approaches for 3D face recognition have used image registration to perform face matching [2, 3, 4, 5]. In this context, the approaches are usually based on the Mean Squared Error (MSE) or Root Mean Squared Error (RMSE), sometimes combined with other measures, to evaluate the quality of the registration in face matching. In all these approaches, the Iterative Closest Point (ICP) is traditionally applied for image registration.

The authors would like to thank the Conselho Nacional de Desenvolvimento Cientfico e Tecnolgico (CNPq) for the financial support.

The ICP is guided by the MSE but it was proved [6] that this measure could allow imprecise local convergence for range image registration, even when improved ICP-based approaches are used [7, 8]. Also, in [6] the authors suggest that the MSE is a good measure for starting the image registration process, but the Surface Interpenetration Measure (SIM) could be more suitable to be used at "the end of the game" to assess the quality of the registration.

The SIM was first used for 3D face matching recently [9]. Although the presented experiments were performed in a limited number of images, these belong to different databases and the final results were quite similar, showing the potential for robustness of this new measure. Also, from those preliminary results, we noted that the SIM produces a better range for discrimination between faces (including same subjects with different expressions) as compared to other metrics.

In this paper, we present novel, more extensive experiments using the SIM in a very known 3D face image database[1]. Our experiments were based in some state-of-art works presented in the literature [2, 3, 4, 5] and the results confirm the potentiality of the SIM as a discriminatory measure for 3D face matching, which may contribute substantially to this field.

This paper is organized as follows. First, we introduce our developed approach by using the SIM in section 2. The experimental results summarizing many comparisons by using different approaches for 3D face matching are presented in section 3, followed by the final remarks in section 4.

2. 3D FACE MATCHING USING THE SIM

We propose to perform 3D face matching by combining range image registration techniques and the SIM. We evaluated two registration approaches: (1) an improved version of the ICP [7, 8], and (2) a Simulated Annealing based approach (SA) [10]. In our experiments the ICP-based approach is guided only by MSE (*i.e.*, the MSE is the measure to evaluate the alignments). For our SA approach, the MSE is used to start the registration process and then the final tuning is

[1]http://www.bee-biometrics.org

made by the SIM [6, 11]. In both approaches, after obtaining the final alignment for the two images, the SIM is computed to evaluate the final 3D face matching.

Although SA is twice slower than ICP, it does not depend on initial pre-alignment as opposed to ICP. Also, by using the SA we can compute the 3D face matching automatically. As discussed in [6], the MSE alone is not the best measure to perform range image registration, so we decided to develop our registration approach by using both MSE and the SIM. We have previously used Genetic Algorithms for range image segmentation with good, precise results [6, 11] but it is extremely time consuming, which also limits many applications. Considering that range image registration by ICP is widely known, we briefly present the SIM and the SA approach for range image registration as follows.

2.1. The Surface Interpenetration Measure

The SIM was developed by analyzing visual results of two aligned surfaces, each rendered in a different color, crossing over each other repeatedly in the overlapping area [11]. The interpenetration effect results from the nature of real range data, which presents slightly rough surfaces with small local distortions caused by limitations of the acquiring system. Because of this, even flat surfaces present a "roughness" in range images. With this, we can assume that independently of the shape of the surfaces the interpenetration will always occur. We also observed that two images acquired from the same object surface with the same scanner position and parameters provide two different range images.

By quantifying interpenetration, one can more precisely evaluate the registration results and provide a highly robust control [6]. To do this we developed the following measure based on the surface normal vector, computed by a local least squares planar fit, at each point. After the alignment of two images, A and B, we identify the set of interpenetrating points in A with respect to B. For each point $p \in A$ we define a neighborhood N_p to be a small $n \times n$ window centered on p. With q denoting a point in the neighborhood N_p, c the corresponding point of p in image B and $\vec{n}_c$ the local surface normal at c, we define the set of interpenetrating points as: $C_{(A,B)} = \{p \in A \mid [(\overrightarrow{q_i - c})\cdot\vec{n}_c][(\overrightarrow{q_j - c})\cdot\vec{n}_c] < 0\}$; where $q_i, q_j \in N_p$ and $i \neq j$. This set comprises those points in A whose neighborhoods include at least one pair of points separated by the local tangent plane, computed at their correspondents in B. With this, we then define the SIM as the fraction of interpenetrating points in A: $SIM_{(A,B)} = \frac{|C_{(A,B)}|}{|A|}$.

In this paper we applied some constraints to the SIM as suggested in [11] to avoid incorrect corresponding points, and to obtain more precise alignments. We included the constraint $m = 5$ as the maximum angle allowed between the normal vectors at c and p, $\vec{n}_c$ and $\vec{n}_p$, respectively. Then, we have $p \in C$ only if $\cos^{-1}(\vec{n}_c \cdot \vec{n}_p) \leq m$. Also, we used a constraint to eliminate the corresponding points on the surfaces boundaries. In this case, $p \in C$ if $c \notin D$, with D the set of boundary points in B, with the thickness of the boundary defined as $s = 1$;

Registrations of two range images presenting good interpenetration have high SIM values. Our experimental results show that erroneous alignments produce low SIM values and that 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, the SIM is a far more sensitive indicator of alignment quality when comparing "reasonable" alignments [6, 9].

2.2. The SA-based approach for range image registration

Simulated Annealing (SA) is a stochastic algorithm for local search in which, from an initial candidate solution, it generates iterative movements to a neighbor solution that represents a better solution to the problem as compared to the current one. The main difference between SA and other local search algorithm, *e.g.* Hill Climbing, is that SA can accept a worse solution than the current candidate in the iterative process. Then, SA does not remain "tied" to local minima and because of this it has better chances to reach its goal, which is a solution close enough to the global one.

In order to apply SA on registration of two range images, six parameters (three parameters each for rotation and translation relative to a 3D coordinate system) are needed to define the candidate solutions as a "transformation vector" that, when applied to one image, can align it with the other.

Our SA-based approach was developed by using the GSL Library[2] and basically has three main stages: (1) an initial solution is obtained by aligning the centers of mass of the two face images; (2) a coarse alignment is performed using a SA-based searching procedure to minimizes a robust evaluation measure, based on the MSAC robust estimator [12] combined with the MSE of the corresponding points between the two face images; and (3) a precise alignment is obtained by a SA-based searching procedure with the SIM as the evaluation measure, where the goal is to maximize the interpenetrating points between the two faces. The termination condition in stage 2 was set empirically at $r = 1000$ iterations, but if the best solution does not change within $h = 300$ iterations (*i.e.* the system is frozen) the termination condition is also achieved. In stage 3 we set $r = 500$ and $h = 35$.

In the SA-based searching procedure, small random values within $[-1, 1]$ are introduced to each element of the transformation vector in an attempt to reach better neighbor solutions. To evaluate the quality of the obtained alignments, the above described metrics (*i.e.*, MSAC and SIM) are used at each stage of our approach. In stage 2, to evaluate the alignments we used a sampling rate of 3.25% of the valid points for each region. In fact, it was not necessary to use all image pixels to obtain a good coarse alignment.

[2]GNU Scientific Library - http://www.gnu.org/software/gsl

For stage 3 we use all valid pixels for each region because we are searching for precise solutions. The initial solution for the second stage corresponds to the best solution obtained on the first stage. In both stages, the "temperature" of the SA is reduced very slowly and only one iteration is performed for each allowed "temperature" until the final one is achieved, as suggested by [13] otherwise the system becomes frozen. The initial "temperature" was defined as $t = 0.002$ and $t = 0.15$ for stage 2 and stage 3, respectively. Based on experimental results the threshold value for the MSAC was empirically defined as 0.7, which represents an inlier boundary distance for the corresponding points between images.

3. EXPERIMENTAL RESULTS

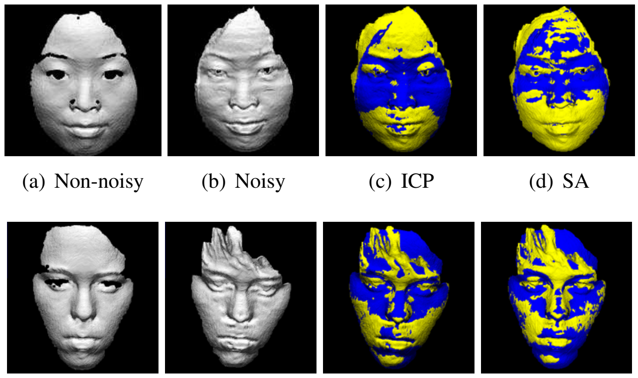
The experiments were performed using 778 images with 640x480 pixels each one, from the BEE 3D face database[3]. We matched each image against all others, totalizing 302253 combinations (300988 from different subjects and 1265 from the same subject). We observed that some images present facial expression or noise perturbation, but we did not perform a detailed analysis. The noise effect is shown in Fig. 1 and we believed it is due to the acquisition process. To minimize the noise we applied a 5×5 median filter on the images.

Each image was segmented automatically into three regions: (a) face region, (b) eyes and nose region and (c) nose region. For regions (b) and (c) we used a similar approach suggested in [2, 3, 5, 4] and both regions consist of rectangular areas, the first one around the eyes and nose and the latter around the nose. For region (a) we used our own approach based on the depth of the range maps. Although this approach has not been published yet and details of the algorithm have been also omitted here because of space limits, all segmentation results were successfully validated by visual inspection.

The following figures present the obtained alignments from both registration approaches, where each image was rendered using a different color to show the interpenetrating areas. Fig. 1 shows two examples of face matching using noisy images. The SIM results are very low due to the poor quality of the image and both approaches, ICP and SA, were compromised. As can be seen, it is difficult to assume that the matching faces (*e.g.* Figs. 1(a) and 1(b)) came from the same subject. It seems that the noisy images were obtained with low resolution producing "flatness" surfaces.

In Fig. 2 we show an example of the improvement obtained by our SA method against the ICP, which presents low interpenetrating points. In contrast, our SA method is capable to obtain a precise alignment in the forehead region, where both views are most similar (see also Fig. 3). Finally, Fig. 3 shows an example where both alignments presented low SIM values because of the facial expression in one image.

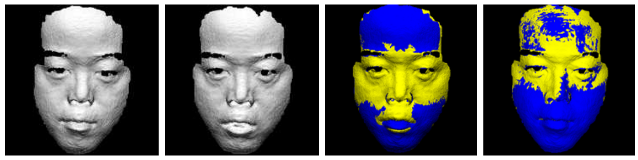
[3]The authors would like to thanks Dr. Jonathon Phillips and Dr. Patrick Flynn for allowing us to use the images.



(a) Non-noisy (b) Noisy (c) ICP (d) SA

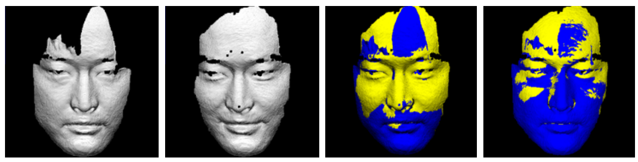
(e) Non-Noisy (f) Noisy (g) ICP (h) SA

Fig. 1. Two samples of face matching of different subjects. In both cases one of the images has noise perturbations: (a) and (e) faces without noise; (b) and (f) faces with noise; (c) and (g) ICP alignments, SIM=4.00% and SIM=6.92%, respectively; and (d) and (h) SA alignments, SIM=8.88% and SIM=11.69%, respectively.



(a) Neutral (b) Expression (c) ICP (d) SA

Fig. 2. Face matching from a same subject using images without noise, but one with slight facial expression: (a) neutral expression; (b) slight opened-mouth expression; (c) ICP alignment, SIM=4.92%; and (d) SA alignment, SIM=23.55%.



(a) Neutral (b) Expression (c) ICP (d) SA

Fig. 3. Face matching from a same subject using images without noise but one with facial expression: (a) neutral expression; (b) smiling expression; (c) ICP alignment, SIM=2.91%; and (d) SA alignment, SIM=11.05%.

Once we obtained all the image registration results, we automatically defined a discriminatory threshold that allowed a FAR (False Acceptance Rate) of 0% for each segmented region. The threshold used for the ICP was: SIM=14.5% for face region; SIM=11.5% for eyes and nose region; and SIM=18.5% for nose region. For our SA approach the defined values were: SIM=16.75% for face region; SIM=14.74% for eyes and nose region; and SIM=22.5% for nose region. As can be seen in Fig. 3, although the SIM values obtained by our SA approach are much higher than by the ICP, for the face region, it was not sufficient to cross the discriminatory threshold that allowed a FAR of 0% (*i.e.* SIM=16.5%). However, we could identify facial expressions or noisy images by analyzing the SIM values as reported in [9].

Table 1 presents the verification rate result for both methods using different regions extracted from the face. By using a FAR of 0% we could eliminate the possibility of identifying an individual that is not authorized in an authentication system with a verification rate of 97.31% by using our SA method applied to the face region. The best result for the ICP-based method was 93.83% for the nose region.

Table 1. Verification rate for each region using a FAR of 0%.

	ICP	SA
Face	84.66	97.31
Eyes	91.86	95.81
Nose	93.83	93.75

Table 2. Verification rate using a FAR of 0% without the noisy images.

	ICP	SA
Face	85.70	98.62
Eyes	92.61	96.59
Nose	94.96	95.12

We observed that the eyes and nose are rigid regions [3, 5] and suffers less with facial expression than the whole face region. However, many times the entire face is more discriminatory than the others. Also, experiments using the SIM show that the nose is not enough discriminatory because it represents only a small fraction of the face.

By analyzing the combinations that compose the false rejection rate (FRR) we observed that most of them include at least one image that suffers from noise (see Fig. 1) or facial expression (see Fig. 3). If we remove only the noisy images the verification rate reaches 98.62% at a FAR of 0% for our approach, as in Table 2. The experimental results using our SA-based method presented promising results and confirmed that the SIM can be very effective for the 3D face matching problem. In fact, according to [14] the performance goal for the Face Recognition Grand Challenge 2006 is a verification rate of 98% at a FAR of 0.1%.

The results shown that our approach returns better results for both face and eyes regions than ICP does. For the nose region the results are quite similar for both methods. The ICP method fails to produce precise alignments for face region because the convergence of the algorithm is driven by MSE and the registration process usually get stuck in local minima [6].

4. FINAL REMARKS

In this paper we introduced a novel, robust approach for 3D face matching based on Simulated Annealing (SA) and the Surface Interpenetration Measure (SIM). The experimental results showed the effectiveness of our approach, which has a verification rate of 97.31% at a FAR (False Acceptance Rate) of 0% and 98.62% at FAR of 0% when combinations with noisy images were eliminated from the FRR (False Rejection Rate) set. Although the use of the SIM in our SA-based approach had presented better results than the ICP-based one, we proved that the SIM, when used as a 3D face matching measure, works well even with the alignments obtained by the ICP. As a future work we plan to use the SIM to identify facial expressions and to develop a robust approach for the 3D face identification problem using a combination of face regions.

5. REFERENCES

[1] K. W. Bowyer, K. I. Chang, and P. J. Flynn, "A survey of approaches to three-dimensional face recognition," in *Proc. of 17th Intl. Conf. on Pattern Recognition*, 2004, pp. 358–361.

[2] K. I. Chang, K. W. Bowyer, and P. J. Flynn, "Adaptive rigid multi-region selection for handling expression variation in 3D face recognition," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition - Workshops*, 2005, vol. 3, p. 157.

[3] K. I. Chang, K. W. Bowyer, and P. J. Flynn, "Effects on facial expression in 3D face recognition," in *Proc. SPIE - Biometric Tech. for Human Identification*, 2005, vol. 5779, pp. 132–143.

[4] X. Lu, A. K. Jain, and D. Colbry, "Matching 2.5D face scans to 3D models," *IEEE PAMI*, vol. 28, no. 1, pp. 31–43, 2006.

[5] X. Lu and A. K. Jain, "Integrating range and texture information for 3d face recognition," in *Proc. of 7th IEEE Workshop on Applications of Computer Vision*, 2005, pp. 156–163.

[6] L. Silva, O. R. P. Bellon, and K. L. Boyer, "Robust range image registration using the surface interpenetration measure and enhanced genetic algorithms," *IEEE PAMI*, vol. 27, pp. 762–776, 2005.

[7] S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *Proc. of 3rd Intl. Conf. on 3D Digital Imaging and Modeling*, 2001, pp. 145–152.

[8] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy, "Geometrically stable sampling for the ICP algorithm," in *Proc. of 4th Intl. Conf. on 3D Digital Imaging and Modeling*, 2003, pp. 260–267.

[9] L. Silva, O. R. P. Bellon, and C. C. Queirolo, "3D face matching using the surface interpenetration measure," in *Lecture Notes in Computer Science*, vol. 3617, pp. 1051–1058. Springer-Verlag, 2005.

[10] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[11] L. Silva, O. R. P. Bellon, and K. L. Boyer, *Robust Range Image Registration Using Genetic Algorithms and the Surface Interpenetration Measure*, vol. 60 of *Machine Perception and Artificial Intelligence*, World Scientific Publishing, 2005.

[12] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *CVIU*, vol. 78, pp. 138–156, 2000.

[13] M. Lundy and A. Mees, "Convergence of an annealing algorithm," *Mathematical Programming: Series A and B*, vol. 34, no. 1, pp. 111–124, 1986.

[14] P. J. Phillips, P. J. Flynn, W. T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. J. Worek, "Overview of the face recognition grand challenge," in *Proc. of IEEE CVPR*, 2005, pp. 947–954.