



HAL
open science

Similarity-Based Matching for Face Authentication

Christophe Rosenberger, Luc Brun

► **To cite this version:**

Christophe Rosenberger, Luc Brun. Similarity-Based Matching for Face Authentication. International Conference on Pattern Recognition (ICPR'2008), Dec 2008, Tampa, Florida, United States. pp.1-4, 10.1109/ICPR.2008.4761860 . hal-00846773

HAL Id: hal-00846773

<https://hal.science/hal-00846773>

Submitted on 20 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Similarity-Based Matching for Face Authentication

Christophe Rosenberger Luc Brun

GREYC Laboratory
ENSICAEN - Université de Caen - CNRS
6 boulevard Maréchal Juin, 14000 Caen, France
{christophe.rosenberger, luc.brun}@greyc.ensicaen.fr

Abstract

We propose in this paper a face authentication method based on a similarity measure. The SIFT descriptor is used to define some interest keypoints characterized by an invariant parameter. A graph is then built where nodes correspond to these keypoints. We model the authentication problem as a graph matching process. Experimental results on the AR database show an EER equals to 12% with only one image used for the enrollment and with images simulating real conditions.

1. Introduction

Biometrics technologies are now present in our daily life environment, for instance: passport, access control and even when we are walking in the street. Among all the existing biometric modalities, face authentication/recognition remains a great challenge. Indeed, this biometric modality is a contactless technic which does not require the cooperation of the individuals.

Many researches have been done in the field of face recognition in order to improve the robustness of algorithms face to different variations which simulates real conditions [14, 8, 1]. Classical approaches are generally based on data reduction with different methods like PCA [13] or more sophisticated ones like fisherface [11]. Many methods use Gabor wavelets as pattern for face recognition [12]. Another solution consists in using local features computed on the original image [5]. The methods from the state of the art obtain good results but we can identify two important problems that limit the use of these methods for real applications. The first one concerns the number of required images for the enrolment phase. To obtain good results,

researchers generally test their methods on benchmark databases and use a percentage of the images sample of an individual for this step. In a real application, with a large database one cannot acquire as many images because it would cost too much (storage, personal needed). It is the same problem than for the fingerprint processing, only one image is generally captured. The second limitation is that many problems may occur like variations of illumination during the acquisition phase. Moreover, many aspects of an individual can change with time (hair, wearing glasses, etc.). In order to contribute to solve this problem, we propose to develop a robust face authentication algorithm which uses a single image for the enrolment step.

This paper is organized as follows. In section two, we detail the proposed method combining invariant and a similarity based matching method. Section three illustrates the efficiency of the proposed method on a well known benchmark namely the AR database. We conclude and give some perspectives of this work in section four.

2. Developed method

We assume in this paper that we process an image of an individual with any background information. The ability of the algorithm to realize a correct authentication in the other case is considered as a propriety of robustness. We propose to use the well known SIFT descriptor to characterize the individual's face.

The invariant descriptor developed in the SIFT algorithm [9], is applied locally on key-points and is based upon the image gradients in a local neighborhood. A recent work [6] shows that the SIFT algorithm is the most efficient local descriptor. The descriptor is created by sampling the magnitudes and orientations of the

image gradients in a neighborhood of each key-point and building smoothed orientation histograms that contain the important aspect of the neighborhood. Each local descriptor is composed on a 4x4 array (histogram) composed each with 8 orientation components. A 128-elements vector is then built for each key-point. We used in this article the implementation provided by Lowe [9].

Each individual is so described by a set of invariant features $Y(I) = \{k_i = (s_i, x_i, y_i)\}, i = 1 : N(I)$ where s_i is the SIFT invariant descriptor computed near the i th keypoint $k_i, (x_i, y_i)$ its position in the original image I and $N(I)$ the number of detected keypoints in the image I (see Figure 1). The recognition problem of an individual given one set $Y(I)$ is to find out the most similar individual in the database $\{Y(I_i), i = 1 : M\}$. Many methods have been proposed using this kind of information for the face recognition process. In [4], the recognition decision is realized by considering the majority vote as matching decision. In this case, each key-point is considered independently. The objective of this paper is to take into account the relationship between these keypoints in order to improve the efficiency of the face authentication process.

2.1 Pattern modeling

We choose to take into account the spatial relationship between keypoints by using graph representations. We connect keypoints by considering their "local neighborhoods". Each keypoint corresponds to a vertex in the neighborhood graph which connects "close" vertices. We considered the following types of graphs [3] which encode different notions of proximity between points:

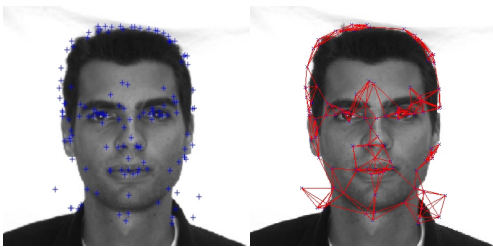


Figure 1. An example of keypoints detection with SIFT and the associated graph definition (MKNG)

- Epsilon-neighborhood graph (ENG): Two vertices are connected if the distance of the correspond-

ing keypoints is less than epsilon. The epsilon-neighborhood graph is used both as a weighted or unweighted graph. In the weighted case, the edge weights equal to the similarity values of the adjacent points;

- A completely connected graph (CCG). such a graph corresponds to an epsilon-neighborhood graph with epsilon greater than the maximal distance between keypoints.
- The symmetric k-nearest neighbor graph (SKNG): two vertices v_1, v_2 are connected if v_1 is among the k-nearest neighbors of v_2 or vice versa. We use a weighted symmetric k-nearest neighbor graph, that is all edges are weighted by the similarity of the adjacent points. The parameter of this graph is k, (number of neighbors considered).
- The mutual k-nearest neighbor graph (MKNG): two vertices v_1, v_2 are connected if v_1 is among the k-nearest neighbors of v_2 and vice versa. Similarly, we define a weighted mutual k-nearest neighbor graph, that is all edges are weighted by the similarity of the adjacent points.

We have so four graph representations for an individual's image using the keypoints associated to the SIFT invariant descriptor value.

2.2 Dissimilarity measure

The dissimilarity between two sets of points such as $Y(I_1)$ and $Y(I_2)$ is often modeled as a weighted bipartite graph matching problem which may be optimally solved using the Hungarian method [7]. However, such a method optimizes a global matching score and may thus be sensible to outliers. We thus rather use the following matching method based on a decision criterion first proposed by Lowe [9]:

Given two points $x \in Y(I_1)$ and $y \in Y(I_2)$, we say that x is associated to y iff :

$$d(x, y) = \min_{z \in Y(I_2)} d(x, z) \text{ and } d(x, y) \leq C d(x, y')$$

where C is an arbitrary threshold, $d(\cdot, \cdot)$ denotes the Euclidean distance between the SIFT descriptors and y' denotes any point of $Y(I_2)$ whose distance to x is minimal but greater than $d(x, y)$:

$$d(x, y') = \min_{z \in Y(I_2), d(x, z) > d(x, y)} d(x, z)$$

In other words, x is associated to y if y is the closest point from x in $Y(I_2)$ according to the Euclidean distance between SIFT descriptors and if the second smallest value of this distance $d(x, y')$ is significantly greater

than $d(x, y)$. The significance of the necessary gap between $d(x, y)$ and $d(x, y')$ is encoded by the constant C . In the same way, we say that $y \in Y(I_2)$ is associated to $x \in Y(I_1)$ iff x is the closest point from y among $Y(I_1)$ according to the Euclidean distance between SIFT descriptors and if the second smallest value of this distance $d(y, x')$ satisfies $d(y, x) < Cd(y, x')$. Then, we will say that x is matched to y iff x is associated to y and y is associated to x . If $Y(I_1)$ and $Y(I_2)$ do not define any matching point, we consider that both sets have an infinite distance. Otherwise, the set of matching points M_1 of Y_1 and M_2 of Y_2 may be computed sequentially under a mild assumption thanks to the following proposition:

Proposition 1 *If we cannot find two couples of points within $Y(I_1) \times Y(I_2)$ with a similar distance, the relationship “is matched to “ induces a bijective mapping ϕ from M_1 to M_2 .*

Proof:

Let us suppose that it exists $(y, y') \in M_2^2$ and $x \in M_1$, such that x is matched to y and y' . Such a relationship would implies that x is associated to both y and y' . The distances $d(x, y)$ and $d(x, y')$ should thus both correspond to the minimum of $\{d(z, x), z \in Y(I_2)\}$ and we should have $d(x, y) = d(x, y')$ which is refused by hypothesis. We can thus define an application ϕ from M_1 to M_2 such that $y = \phi(x)$ only if x is matched to y . This application is surjective, since $y \in M_2$ only if it exists at least one $x \in Y(I_1)$ which is matched to y . Finally, using the fact that all distances are different one can easily show that ϕ is injective. It defines thus a bijective mapping from M_1 to M_2 . ■

The above procedure either classify both sets $Y(I_1)$ and $Y(I_2)$ as dissimilar (case $M_1 = M_2 = \emptyset$) or provides a set of $N = |M_1| = |M_2|$ couples $(x, \phi(x))$ of matched points. Note that the application ϕ only used the information provided by the SIFT invariant descriptor. The second step has for objective to take into account the spatial relationship of the points among M_1 and M_2 . We thus use one of the four pattern representations described in the previous section by using the locations of each matched keypoint. Note that the matching between M_1 and M_2 being encoded by the application ϕ , one may encode the adjacency matrices of both graphs in such a way that a line or a column encoding the adjacency relationships of point x of the first matrix encodes the adjacency relationships of $\phi(x)$ in the second matrix. We thus do not need to compute a permutation matrix in order to find a matching between both graphs since we consider that this matching is already encoded by ϕ . Given both adjacency matrices, and the match-

ing relationship ϕ , the quality of the match is estimated using the squared Euclidean distance [2] which provides a global measure of the dissimilarity between both graphs. Such a distance may be interpreted as an upper bound of an edit distance between both graphs. Indeed given the mapping ϕ between the vertices of both graphs, one can construct a distance by associating a cost equal to the squared difference of edge’s weights for an edge substitution and the square of an edge’s weight for an edge deletion.

3. Experimental results

In order to facilitate the research in this field, many benchmark databases have been created to quantify the performances of a proposed algorithm and compare with the state of the art. The AR face database [10] is one of the most interesting one because it is composed of a large amount of individuals (120) and samples (26 per individual). Figure 2 shows 5 images of the same individual.



Figure 2. Some examples of images in the AR database

In this work, we use only one image for the enrolment step. This image is unaltered and the individual has no particular expression. We make for each individual $25+119*26=3119$ matching procedures. In total, the face authentication method we propose has been tested with 374280 face authentication tests. We tested the four graph representations. Two parameters have to be set. The first one $C \in [0, 1]$ is used in the first step of the keypoints matching. High values of C induce a restrictive selection of matching points. The second parameter is k for the three graph representations. It corresponds to the number of neighbors considered in the graph definition. The dissimilarity measure gives a coefficient that

can be interpreted as a confidence criterion on the fact that two faces belong to different individuals. Figure 3 gives the experimental results on the AR face database on a ROC curve. We can set a threshold to decide based on this dissimilarity measure if two images represent the same individual. By modifying this threshold, we can quantify the behavior of an algorithm. We use the FAR (False Acceptance Rate) and FRR (False Rejection Rate) criteria to analyse the behavior of an algorithm. The value of C has an important impact on the results as it permits to select reliable matched keypoints. As we can see in Figure 3, the representation based on a completely connected graph (CCG) gives the worst result. This means that the selection of edges is necessary. All other representations provide approximately similar behavior. The more selective graph representation, based on mutual k-nearest neighbor graph (MKNG) with $k=5$ and $C=0.8$ provides the lowest value of the EER (Equal Error Rate) which is equals to 12%.

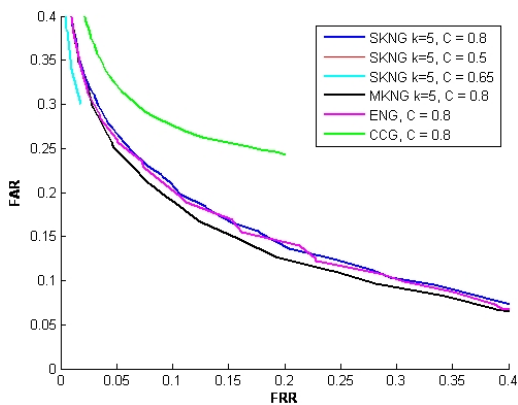


Figure 3. Verification results for the different representations

4. Conclusion and perspectives

We proposed in this paper a face authentication method based on a similarity measure. A graph representation based on SIFT invariant keypoints is used. The SIFT invariants provide a representation of some parts of the individual's face. The graph definition allows us to take into account their spatial relationships. We obtained on the AR benchmark containing the face of 120 individuals with different alterations an EER equals to 12%. The main advantage of the proposed method is its ability to obtain this performance with only one image used in the enrolment phase. We intend

to use this similarity measure as a kernel in a support vector machine in order to improve the efficiency of the enrolment step.

References

- [1] T.-W. Chen, S.-C. Hsu, and S.-Y. Chien. Automatic feature-based face scoring in surveillance systems. In *ISM '07: Proceedings of the Ninth IEEE International Symposium on Multimedia*, pages 139–146, Washington, DC, USA, 2007. IEEE Computer Society.
- [2] P. Dollar. Piotr's image & video toolbox for matlab, 2007. <http://vision.ucsd.edu/~pdollar/toolbox/doc/>.
- [3] M. Hein and U. von Luxburg. Graphdemo: a matlab gui to explore similarity graphs and their use in machine learning, 2007. <http://www.ml.uni-saarland.de/GraphDemo/GraphDemo.html>.
- [4] B. Hemery, C. Rosenberger, C. Toinard, and B. Emile. Comparative study of invariant descriptors for face recognition. *8th International IEEE Conference on Signal Processing (ICSP)*, 2006.
- [5] G. Heusch and S. Marcel. Face authentication with salient local features and static bayesian network. *Lecture Notes in Computer science*, page 878887, 2007.
- [6] C. S. K. Mikolajczyk. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10:1615–1630, 2005.
- [7] S. Kumar and D. G. M. Sallam. Matching point features under small nonrigid motion. *Pattern Recognition*, 34:2353–2365, 2001.
- [8] X. Liu and T. Chen. Pose-robust face recognition using geometry assisted probabilistic modeling. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 1:502–509, 2005.
- [9] D. Lowe. Distinctive image features from scale-invariant key-points. *International Journal for Computer Vision*, 2:91–110, 2004.
- [10] A. Martinez and R. Benavente. The ar face database. *CVC Tech. Report*, 24, 1998.
- [11] J. P. H. Peter N. Belhumeur and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, 19(7):711–720, July 1997.
- [12] L. Shen and L. Bai. A review on gabor wavelets for face recognition. *Pattern Analysis & Applications*, 9(2):273–292, 2006.
- [13] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 586–591, 1991.
- [14] S. Y. L. Z. Yuxiao Hu, Dalong Jiang and H. zhang. Automatic 3d reconstruction for face recognition. *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, pages 843–848, 2004.