MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11-13, 2002, Nara-ken New Public Hall, Nara, Japan

13—10

A Human Face Image Searching System using Sketch

C H Man and Pong C Yuen
Department of Computer Science
Hong Kong Baptist University
chman@comp.hkbu.edu.hk, pcyuen@comp.hkbu.edu.hk

Abstract

This paper addresses the problem in mug shot searching in suspect database. The objective is to improve the efficiency of the suspect identification process. The input is a sketch constructed by a witness. The output is the mug shot ranking for the suspect database. In the system, local and global feature measurements between sketch and mug shot faces are adopted. The local distance measurement compares facial primitives individually while the global distance measurement compares the geometric distances among feature points on the face. All mug shots are ranked based on the combined results. Experimental results show that 70% accuracy is achieved by considering the top 15% ranks.

1 Introduction

Sketch construction and mug shot searching are two common ways being used for suspect identification all over the world. A computerized system for the sketch construction [5] has been developed and employed and the suspect databases are usually in an electronic form. However, as far as we know, these two systems are not fully connected. Moreover, the size of a suspect database in police force is often in the order of tens of thousand. It is inefficient for human to search from such large among of mug shots. In this paper, we are going to combine sketch construction and mug shot searching so that it is more efficient for the suspect identification process. Our approach is to rank the mug shots in suspect database according to the similarity against the sketch so that the mug shot searching process does not need to cover the whole database.

The sketch construction software being used in Hong Kong police force [5] allows user to select from around 100 choices for each of facial primitive. The sketch constructed by that system is mainly formed by 6 facial primitives which include hair, eyebrows, eyes, nose, mouth and face outline. Primitives consist of only edges rather than skin texture. One example for each facial primitive is shown in figure 1(a). Figure 1(b) shows a sample of sketch constructed by the software.

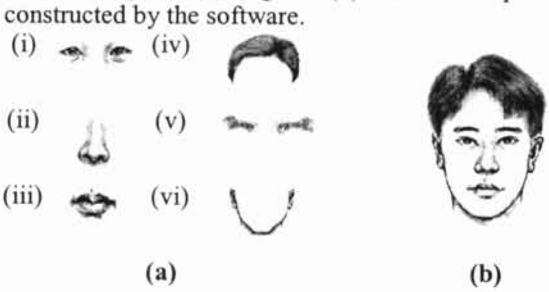


Figure 1. (a) The 6 facial primitives of a sketch (b) A sample sketch

Our problem are different from other research works on mug shot searching problem [1], [2], [3] and [4]. The sketches they considered have skin texture while those in our problem contain only edges. Also each of their sketches for testing is drawn by artist. However each of the sketches in our problem is composed by 6 primitives and there are around 100 choices for each of the primitives. The flexibility will limit the similarity. That is why the sketch often does not look like the actual mug shot. That is one of the major problems to be overcome.

The proposed system adopts local feature comparison and global feature comparison as distance measurements between faces. The idea of the former one is to compare facial primitives individually. Since all sketch primitives are chosen by a witness, the sketch information is very significant. Then the measurement can be broken down into comparison of eyebrows, eyes, nose, mouth and face outline. The latter one considers the interrelation amount the primitives, which are not concerned in local feature comparison. It measures the geometric differences among feature points on the face.

The paper is organized as follows. An overview of the proposed system is given in Section 2. Details on local feature comparison, global feature comparison and combined function are discussed in Sections 3 to 5 respectively. Sections 6 and 7 report the experimental results and conclusion.

2 An Overview of the Searching System

The input to our system is a sketch while the output is a list of mug shots ranked by the similarity with the incoming sketch. The block diagram is shown in figure 2. Instead of an image file, the input is a set of parameters, which are the codes of the primitives, primitive positions, horizontal and vertical size adjustments of primitives. When these parameters are passed into the system, both local and global distances between the incoming sketch and each mug shot in the suspect database are calculated. Mug shots are ranked according to the combined distance.

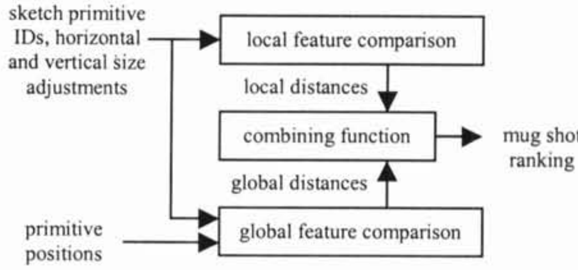


Figure 2. System overview

The local feature comparison measures the dissimilarity between primitives. Since all sketch primitives are strictly chosen by witness, they are significant. Then the measurement is broken down into comparison of eyebrows, eyes, nose, mouth and face outline. Hair is not considered since it is not a consistent feature. The global feature comparison measures the geometric facial distances. It measures the inter-primitive relation. Finally the local and global distances are combined. Distances of all mug shots in suspect database are ranked.

3 Local Feature Comparison

To perform local feature comparison, a face is divided into primitives including eyebrows, eyes, nose, mouth and face outline. Hair is not considered since it is not a consistent feature. Distance between each pair of primitives is computed individually. The overall local distance is the weighted sum of all primitive distances

All types of primitives are compared using the same strategy. For each individual primitive comparison, only the outline is used as edges provide rich information for face recognition [6]. Moreover, we adopt the point distribution model in active sharp model [8] to calculate the local distance. Therefore the local feature comparison consists of two stages, namely training stage and runtime stage. The block diagram of the local feature comparison is shown in Figure 3.

In the training stage, normalized sketch point profiles $\boldsymbol{P_S}=\{\boldsymbol{P^1_S}, \boldsymbol{P^2_S} \ldots\}$ and normalized mug shot point profiles $\boldsymbol{P_M}=\{\boldsymbol{P^1_M}, \boldsymbol{P^2_M}, \ldots\}$ are extracted from sketch database and mug shot database respectively. Eigenprimitives $\boldsymbol{E}$ are extracted from $\boldsymbol{P_S}$. The mug shot local feature vectors, $\boldsymbol{F_M}$, are then be found by projecting $\boldsymbol{P_M}$ to $\boldsymbol{E}$.

In the runtime stage, a sketch point profile $\boldsymbol{P_{income}}$ is constructed by $\boldsymbol{P_S}$ according to the incoming primitive ID, the horizontal and vertical adjustment information. The feature vector of incoming sketch primitive $\boldsymbol{F_{income}}$ is constructed by projecting $\boldsymbol{P_{income}}$ to $\boldsymbol{E}$. The projection process is same as that of mug shots. The primitive distance between sketch and mug shot ID i, $\boldsymbol{d^i_{pri}}$, is computed by the difference of $\boldsymbol{F_{income}}$ and $\boldsymbol{F^i_M}$.

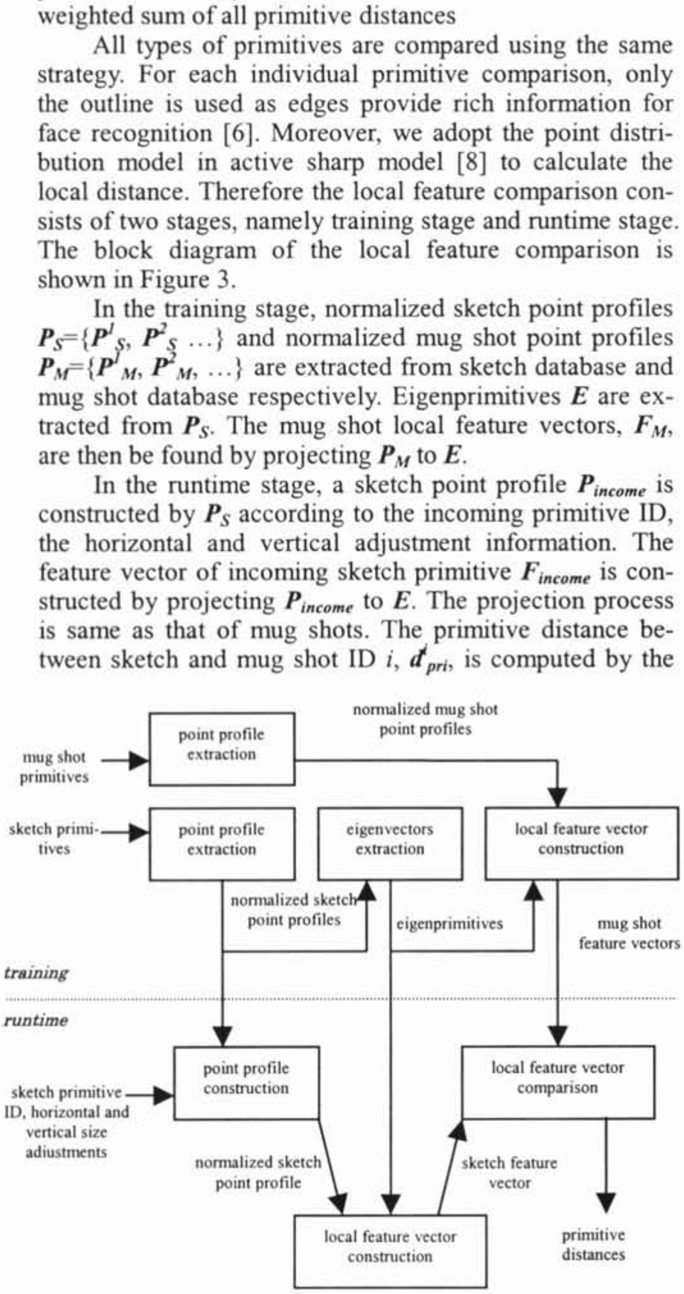


Figure 3. Facial primitive comparison

3.1 Point Profile Extraction

A point profile in our system is the feature points on the primitive outline. Feature points are first marked manually. Cubic spline is then used for extending the number of points. Figure 4 shows the samples of different kinds of point profiles altogether. $\boldsymbol{P}$ is defined as such point profile with translation, rotation and scaling normalization.

Figure 4. Samples of different kinds of point profiles

3.2 Eigenprimitive Calculation

Eigenprimitive E is found by principal component analysis on the normalized sketch point profiles $\boldsymbol{P_S} = \{\boldsymbol{P^1_S}, \ldots, \boldsymbol{P^N_S}\}$ where N is the total number of sketch primitives. The c variance of $\boldsymbol{P_S}$ is:
where μ is the mean of $\boldsymbol{P_S}$:

$$\Omega = \frac{1}{N}\sum_{n=1}^{N}[P_s^n-\mu][P_s^n-\mu]^T$$

$$\mu = \frac{1}{N}\sum_{n=1}^{N}P_s^n$$

Find k dominant eigenvectors $\boldsymbol{E} = \{\boldsymbol{E_1}, \ldots, \boldsymbol{E_k}\}$ such that:

$$\Omega E_j = \lambda_j E_j$$

where λ_j is the jth largest eigenvalue, $1 \leq k \leq N$.

3.3 Local Feature Vector Construction

Given a point profile $\boldsymbol{P}$, its local feature vector F is calculated by projecting $\boldsymbol{P}$ to the eigenprimitives $\boldsymbol{E}$:

$$F = E^T(P-\mu)$$

The size of $\boldsymbol{F}$ is much smaller than that of $\boldsymbol{P}$ so that it saves the storage and computational time for comparison.

3.4 Point Profile Constructions

To get the incoming sketch point profile in runtime, the $\boldsymbol{P^i_S}$ is first retrieved from database where i is the incoming sketch primitive ID. After the horizontal and vertical adjustments, the point profile needs to be normalized again for scaling.

3.5 Primitive Feature Vector Comparison

The distance between the local feature vector of incoming sketch F_{income} and that of the mug shot ID i, F^i_M, is calcu-

$$d^i = d(F_{income}, F^i_M) = \sqrt{\sum_j (F_{income}(j) - F^i_M(j))^2}$$

lated by Euclidian distance:

3.6 Local Distance Calculation

The overall local distance between incoming sketch and mug shot ID i is the weighted sum of all primitive distances:

$$D^i_{local} = w_{eyebrows} \times (d^i_{left\ eyebrow} + d^i_{right\ eyebrow})/2 + w_{eyes} \times (d^i_{left\ eye} + d^i_{right\ eye})/2 + w_{nose} \times d^i_{nose} + w_{mouth} \times d^i_{mouth} + w_{face} \times d^i_{face}$$

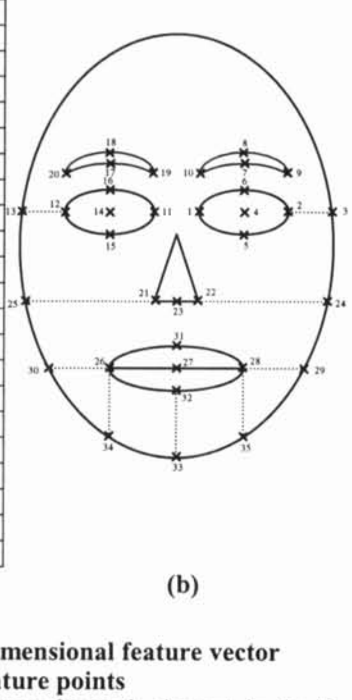
where w_{pri} is the weight for primitive ***pri***,
d_{pri} is the distance between primitive ***pri*** pair.

4 Global Feature Comparison

Global feature comparison measures the geometric difference between the incoming sketch and a mug shot. Each face is described by a vector $\boldsymbol{G}$, which is formed by distances among facial feature points [7]. In this section, (i,j) is defined as the geometric distance between point i and point j. Figure 5(a) shows the 25 geometric features used in our system. These features are computed from those facial feature points shown in figure 5(b). All the 25 distances are normalized by (4,14)+(14,27)+(4,27) to provide scale invariance. The scale normalized geometric feature vector of the mug shot ID i is defined as $\boldsymbol{G_i} = \{\boldsymbol{G_i^1}, \ldots, \boldsymbol{G_i^{25}}\}$.

Feature	Distance
G^1	0.5*((1,2)+(11,12))
G^2	0.5*((5,6)+(15,16))
G^3	(3,13)
G^4	(24,25)
G^5	(29,30)
G^6	(32,33)
G^7	(23,31)
G^8	(21,22)
G^9	0.5*((13,25)+(3,24))
G^{10}	0.5*((25,30)+(24,29))
G^{11}	0.5*((30,34)+(29,35))
G^{12}	0.5*((1,22)+(11,21))
G^{13}	(10,19)
G^{14}	0.5*((2,9)+(12,20))
G^{15}	0.5*((9,10)+(19,20))
G^{16}	0.5*((11,19)+(1,10))
G^{17}	0.5*((7,8)+(17,18))
G^{18}	0.5*((1,28)+(11,26))
G^{19}	0.5*((12,13)+(2,3))
G^{20}	(31,32)
G^{21}	0.5*((5,6)+(15,16))
G^{22}	0.5*((5,7)+(15,17))
G^{23}	0.5*((33,34)+(33,35))
G^{24}	0.5*((26,34)+(28,35))
G^{25}	0.5*((1,23)+(11,23))

(a)



(b)

Figure 5. (a) The 25-dimensional feature vector (b) Facial feature points

Since the scales are different for each element in the feature vector, the distance between $\boldsymbol{G_{sketch}}$ and $\boldsymbol{G_i}$ cannot be calculated by Euclidian distance. In our system, the global distance between the sketch and the mug shot ID i is calculated by:

$$D^i_{global} = \frac{1}{25}\sum_{k=1}^{25}\left(\frac{G^k_{sketch} - G^k_i}{\sigma_{G^k}}\right)^2$$

where σ_{G^k} is the standard deviation of $\boldsymbol{G^k}$ of all mug shots in database.

5 Combined Function

The local and global distances are normalized before combining. The scales for local feature distance and global feature distance are different. Therefore, they are normalized by the standard deviations themselves.

A weighted sum of local and global distances is proposed for combination. The combining function for local feature distance $\boldsymbol{D^i_{local}}$ and global feature distance $\boldsymbol{D^i_{global}}$ is given by:

$$D^i_{combine} = \omega_{local} \times \frac{D^i_{local}}{\sigma_{local}} + (1 - \omega_{local}) \times \frac{D^i_{global}}{\sigma_{global}}$$

where σ_{local} is the standard deviation of $\boldsymbol{D_{local}}$,
σ_{global} is the standard deviation of $\boldsymbol{D_{global}}$,
$0 \leq \omega_{local} \leq 1$.

$\boldsymbol{D^i_{combine}}$ is the dissimilarity between the sketch and the mug shot ID i. In other words, the smaller of $\boldsymbol{D^i_{combine}}$ is, the more similar to the sketch of the mug shot is. Mug shots are displayed in ascending order of $\boldsymbol{D^i_{combine}}$ so that the mug shot searching process does not need to cover the whole database.

6 Experimental Results

To evaluate the proposed system, sketches were constructed according to the mug shots in the testing database. Each constructed sketch is input into the system. The output shows the ranking of the target for that sketch. The higher the ranking indicates the better of the result is. 115 persons with 91 males and 24 females are used for testing. 92 sketches have been constructed for evaluations.

The results presented in this section are divided into 3 parts. The first part reports the result using the local feature comparison. The second part is for the global feature comparison. The third part is for the combined results.

For the local feature comparison, $w_{eyebrows}$, w_{eyes}, w_{nose}, w_{mouth} and w_{face} are set to 1 such that all primitives are equal weighting. Experimental results show that 33.7%, 50.0% and 95.7% of the testing cases are ranked within top 6 (top 5%), top 12 (top 10%) and top 58 (top 50%) respectively. Figure 6(a) shows the individual primitive comparison results of eyebrows, eyes, nose, mouth and face outline as well as the equal weighted sum overall result.

For the evaluation of global feature comparison, experimental results show that 42.4%, 55.4% and 88.0% of the testing cases are ranked within top 6 (top 5%), top 12 (top 10%) and top 58 (top 50%) respectively. Figure 6(b) shows results of all the 25 individual facial geometric features as well as the overall result.

For the evaluation of combined strategy, ω_{local} is set to 0.5 such that local and global distances are in equal weighting. Experimental results show that 47.8%, 60.9% and 96.7% of the testing cases are ranked within top 6 (top

5%), top 12 (top 10%) and top 58 (top 50%) respectively. Figure 6(c) plots the local, global and combined results for comparison. It can be seen that the combined results always give better accuracy than that of either local or global measurement.

7 Conclusion

We have proposed and developed an algorithm to measure the similarity of a sketch and a mug shot image. The proposed algorithm makes use of both local facial features and global facial features to draw the final conclusion. A small database with 115 persons is used to evaluate the flexibility of the proposed method. The experimental results show that the accuracy is close to 100% with top 50% ranks. That means, the searching time can be reduced by half with the use of our proposed method. Moreover, the accuracy is higher than 70% with top 15% ranks.

We have also found that one of the main errors is that the constructed sketch does not look like the actual mug shot. Therefore, the next step of this project is to develop an interactive searching system with human involvement such that the error due to the improper selection of local features in sketch construction can be minimized.

Moreover, the current results have not yet taken the advantages of weight adjustments. With the human involvement, the user/witness can adjust the weight of based on his/her confidence on particular features. In turn, the results can be further improved.

Acknowledgments

The authors would like to thank the Chinese University of Hong Kong for the support of the software AI-CAMS-FIT [5]. This research is supported by Hong Kong Baptist University.

References

[1] R. Uhl and Lobo N. da Vitoria, "A framework for recog-niz-ing a facial image from a police sketch", *IEEE Conference on Computer Vision and Pattern Recognition*, San Fran-cisco, June 1996.

[2] W. Konen, "Comparing Facial Line Drawings with Gray-Level Images: A Case Study on PHANTOMAS", *International Conference on Artificial Neural Networks*, pp. 727-734, 1996.

[3] E. Baker, "The Mug-Shot Search Problem: A Study of the Eigenface Metric, Search Strategies, and Interfaces in a System for Searching Facial Image Data", *Harvard Univer-sity Division of Engineering and Applied Sciences Technical Report TR-16-98*, December 1998.

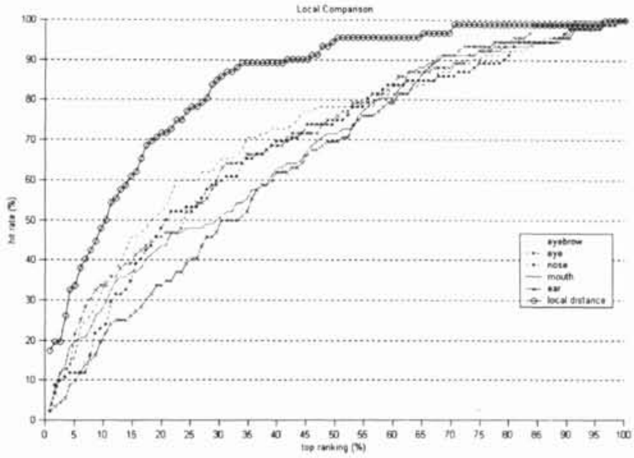
[4] E. Baker and M. Seltzer, "The Mug-Shot Search Problem", *Vision Interface `98 Proceedings*, pages 421-430. June 1998.

[5] J.W. Brahan, K.P. Lam, H. Chan and W. Leung, "AICAMS - Artificial Intelligence Crime Analysis and Management System". *Applications and Innovations in Expert Systems V, Journal of Knowledge Based System*. 143-153, 1997.

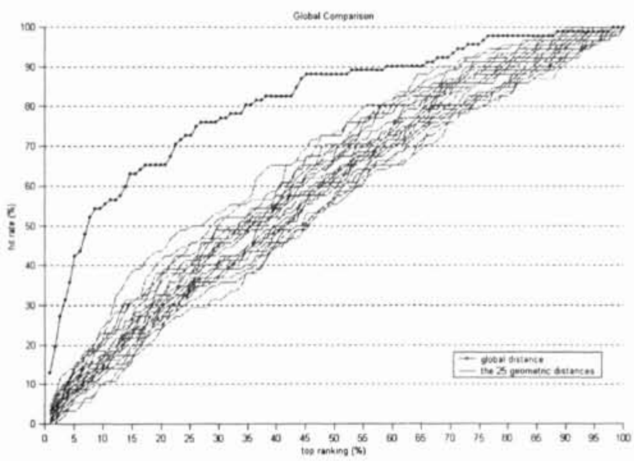
[6] Yongsheng Gao, Maylor K.H.Leung: "Face Recognition Using Line Edge Map" *IEEE transactions on pattern analysis and machine intelligence*, Vol.24, No. 6, June 2002.

[7] Ingemar J. Cox, Joumana Ghosn, Peter N. Yianilos: "Feature-Based Face Recognition Using Mixture-Distance". *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1995.

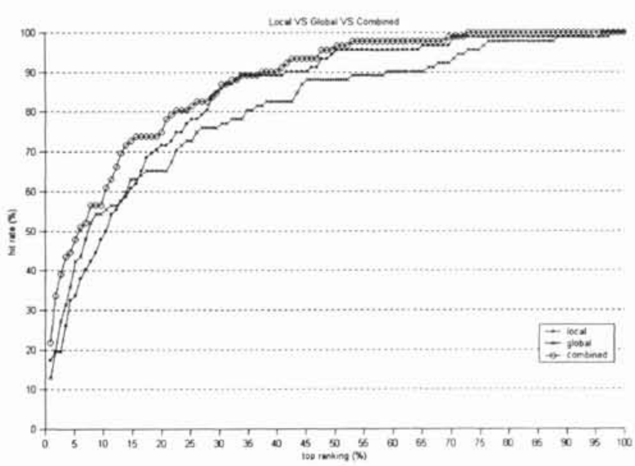
[8] T.F.Cootes and C.J.Taylor, "Statistical Models of Appearance for Computer Vision" *Internal report 2001*. (Available in http://www.isbe.man.ac.uk.)



(a)



(b)



(c)

Figure 6. (a) Local feature comparison result (all w_{pri}=1)
(b) Global feature comparison result
(c) Combined result (ω_{local}=0.5)