MULTI-VIEW COMBINED VISUAL SIMILARITY RETRIEVAL FOR FACE DATABASE

Yongsheng Gao and Maylor K. H. Leung
School of Computer Engineering
Nanyang Technological University, Singapore 639798
asysgao@ntu.edu.sg, asmkleung@ntu.edu.sg

Abstract   This work proposed a new multi-view combined visual similarity retrieval approach for face database. The search and retrieval employ a novel facial corner coding and matching approach, which delivers fast searching speed for database retrieval. The experimental results demonstrate that the facial corners can be a useful identity feature for human face database retrieval. The profile view of a face provides a complementary structure of the face that is not seen in the frontal view. The proposed multi-view modified Hausdorff distance combines the facial corners of multiple views and achieved an improved accuracy. This technique is not limited to two view combined retrieval but a generic approach for multiple view combined similarity retrieval.

Keywords: Visual similarity, image retrieval, multi-view fusion.

1.   INTRODUCTION

Most of the current studies of human face database retrieval focus on frontal faces [1-14]. Little attention has been paid to multiple view combined face database retrievals. In mug shots face searching and retrieval for law enforcement, the face profile view is another important source of identity features for the retrieval of faces. The profile view of a face provides a complementary structure of the face that is not seen in the frontal view. The retrieval system combining both frontal and profile views of faces can improve the system accuracy. And it would be more foolproof because it is difficult to fool the profile face identification by a mask or make-up.

This work proposed a new multi-view combined visual similarity retrieval approach for face database. The search and retrieval employ a novel facial corner coding and matching approach, which delivers fast searching speed for database retrieval. Unlike other face profile matching techniques that make use of the profile silhouette only, this method integrates the outline and interior facial parts in the profile matching. The proposed multi-view modified Hausdorff distance fuses the frontal and profile view faces to achieve an improved system performance. The technique is not limited to two view combined retrieval but a generic approach for multiple view combined similarity retrieval. This study is believed to be the first piece of work for multi-view combined visual similarity retrieval of face database and investigating the facial corners for human face matching.

In the following, a facial corner feature extraction and matching method is presented in Section 2. In Section 3, the combined image similarity retrieval of multiple views using corners is presented. A multi-view modified Hausdorff distance is proposed and described in detail. Encouraging experimental results are reported in Sections 4. Finally, the paper is concluded in Section 5.

2. FACIAL CORNER EXTRACTION AND MATCHING

Corners and vertices represent a class of information in image analysis. These features can be used to identify rigid objects in the scene, for example, for stereoscopic matching, geometrical calibration, due to their accuracy and less ambiguous locations. However, for non-rigid objects with smooth surfaces, such as human faces, it remains an unexplored question whether corners or dominant points can be used for identification and how well they perform. This work is believed the first attempt to apply corner coding and matching on human face retrieval.

2.1 Facial Corner Detection

Edges are the most fundamental features of objects in the 3D world. The edges in an image reflect large local intensity changes that are caused by the geometrical structure of the object, the characteristics of surface reflectance of the object, and viewing direction. Edge detection is an important technique in computer vision and pattern recognition systems since the edges usually correspond to some important properties of objects, such as object shape and boundary.

Edge extraction has been studied most extensively, and many reliable algorithms have been proposed and implemented [15-26]. A popular scheme for edge detection is to find the extrema in the blurred spatial derivates of the intensity images. The idea of blurred Laplacian filtering is extensively elaborated by Marr *et al.* [15, 16]. The Canny edge detection algorithm [20] is considered a "standard method" used by many researchers. Based on the assumption that intensity edges are ideal step edges corrupted by Gaussian noise, the Canny edge detector uses linear filtering with a Gaussian kernel to smooth noise and then computes the edge strength and direction for each pixel in the smoothed image. The edge detection algorithms proposed by Chen and Yang [21], Nevatia *et al.* [22], Nalwa *et al.* [23], Iverson *et al.* [24], Bergholm [25] and Rothwell *et al.* [26] are only a few examples from a large volume of these algorithms that had been presented to the vision community. The continuing development of edge detectors is producing increasingly complex edge detection algorithms. However, the increased sophistication of newer algorithms is not producing a commensurate improvement in performance [17] since the early algorithms of Roberts [18] and Sobel [19]. M. D. Heath *et al.* [27, 28] had investigated the performance of different edge detectors. They compared the edge detectors based on experimental psychology and statistics, in which humans rated the outputs of low level vision algorithms. One of their clear results is that "no one single edge detector was best overall; for any given image it is difficult to predict which edge detector will be best" [28].

The proposed face retrieval method using facial corners does not rely on any specific edge detector. It is a general method that any edge detector can be used. In this study, an edge detector based on the algorithm of Nevatia [22] is used followed by a thinning process to generate one pixel wide edge curves.

To generate facial corners, the dynamic two strip algorithm (Dyn2S) [30] is utilized to detect dominant points on the facial edge curves. In Dyn2S algorithm, a strip is fitted to the left and right of each point on the curve, and the points inside each strip are approximated as a straight line. The orientation and width of the strip are adjusted automatically. Longer and narrower strips are favored. A measure of "merit" based on the strips lengths and widths, and the angle between the strips is calculated. This merit provides an objective evaluation of the prominent strength of each point. The results of applying these processes on the frontal and profile faces are illustrated in Figure 1.

Figure 1 Facial corners extracted from frontal and side view faces.

2.2   Corner Matching

The Hausdorff distance is a shape comparison metric based on binary images. It is a distance defined between two point sets. Unlike most shape comparison methods that build a one-to-one correspondence between a model and a query image, the Hausdorff distance can be calculated without explicit point correspondence. Huttenlocher *et al.* [31] argued that the Hausdorff distance for binary image matching is more tolerant to perturbations in the locations of points than binary correlation techniques since it measures proximity rather than exact superposition. Takács [29] used Hausdorff distance to measure the similarity of two binary frontal face images, and achieved 92% accuracy in identifying the input. He argued that the process of face recognition might start at a much earlier stage and edge images could be used for fast screening of faces without the involvement of high-level cognitive functions. Here, we extend the Hausdorff distance measure from binary image matching to facial corner point matching in order to further reduce the computational time to cater for the high speed requirement in image database retrieval.

Given two finite facial corner point sets $M = \{ m_1, m_2, ..., m_k \}$ (representing a model in the database) and $T = \{ t_1, t_2, ..., t_n \}$ (representing a query image), the Hausdorff distance is defined as

$$H(M,T) = \max(h(M,T), h(T,M)) \tag{1}$$

where

$$h(M,T) = \max_{m_i \in M} \min_{t_j \in T} \left\| m_i - t_j \right\| \tag{2}$$

and $\left\| m_i - t_j \right\|$ is Euclidean norm between the points m_i and t_j. The function $h(M,T)$ is called the directed Hausdorff distance from M to T. It identifies the point $m_i \in M$ that is the farthest from any point of T and measures the distance from m_i to its nearest neighbor in T. The Hausdorff distance $H(M,T)$ is the maximum of $h(M,T)$ and $h(T,M)$. Thus, it measures the degree of mismatch between two sets by measuring the distance of the point of M that is farthest from any point of T and vise versa.

The Hausdorff distance defined as (1) and (2) is very sensitive to outlier points. A few outlier points, even only a single one, can perturb the distance greatly, though the two objects might be very similar. Dubuisson and Jain [32] investigated 24 forms of different Hausdorff distance measures and indicated that a Modified Hausdorff Distance (MHD) measure has the best performance. The directed MHD is defined as

$$h(M,T) = \frac{1}{p} \sum_{m_i \in M} \min_{t_j \in T} \left\| m_i - t_j \right\| \tag{3}$$

where p is the number of points in M. The definition of the undirected MHD is the same as (1). The MHD is robust to outlier points that might result from segmentation errors. In this study, a new multi-view modified Hausdorff distance (MVMHD) is proposed for the combined image similarity retrieval of multiple views for human face databases.

3.   COMBINATION OF MULTIPLE VIEWS

Let $M = \{M^1, M^2, \cdots, M^n\}$ be an object model in the database, which is composed of n views from n specific viewing angles, and $T = \{T^1, T^2, \cdots, T^n\}$ be a query object composed of n views of the same viewing angles as in the model. Define $M^1 = \{m_1^1, m_2^1, \cdots, m_{p_1}^1\}$, $M^2 = \{m_1^2, m_2^2, \cdots, m_{p_2}^2\}, \ldots, M^n = \{m_1^n, m_2^n, \cdots, m_{p_n}^n\}$ to be n point sets representing n model views of object M in the database and $T^1 = \{t_1^1, t_2^1, \cdots, t_{q_1}^1\}$, $T^2 = \{t_1^2, t_2^2, \cdots, t_{q_2}^2\}, \ldots, T^n = \{t_1^n, t_2^n, \cdots, t_{q_n}^n\}$ to be n point sets representing n views of a query object T from the same viewing angles as in the model M. $p_1, p_2, \ldots, p_n$ and $q_1, q_2, \ldots, q_n$ are point numbers in model and query views, respectively. A multi-view modified Hausdorff distance (MVMHD) is defined as the distance between the model image set M and the query image set T:

$$H_{MVMHD}(M,T) = \max(h_{MVMHD}(M,T), h_{MVMHD}(T,M)) \qquad (4)$$

The directed MVMHDs from M to T and from T to M are defined as (5) and (6).

$$h_{MVMHD}(M,T) = \frac{1}{\sum\limits_{k=1}^{n} N_{M^k}} \sum\limits_{k=1}^{n} \sum\limits_{m_i^k \in M^k} \min_{t_j^k \in T^k} \left\| m_i^k - t_j^k \right\| \qquad (5)$$

$$h_{MVMHD}(T,M) = \frac{1}{\sum\limits_{k=1}^{n} N_{T^k}} \sum\limits_{k=1}^{n} \sum\limits_{t_i^k \in T^k} \min_{m_j^k \in M^k} \left\| t_i^k - m_j^k \right\| \qquad (6)$$

where the superscript k stands for the kth view of the model set M or the query set T, and n is the number of views in representing an object. For a point $m_i^k \in M^k$ in kth view of the model M, $h_{MVMHD}(M,T)$ identifies its nearest neighbor point in the kth view T^k of T and measures the distance from m_i^k to the identified point. All the points in each view of M contribute equally to the calculation of the directed MVMHD.
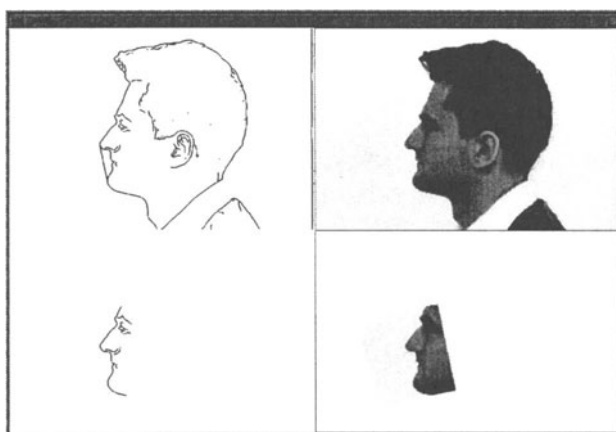
4.   EXPERIMENTAL RESULTS

A face database [33] of 30 persons with 2 frontal and 2 profile views per person from the University of Bern was used to test the capability of the proposed MVMHD approach on multi-view combined face image retrieval. Each image has the size of 512×342 pixels with 256 gray levels, and with variations of the head position, size and contrast. One profile and one frontal face images of all individuals were used as queries while the others were used as models. Though there are 30 pairs of models and queries, one can have 60 retrieval experiments (one query is matched against 30 models in each retrieval experiment) if the roles of models and queries are interchanged. Sample sets of the frontal and profile faces are shown in Figure 2.

Figure 2 Example sets of frontal and profile view faces in the database.

In order to remove the influences of hair, neck, and other foreign objects such as the collar of shirts, the face retrieval process should be restricted to the personal identity area with eyebrows, eyes, nose, mouth, and chin only. For this purpose, an automatic profile location algorithm is employed to detect the nose tip and chin points. The distance between these two points was used as a reference to normalize image size, align face position and crop facial area. An example of the profile cropping process is illustrated in Figure 3. For the frontal faces, the eye locations are detected for normalization, alignment and cropping.

Figure 3 Illustration of profile face cropping.

A prototypical multi-view combined face retrieval system using the proposed MVMHD method was implemented (Figure 4). The experimental results of combined similarity retrieval of frontal and profile views are summarized in Table 1 together with the retrieval accuracy using single frontal view and profile view. The proposed facial corner matching can correctly retrieve 96.67% of frontal view faces and 90% of profile view faces. The combined retrieval on facial corner features using the multi-view modified Hausdorff distance improved the system performance up to 100%. These results show that the facial corners can be a useful identity feature for human face retrieval, which delivers fast searching speed. And the combined retrieval of multiple views using the proposed MVMHD demonstrated a better system performance.
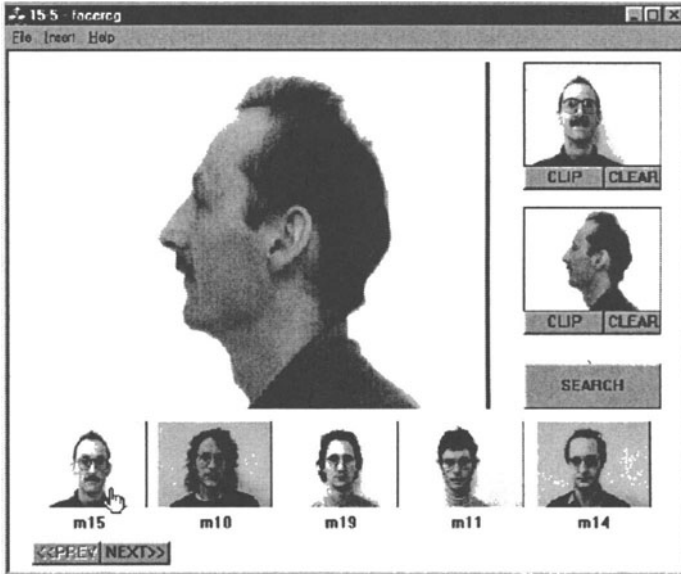
Figure 4 An example of multi-view combined face database search and retrieval result. The frontal and profile views in the right windows are query views. The images at the bottom are five best matched individuals (the leftmost is the best match) retrieved from the model database.

Table 1 The retrieval results of single view (frontal and side view) and combined views.

	Single view		Combined views
	Frontal view	Side view	
Retrieval accuracy	96.67%	90.00%	100%
Average distance of all matches	5.87	9.16	6.57
Average distance of correct matches	3.18	3.99	3.34

5.   CONCLUSION

So far, no single face matching method can guarantee 100% accuracy. However, a face retrieval system that integrates the matching of frontal and profile faces could achieve better accuracy. This study is believed to be the first piece of work for multi-view combined visual similarity retrieval of face

database and investigating the facial corners for human face matching. The experimental results demonstrate that the facial corners can be a useful identity feature for human face database retrieval, which delivers fast searching speed. The profile view of a face provides a complementary structure of the face that is not seen in the frontal view. The proposed multi-view modified Hausdorff distance combines the facial corners of multiple views and achieved an improved accuracy. This system would be more foolproof because it is difficult to mislead the profile face identification by a mask or make-up.

REFERENCES

[1]   M. Turk and A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, Vol. 3, pp. 71-86, 1991.

[2]   A. Pentland, B. Moghaddam and T. Starner, "View-based and Modular Eigenspaces for Face Recognition", Proc. IEEE Computer Soc. Conf. On Computer Vision and Pattern Recognition, pp. 84-91, 1994.

[3]   S. Lawrence, C. L. Giles, A. C. Tsoi and A. D. Back, "Face Recognition: A Convolutional Neural-Network Approach", IEEE Transaction on Neural Networks, Vol. 8, pp. 98-113, 1997.

[4]   S. H. Lin, S. Y. Kung and L. J. Lin, "Face Recognition/Detection by Probabilistic Decision-Based Neural Network", IEEE Transaction on Neural Networks, Vol. 8, pp. 114-132, 1997.

[5]   S. Y. Kung and J. S. Taur, "Decision-based neural networks with signal/image classification applications", IEEE Transaction on Neural Networks, Vol. 6, pp. 170-181, 1995.

[6]   M. Lades, J. C. Vorbrüggen,, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Würtz and M. Konen, "Distortion Invariant Object Recognition in the Dynamic Link Architecture", IEEE Transaction on Computers, Vol. 42, pp. 300-311, 1993.

[7]   L. Wiskott and C. von der Malsburg, "Recognizing Faces by Dynamic Link Matching", Neuroimage 4, pp. S14-S18, 1996.

[8]   F. Samaria and F. Fallside, "Face Identification and Feature Extraction using hidden markov models", G. Vernazza, editor, Image Processing: Theory and Application, Elsevier, 1993.

[9]   Y. Kaya and K. Kobayashi, "A Basic Study on Human Face Recognition", Frontiers of Pattern Recognition, S. Watanabe, Ed. pp. 265, 1972.

[10] R. Bruneli and T. Poggio, "Face Recognition: Features versus Templates", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, pp. 1042-1052, 1993.

[11] I. J. Cox, J. Ghosn and P. N. Yianios, "Feature-based face recognition using mixture-distance", Computer Vision and Pattern Recognition, Piscataway, NJ: IEEE Press, 1996.

[12] B. S. Manjunath, R. Chellappa and C. von der Malsburg, "A feature based approach to face recognition", Proc. IEEE Computer soc. Conf. on Computer Vision and Patt. Recog., pp. 373-378, 1992.

[13] M. Bichsel, "Strategies of Robust Object Recognition for Identification of Human Faces", Ph.D. thesis, Eidgenossischen Technischen Hochschule, Zurich, 1991.

[14] C. L. Kotropoulos, A. Tefas and I. Pitas, "Frontal Face Authentication Using Discriminating Grids with Morphological Feature Vectors", IEEE Transactions on Multimedia, Vol. 2, pp. 14-26, 2000.

[15] D. Marr, "Vision: A Computational Investigation into the Human Representation and Processing of Visual Information" W. H. Freeman, San Francisco, CA., 1981.

[16] D. Marr and E. C. Hildreth, "Theory of Edge Detection", Proc. Royal Soc. of London, Vol. B207, pp. 187-217, 1980.

[17] K. L. Boyer and S. Sarkar, "Assessing the State of the Art in Edge Detection:1992" SPIE, Application of Artificial Intelligence X: Machine Vision and Robotics, Vol. 1708, pp. 353-362, 1992.

[18] L. G. Roberts, "Machine Perception of Three-Dimensional Solids", Optical and Electro-Optical Information Processing, J. T. Tippett *et al.* eds., pp. 159-197, Cambridge Mass.: MIT Press, 1965.

[19] I. E. Sobel, "Cameral Model and Machine Perception", PhD thesis, Stanford Univ., 1970.

[20] J. Canny, "A Computational Approach to Edge Detection", IEEE Transaction on PAMI, Vol. 8, pp. 679-698, 1986.

[21] G. Chen, and Y. H. Yang, "Edge Detection by Regularized Cubic B-Spline Fitting", IEEE Transactions on Systems, Man and Cybernetics, Vol. 25, pp. 636–643, 1995.

[22] R. Nevatia and K. R. Babu, "Linear Feature Extraction and Description", Computer Graphics and Image Processing, Vol. 13, pp. 257-269, 1980.

[23] V. S. Nalwa and T. O. Binford, "On Detecting Edges", IEEE Transaction on PAMI, Vol. 8, pp. 699-714, 1986.

[24] L. A. Iverson and S. W. Zucker, "Logical/Linear Operators for Image Curves", IEEE Transactions on PAMI, Vol. 17, pp. 982-996, 1995.

[25] F. Bergholm, "Edge Focusing", IEEE Transactions on PAMI, Vol. 9, pp. 726-741, 1987.

[26] C. A. Rothwell, J. L. Mundy, W. Hoffman and V. D. Nguyen, "Driving Vision by Topology", Int'l Symp. Computer Vision, pp. 395-400, Coral Gables, Fla., Nov. 1995.

[27] M. D. Heath, S. Sarkar, T. Sanocki and K. W. Bowyer, "A Robust Visual Method for Assessing the Relative Performance of Edge-Detection Algorithms", IEEE Transactions on PAMI, Vol. 19, pp. 1338-1359, 1997.

[28] M. D. Heath, S. Sarkar, T. Sanocki and K. W. Bowyer, "Comparison of Edge Detectors: A Methodology and Initial Study", Computer Vision and Image Understanding, Vol. 69, pp. 38-54, 1998.

[29] B. Takács, "Comparing Face Images using the Modified Hausdorff Distance", Pattern Recognition, Vol. 31, pp. 1873-1881, 1998.

[30] M. K. H. Leung and Y. H. Yang, "Dynamic two-strip algorithm in curve fitting", Pattern Recognition, Vol. 23, pp. 69-79, 1990.

[31] D. P. Huttenlocher, G. A. Klanderman and W. J. Rucklidge, "Comparing Images Using the Hausdorff Distance", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, pp. 850-863, 1993.

[32] M. P. Dubuisson and A. K. Jain, "A Modified Hausdorff Distance for Object Matching", Proceedings 12th International Conference on Pattern Recognition, Jerusalem, Israel, pp. 566-568, 1994.

[33] Bern University face database. ftp://iamftp.unibe.ch/pub/Images/FaceImages/

BIOGRAPHIES

Yongsheng Gao received the BSc and MSc degrees in Electronic Engineering from Zhejiang University, China, in 1985 and 1988 respectively, and his PhD degree in Computer Engineering from Nanyang Technological University, Singapore. Currently, He is an assistant professor with Nanyang Technological University. His research interests include computer vision, pattern recognition and face recognition.

Maylor K.H. Leung received the BSc degree in physics from the National Taiwan University in 1979, and the BSc, MSc and PhD degrees in computer science from the University of Saskatchewan, Canada, in 1983, 1985 and 1992 respectively. Currently, Dr. Leung is an associate professor with Nanyang Technological University, Singapore. His research interests include structural pattern processing, face recognition, motion analysis and string processing.