Matching Score based Face Registration

B.J.Boom, G.M. Beumer, L.J. Spreeuwers and R.N.J. Veldhuis
b.j.boom@utwente.nl
University of Twente, EEMSC, Signals & Systems
P.O. box 217, 7500 AE, Enschede, The Netherlands

Abstract— **Accurate face registration is of vital importance to the performance of a face recognition algorithm. We propose a new method: matching score based face registration, which searches for optimal alignment by maximizing the matching score output of a classifier as a function of the different registration parameters (translation, rotation, scale). We compare this method with our previously developed methods, namely MLLL based on maximizing the likelihood ratio in combination with BILBO which corrects outliers in the found landmarks and a Viola-Jones based landmark detector. We determine the accuracy of the registration methods and give an indication of the speed of the methods. Futhermore, we investigate the influence of the registration on the task of face verification.**

I. Introduction

Several papers have shown that good registration is essential for a good face-recognition performance [1],[2]. We have developed a new face registration approach, which is not based on landmarks but makes use of the face recognition algorithm as a evaluation criterion to improve the face registration. By using a simple search algorithm we vary the 4 alignment parameters, namely scale, rotation and translation in x- and y-direction. The new face alignments are evaluated by our face registration algorithm and, using the search algorithm, we find the optimal face registration.
In section II we explain the first implementation of our new approach. In section III we describe our experimental setup and the results. We compare them to some landmark based methods in RMS error and EER. The final section contains a discussion about the matching score based face registration.

II. Matching Score based Face Registration

Matching Score based Face Registration (MSFR) is based on the assumption that a face classifier will give a higher output (matching score) if the face is better aligned to the reference face. To verify this assumption we have performed several experiments in which we varied one of the registration parameters (translation, rotation and scale) of a manually labelled face image. A typical example is given in Figure 1 which shows 4 graphs, where we determined the matching score for a single face image while varying one of the parameters. The parameters are varied relative to a registration based on manually labelled landmarks. In the graphs this corresponds with scale = 1, angle = 0 and translation in x- and y-direction = 0. Figure 1 shows that the manual registration is rather good although the matching score can be improve by using a slightly different translation in x-direction.

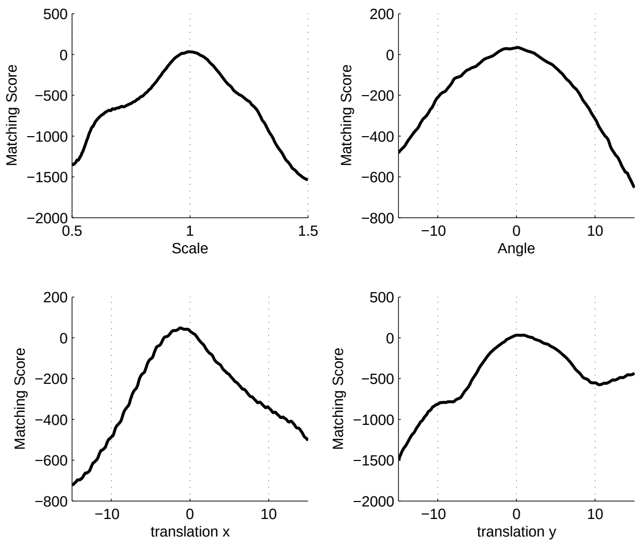

Fig. 1. Matching Score of a face while varying one of the parameters

In a fully automated system we use the location given by a face detection algorithm [3] instead of the manually labelled data. Using this region as a starting point, we try to find a better alignment of the face. In our case, the matching score is determined by the face recognition algorithm based on the log-likelihood ratio [4]. This face recognition algorithm is trained on a manually labelled training set. We maximize the matching score using a search algorithm, which varies the translation, rotation and scale parameters of a face image to find the optimal alignment. A schematic representation of the entire system is given in Figure 2.

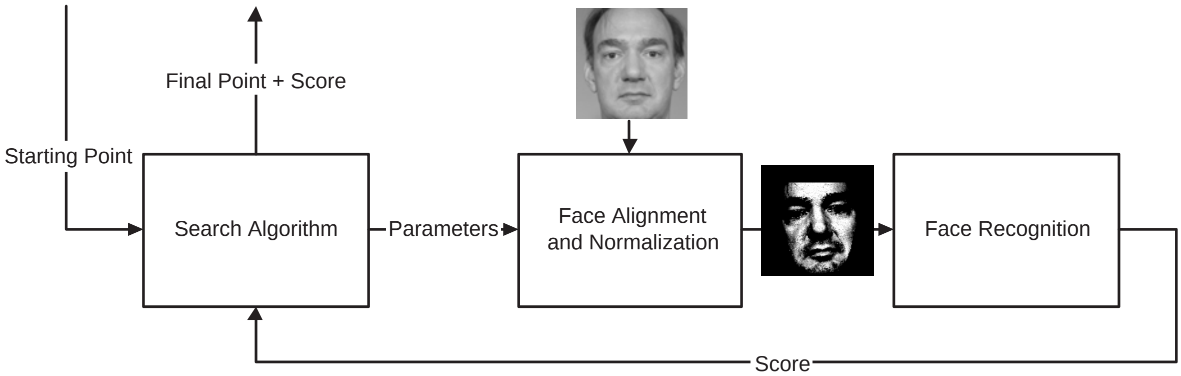

Fig. 2. Scheme of matching score face registration

A. Search Algorithm

We have performed experiments with two search algorithms, namely the Nelder-Mead search algorithm [5] and a variant of the Powell-Brent search algorithm [6],[7]. Best results are obtained using the first algorithm, which seems to be less sensitive to local minima. In this paper we will therefore only focus on the results of the search algorithm proposed by Nelder-Mead in [5], also known as downhill simplex method. The starting point of the search algorithm in our case is the location given by the face detection algorithm. The search algorithm will search in 4 dimensions (translation in x- and y-direction, rotation and scale), where the first simplex (geometrical figure in N dimensions consisting of $N+1$ points) is created from the starting point and four points for which we varied a single parameter. For the translations we used a step of 5 pixels, for rotation of 5 degrees and for scale we multiplied with a factor of 1.2.

B. Face Alignment and Normalization

Using the parameters given by the search algorithm and the face image found by the face detection algorithm the face is aligned using a rigid transformation and a scaling transformation. After that, we select a region of interest (ROI) and we normalize the image inside the ROI to zero mean, unit variance. An example of the output of our face alignment and normalization process can been seen in Figure 2.

C. Face Recognition

For face recognition we do feature reduction by subsequently performing PCA [8] and LDA [9]. We use the algorithm proposed in [4] which uses the log-likelihood ratio to classify face images. For a certain class i the similarity score S is calculated by:

$$\begin{aligned} S_{y,i} &= -(y-\mu_{W,i})^T \Sigma_W^{-1} (y-\mu_{W,i}) \\ &\quad + y^T \Sigma_T^{-1} y - \log|\Sigma_W| + \log|\Sigma_T| \end{aligned} \qquad (1)$$

Here y is a vector which is a representation of the face image after feature reduction, Σ_T is the total covariance matrix, Σ_W is the within class covariance matrix and $\mu_{W,i}$ is the ith class average.

III. Experiments and Results

A. Experimental Setup

In our experiments we use the FRGC version 1 database [10]. We used 3772 images in which the face was correctly found by Viola-Jones [3]. These images were taken under controlled conditions and contain 273 individuals. The database is randomly split into two subsets, each consisting of approximately half of the images of each person. One subset is used for training and enrollment, the other is used for testing. The results of the face recognition are measured in Equal Error Rate (EER), at the point of operation where False Accept Rate (FAR) is equal to the False Reject Rate (FRR). To get a better estimate of the expected EER, we repeat the experiments 33 times, randomly splitting the datasets so other subsets are used in the training and test set.

We compare our face registration method with the landmark based face registration methods in [11], namely a Viola-Jones landmark detector and the combination of MLLL and BILBO. We used the BIOID database [12] to train MLLL+BILBO and the Viola-Jones landmark detector. After face registration, the face recognition using landmarks is trained with face images which are aligned using the automatically found landmarks. This makes the algorithm more robust against mistake made by the landmark algorithm.

To train the MSFR we use face images which have been

aligned using the manually labelled landmarks given by the FRGC database. After face registration with MSFR we can use the manually labelled face images or the automatically register face images for training the face recognition algorithm. In this experiment we also use the manually labelled face images for training the face recognition. If the automatically register face images are used for training, the face recognition algorithm gives similar results.
To compare the MSFR with the landmark methods using an EER we also need to calculate the rejection score. This is done by introducing imposters to the face registration algorithm. Because the registration is based on the knowledge of an identity, we have to run the registration algorithm again giving a false identity to calculate a rejection rate. This is not necessary in the case of landmark finding because the registration is done without knowing the identity of the person. The matching score of authorized faces and score of the imposters allow us to calculate the FAR and FRR.

B. Results

We compare our algorithms on accuracy of the face alignment and performance in face recognition. To measure the performance in the accuracy of registration we use the RMS error. For the landmark based methods, the manually labelled landmarks of eyes, nose, mouth are compared with the final output of the landmark detectors. MSFR does not use landmarks, therefore we calculated the position of the eyes, nose, mouth after MSFR and compared these with the manually labelled landmarks. Because a straightforward comparison cannot be used due to the various scales of the face images, the exact calculation of the RMS error is given below:

Calculation of the RMS error:
1. Translate, scale and rotate the groundtruth data so that the eye landmarks are on a horizontal line at a 100-pixels distance form each other.
2. Align the shape found to the corresponding groundtruth shape.
3. Calculate the Euclidian distance between each landmark and its groundtruth equivalent.
4. Remove the bias caused by the different labelling policies in the databases, i.e. tip of the nose (BioID) versus a point between the nostrils (FRGC).
5. Calculate the RMS value of the remaining difference between the found shape and the groundtruth shape, which is now given as a percentage of the inter-eye distance.

The average RMS errors of the registration methods are shown in Table I. The matching score based face registration clearly outperforms the landmark based methods.

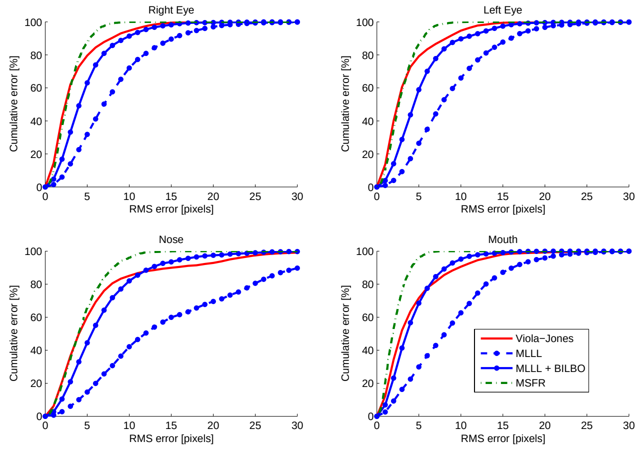

Fig. 3. The cumulative error distribution function of the RMS pixel error, comparing MSFR with landmark based methods

TABLE I
RELATIVE RMS ERROR RESULTS ON THE FRGC DATABASE, COMPARING MSFR WITH LANDMARK BASED METHODS

FRGC	right eye	left eye	nose	mouth
Viola-Jones	3.2	3.3	6.3	4.1
MLLL	6.7	7.2	13.0	7.3
MLLL+BILBO	4.2	4.6	5.8	3.7
MSFR	2.4	2.4	3.6	1.8

In Figure 3 the cumulative error distribution is shown as function of the RMS error. The graphs show the percentage of examples with a RMS error less than the RMS error on the horizontal axis. In these graphs, the distribution of the RMS error of the different methods also becomes visible. Figure 3 clearly shows that the MSFR performs better, especially when the RMS error becomes larger. The larger alignment errors seem to be the reason of the decrease in performance of the face recognition. In [2] it is shown that the EER increases strongly when the error in alignment grows.
Next, the performance in face verification of the MSFR is compared with the landmark based methods. The landmark based methods use the same face recognition algorithm as the matching score based face registration. We also use the same region of interest and face normalization. The only difference is that the face alignment is done by aligning the landmarks to a reference shape. For the landmark based methods, we calculate for each matching

score $(N-1)$ imposter scores, where N is the amount of individuals in the database. To reduce computation times, we calculate for the MSFR for each matching score only one random chosen impostor score. In theory, this does not effect the comparison and the results in EER significantly.

TABLE II
RESULTS OF THE FACE VERIFICATION EXPERIMENT, COMPARING MSFR WITH LANDMARK BASED METHODS AND MANUALLY LABELLED DATA

FRGC	EER [%]	std(EER)[%]
Ground truth data	0.45	0.03
Viola-Jones	4.9	0.1
MLLL	4.0	0.1
MLLL+BILBO	3.6	0.1
MSFR	0.95	0.2

The results are shown in table II, repeating the experiment 33 times for MSFR. Face verification based on groundtruth data gives the best result (EER = 0.45%), but the matching score based face registration works well with EER of 0.95%. This is a great improvement in EER compared to the landmark based method. We have manually inspected the face alignment with the lowest matching scores. Some low matching scores are caused by the search algorithm that makes a wrong step from which it impossible to converge to a good alignment.

At the moment, we only use this method for face verification problems, because the matching score based optimization is a time consuming process. It takes about 20-30 seconds to classify a face image on a Intel Pentium 2.80Ghz. For face identification we have to perform the search process for every person in the database, while for verification we perform it for one person. It is possible to optimize this further, using other search techniques, lower resolution face image, etc.

IV. DISCUSSION

In this paper we discussed a new approach for face registration. This approach uses the output of the face recognition classifier to evaluate the registration. We search for a optimal registration varying the face alignment parameters. Our new face registration approach performs better than the landmark based methods, in both RMS error and EER. The results obtained with this simple search algorithm indicate that this method has potential, but that improvements can be made in speed and accuracy.
The biggest disadvantage at the moment is the operating speed of the method. It takes about 20-30 seconds to classify a face image on a Intel Pentium 2.80Ghz, which makes it, at the moment, not usable for practical applications. We have to take into account that no research has been done yet in optimizing this process. One of the simple solutions that can be used is lowering the resolution of the face images which will improve the speed of the whole system. Other solutions can be optimized code, better search algorithms, faster classifier, etc.

REFERENCES

[1] T.P. Riopka and T. Boult, "The eyes have it," in *Proceedings of ACM SIGMM Multimedia Biometrics Methods and Applications Workshop.*, Berkeley, CA, 2003, pp. 9–16.

[2] G.M. Beumer, A.M.Bazen, and R.N.J. Veldhuis, "On the accuracy of EERs in face recognition and the importance of reliable registration.," in *SPS 2005*. IEEE Benelux/DSP Valley, April 2005.

[3] Paul A. Viola and Michael J. Jones, "Rapid object detection using a boosted cascade of simple features.," in *CVPR (1)*, 2001, pp. 511–518.

[4] R.N.J. Veldhuis, A.M. Bazen, W. Booij, and A.J. Hendrikse, "Hand-geometry recognition based on contour parameters," in *Proceedings of SPIE Biometric Technology for Human Identification II*, Orlando, FL, USA, March 2005, pp. 344–353.

[5] J.A.Nelder and R.Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, pp. 308–313, 1965.

[6] R.P. Brent, *Algorithms for Minimization without Derivatives*, Prentice-Hall, Englewood Cliffs, N.J., 1973.

[7] M.J.D Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Computer Journal*, vol. 7, pp. 155–162, 1964.

[8] M. Turk and A.P. Pentland, "Eigenfaces for recognition.," *Journal of cognative neuroscience*, pp. 71–86, 1991.

[9] Peter N. Belhumeur, Joao Hespanha, and David J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," in *ECCV 2*, 1996.

[10] NIST, "Frgc face db," `http://www.frvt.org/FRGC/`.

[11] G.M. Beumer, Q. Tao, A.M.Bazen, and R.N.J. Veldhuis, "A landmark paper in face recognition," *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, pp. 73–78, 2006.

[12] HumanScan, "Bioid face db," `http://www.humanscan.de/`.