Feature-Level Fusion in Personal Identification

Yongsheng Gao, Member, IEEE, Michael Maggs

School of Microelectronic Engineering, Griffith University, Brisbane, Australia
yongsheng.gao@griffith.edu.au

Abstract

The existing studies of multi-modal and multi-view personal identification focused on combining the outputs of multiple classifiers at the decision level. In this study, we investigated the fusion at the feature level to combine multiple views and modals in personal identification. A new similarity measure is proposed, which integrates multiple 2-D view features representing a visual identity of a 3-D object seen from different viewpoints and from different sensors. The robustness to non-rigid distortions is achieved by the proximity correspondence manner in the similarity computation. The feasibility and capability of the proposed technique for personal identification were evaluated on multiple view human faces and palmprints. This research demonstrates that the feature-level fusion provides a new way to combine multiple modals and views for personal identification.

1. Introduction

Face, fingerprint, palm-print and voice recognition techniques have been successfully used in personal identification. However, A single biometric feature sometimes fails to be exact enough for verifying the identity of a person. Personal identification systems based on a single view from a single sensor have inherent weakness. They generally encounter uncertainty problem when the number of people increases. In face recognition, for example, a frontal view face image may not provide sufficient 3D information to distinguish people who look like each other, especially when the observation is ambiguous. A different view, such as a profile view, can provide a complementary structure of the face to make up for observations that are inherently incomplete or ambiguous in frontal view. Moreover, a face recognition system using multiple views is more foolproof because it is much more difficult to fool a face recognition system combining multiple views from different viewpoints than a conventional frontal face recognition system by a mask or a picture.

Traditionally, the identification decision is drawn on a single modality of biometric information. Recent studies [1], [2] show that the use of multiple modalities to solve automatic personal identification problem leads to tangible benefits in terms of accuracy and robustness. Multi-modal systems can achieve better performance than single modal systems because different modals provide complementary information. For instance, it is very likely that two or more people, among a large population, have similar faces, in particular, under appearance changes due to expression, lighting condition and subject's pose changes, but their fingerprints or palm-prints are different enough to distinguish them, or vice versa.

Brunelli and Falavigna [2] combined acoustic and visual cues for personal identification. Two classifiers based on acoustic features and three based on visual features provided data for an integration module. They integrated multiple classifiers at a hybrid rank/ measurement level using HyperBF networks. Kittler *et al.* [3] proposed a common theoretical framework for classifier combination and fused [4] multiple instances of biometric data to improve the performance of a person identity verification system. The fusion problem is formulated in the framework of the Bayesian estimation theory. Ben-Yacoub *et al.* [5] investigated the support vector machine, the multiplayer perception, the C4.5 decision tree, the Fisher's linear discriminant and the Bayesian classifier to fuse face and speech data for personal identity verification. Dieckmann *et al.* [6] included the lip motion feature in the face and speech identification system to improve the system reliability. Chatzis *et al.* [7] used fuzzy clustering algorithms for decision–level data fusion in a person authentication system. Different modalities of image and speech were combined by fuzzy *k*-means and fuzzy vector quantization algorithms, and median radial basis function network. Lin and Jain [1] integrated faces and fingerprints for personal identification. A decision fusion scheme integrates multiple cues with different confidence measures by assuming that the similarity values between faces are statistically independent of the similarity values between fingerprints. Encouraging results were obtained by their integrated system in the identification mode. The current studies of multi-modal and multi-view personal identification focused on combining multiple classifiers at the decision level. Jain and Dorai [8] suggested that

ideally, one would like to combine the features derived from the individual models, so that a more accurate classification can be obtained using the pooled features. However, the feature-level fusion still remains an unexplored area for personal identification.

To the best of our knowledge, there is hardly reported research work on personal identification using feature-level fusion. In this study, we attempt to explore the idea of combining multiple views and multiple modals at the feature-level for personal identification. A novel feature-level integration method is proposed to combine a pooled line edge map features from multiple views and multiple modals. Line features derived from different views and modals are combined at the feature level to form a single integrated similarity measure. The multi-view line segment Hausdorff distance (MVLHD) is proposed to integrate these homogeneous lines from different sensors, which delivers fast matching speed.

The paper is organized as follows. In Section 2, we introduce the concept of describing an object using line edge maps. In Section 3, a novel multi-view line segment Hausdorff distance is proposed to combine multiple views and multiple modals at the feature level to form a single integrated similarity measure. The approach proposed in Sections 2 and 3 are examined experimentally for personal identification using multiple views of human faces and palmprints in Section 4. The system performance is also tested on multiple views of faces under appearance changes. Finally, conclusions are drawn in Section 5.

2. Line feature description

Edges are the most fundamental features of objects in the 3D world. The edges in an image reflect substantial 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 maps are widely used in various object recognition techniques [9], [10]. On the other hand, cognitive psychological studies [11],[12] also indicated that human beings recognize line drawings as quickly and almost as accurately as gray level pictures. These results show that edges of objects can be an important feature for object recognition. However, pixel-wise edge maps describe spatial information of edge curves but lack discriminative capability of representing structural information. Line features [13],[14],[15] are thus proposed for describing and matching objects because a line has more information than a point [13],[14]. In biometric systems, the line features had been successfully used in lighting insensitive face recognition [15] and palmprint verification [14], [16]. In this research, we employ the line edge maps (LEMs) [15] of an object from different views and sensors as a homogeneous feature descriptor. The LEM provides both spatial information of a template matching map and local structural information of a geometrical feature matching representation. It is an object descriptor in accordance with the argument of Bruneli and Poggio [17] that successful object recognition approaches might need to combine aspects of feature-based approaches with template matching method. A multi-view line segment Hausdorff distance is proposed to combine the above lines at the feature level to form a single integrated similarity measure.

In this study, an edge detector based on the algorithm of [18] is used followed by a thinning process to generate one-pixel wide edge curves. To generate the Line Edge Map (LEM), the dynamic two strip algorithm (Dyn2S) [19] is utilized to detect dominant points on the edge curves. The result of applying these processes on a face is illustrated in Figure 1. Subsequently, line features derived from different views and modals are combined at the feature level using the proposed multi-view line segment Hausdorff distance (MVLHD). The MVLHD provides an integrated similarity measure using the pooled line features from all views of the object.

Figure 1. An example of LEM.

3. Multi-view line segment Hausdorff distance

Consider an object recognition problem where objects are represented by a pooled feature from multiple views and modals. Assume that $M = \{M^1, M^2, \cdots, M^n\}$ is an exemplar object consisting of n views from n different viewpoints of multiple sensors and $T = \{T^1, T^2, \cdots, T^n\}$ is a test object consisting of n views from the same viewpoints of the sensors as in the exemplar object. 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 line sets representing the features in the n exemplar views of object M 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 line sets representing the features in the n views of the test object T from the same viewpoints of the sensors as in the exemplar object M. $p_1, p_2, \ldots, p_n$ and $q_1, q_2, \ldots, q_n$ are line numbers in the exemplar and the test views, respectively. Obviously, a

line $t_i^k \in T^k$ in the kth view of the test object T has to be only allowed to find its match in the corresponding view M^k of the exemplar object M.

The difference between lines m_i^k and t_j^k can be depicted by a vector $\bar{d}(m_i^k, t_j^k) = \left[d_\theta(m_i^k, t_j^k) \;\; d_{//}(m_i^k, t_j^k) \;\; d_\perp(m_i^k, t_j^k) \right]$, where the superscript k stands for the kth view. $d_\theta(m_i^k, t_j^k)$, $d_{//}(m_i^k, t_j^k)$, $d_\perp(m_i^k, t_j^k)$ are angular, parallel and perpendicular distances, respectively. The angular distance measures the structural/directional difference between m_i^k and t_j^k. The parallel and perpendicular distances measure the differences of geometrical locations.

Definition 1 (Angular distance): $d_\theta(m_i^k, t_j^k) = f(\theta(m_i^k, t_j^k))$. $\theta(m_i^k, t_j^k)$ computes the smaller intersecting angle between lines m_i^k and t_j^k. $f()$ is a non-linear penalty function to map the angle to a scalar. It is desirable to encourage small angular variation (which is most likely the intra-class variation) but penalize heavily on large deviation (which is most likely the inter-class difference). In this study, the tangent function is used.

Definition 2 (Parallel distance between horizontal lines): $d_{//}(m_i^k, t_j^k) = \min[diag(\mathbf{L}_{//mt})]$. Given two horizontal lines m_i^k and t_j^k in the kth view of M and T, the horizontal distances between any two of the four ends of lines m_i^k and t_j^k along the horizontal direction are defined by the matrices

$$\mathbf{L}_{//mt} = \begin{bmatrix} l_{m_l t_l} & l_{m_l t_r} \\ l_{m_r t_l} & l_{m_r t_r} \end{bmatrix} \quad \text{and} \quad \mathbf{L}_{//tm} = \begin{bmatrix} l_{t_l m_l} & l_{t_l m_r} \\ l_{t_r m_l} & l_{t_r m_r} \end{bmatrix}$$

where the subscripts l and r stand for left and right ends of lines m_i^k and t_j^k. Since $l_{m_l t_r} = l_{t_r m_l}$, we have $\mathbf{L}_{//mt} = \mathbf{L}^T_{//tm}$. $diag(\mathbf{L}_{//mt})$ denotes the set of diagonal elements of matrix $\mathbf{L}_{//mt}$. $d_{//}(m_i, t_j)$ is defined as the minimum distance in $diag(\mathbf{L}_{//mt})$, which means it only reflects the smallest shift of the line end-points. If one of the end-points of m_i^k is located exactly at the same horizontal location of the corresponding end-point of t_j^k and the other one shifts, the parallel distance remains zero no matter how far the other end-point of m_i^k shifts. This introduces tolerance to the shifting of line end-points due to segmentation errors.

Definition 3 (Parallel distance between horizontal lines): $d_{//}(m_i^k, t_j^k) = 0$, if $(m_l < t_l \cap m_r > t_r) \cup (m_l > t_l \cap m_r < t_r)$. In order to cater for the effect of shifting of line end-

points, $d_{//}(m_i^k, t_j^k)$ is reset to zero if one line is within the range of the other.

Definition 4 (Perpendicular distance between horizontal lines): $d_\perp(m_i^k, t_j^k)$ is the vertical distance between two horizontal lines m_i^k and t_j^k.

Definition 5 (Distance between arbitrary lines): In general, m_i^k and t_j^k would not be parallel and horizontal. The shorter line is then rotated with its midpoint as rotation center to make it parallel to the other line. Subsequently, the coordinate system is rotated such that the two lines are horizontal before applying Definitions 2-4 to compute $d_{//}(m_i^k, t_j^k)$ and $d_\perp(m_i^k, t_j^k)$.

The distance between two lines m_i^k and t_j^k is defined as

$$d(m_i^k, t_j^k) = \sqrt{(W \cdot d_\theta(m_i^k, t_j^k))^2 + d_{//}^2(m_i^k, t_j^k) + d_\perp^2(m_i^k, t_j^k)} \quad (1)$$

as $d_\theta(m_i^k, t_j^k)$, $d_{//}(m_i^k, t_j^k)$, $d_\perp(m_i^k, t_j^k)$ are independent. W is the weight of the angular distance to balance the contributions of the angle and the displacements. The directed Multi-View Line Segment Hausdorff Distance (dMVLHD) from M to T and from T to M are defined as follows.

$$dMVLHD(M,T) = \sum_{k=1}^{n} \frac{N^{M^k}}{\sum_{k=1}^{n} N^{M^k}} \bullet \left[\sum_{m_i^k \in M^k} \left(\frac{l_{m_i^k}}{\sum_{m_i^k \in M^k} l_{m_i^k}} \bullet \min_{t_j^k \in T^k} d(m_i^k, t_j^k) \right) \right]$$

$$(2)$$

$$dMVLHD(T,M) = \sum_{k=1}^{n} \frac{N^{T^k}}{\sum_{k=1}^{n} N^{T^k}} \bullet \left[\sum_{t_i^k \in T^k} \left(\frac{l_{t_i^k}}{\sum_{t_i^k \in T^k} l_{t_i^k}} \bullet \min_{m_j^k \in M^k} d(t_i^k, m_j^k) \right) \right]$$

$$(3)$$

where $l_{m_i^k}$ is the length of line m_i^k. N^{M^k} and N^{T^k} are the numbers of lines in the kth view of the model M and the test object T, respectively. The distance between the line m_i and its matched pair (i.e., the line t_j with minimum distance $\min\limits_{t_j^k \in T^k} d(m_i^k, t_j^k)$) is weighted by the normalized length of m_i^k (i.e., $l_{m_i^k} \big/ \sum\limits_{m_i^k \in M^k} l_{m_i^k}$) because its contribution to *dMVLHD(M,T)* is assumed to be proportional to the length of the line. The contribution from the kth view is weighted by its number of lines in percentage (i.e., $N^{M^k} \big/ \sum\limits_{k=1}^{n} N^{M^k}$). This is based on the assumption that the contribution from each view is proportional to the ratio of the number of features (i.e., lines) in that view over the total number of lines. When

the two lines to be matched are perpendicular (i.e., $\theta(m_i^k, t_j^k) = 90°$), which is the most unlikely match, the searching of t_j^k with minimum distance $d(m_i^k, t_j^k)$ will never match such a perpendicular pair. Hence, the calculation of tan90° is avoided by skipping it over. For a line $m_i^k \in M^k$ in the kth view of the exemplar object M, $dMVLHD(M,T)$ identifies its nearest neighboring line in the kth view T^k of the test object T and measures the distance from m_i^k to the identified line. $dMVLHD(M,T)$ in effect assigns each line in M a matched pair based on its distance to the nearest line in its corresponding view in T, then it uses the line length weighted distances of such lines as the measure.

Finally, the Multi-View Line Segment Hausdorff Distance (MVLHD) is defined as the maximum of $dMVLHD(M,T)$ and $dMVLHD(T,M)$ to represent the distance between the exemplar object (image set) M and the test object (image set) T:

$$MVLHD(M,T) = \max(dMVLHD(M,T), dMVLHD(T,M))$$
(4)

4. Experiments

Three experiments had been conducted to examine the performance of the proposed feature-level fusion approach for personal identification. They are (a) personal identification by fusing the frontal face, the profile face and the palm print, (b) personal identification by fusing 3 different view faces under appearance changes due to smiling expression, and (c) personal identification by fusing 3 different view faces under appearance changes due to speaking action.

In the first experiment, a face-profile-palmprint database was collected to evaluate the performance of the proposed system. The database contains 210 images of 35 people (29 males and 6 females), which is composed of two sets of images taken in two different sessions. In the first session, three images (one frontal face, one profile face and one palmprint) were captured from each person and stored as his/her model set in the database. In the identification stage, the frontal face, the profile face and the palmprint were captured again as the test image set in the second session, and compared to each model set in the database. The best matched model in the database with minimum MVLHD value is considered as the identified person. The experiment was conducted in a laboratory with fluorescent lights on top and windows on one side. The face images were taken with a neutral expression by two cameras at the same time.

The other two experiments were conducted on a public available face database [20] from the University of Stirling to evaluate the effectiveness of the proposed approach for multi-view fused personal identification

under expression changes and during speaking action. The database contains 311 images of 35 people (18 females and 17 males). 31 people (16 females and 15 males) have complete image set, which contains three poses (frontal view, ¾ view and profile view) and three expressions (neutral expression, smiling and speaking), and thus can be used as our testing data sets.

4.1 Identification using faces and palmprints

The performance of the proposed MVLHD system was examined by fusing frontal view face, profile view face and palmprint. The identification accuracies are summarized in Table 1. The single view LHD matching method [15] can correctly identify 82.86% of the frontal faces, 77.14% of the profile faces and 82.86% of the palmprints. The combined matching using MVLHD correctly identified 97.14% of the people. A significant 20% improvement in accuracy was obtained using the proposed approach. 100% of the people were identified in the top 2 identification scheme. In the top 2 identification scheme, the correct match is counted when the identical person of the input is among the best 2 matched persons from the models.

4.2 Identification under expression changes and during speaking action

The system is also evaluated using a public face database [20] with appearance changes due to smiling and speaking. In the experiments, the frontal, ¾ and profile views of neutral expression from each person were used as the three exemplar views of one model of the person. The algorithm was tested using the three views taken under non-rigid deformations due to smiling expression and speaking action, respectively. Note that each experiment was a single model based object matching. Each model or input object was represented by three images from different viewpoints. There is significant non-rigid distortion between the input object and the corresponding model.

The identification accuracies of the proposed MVLHD feature-level fusion approach on non-rigid distortions due to smiling and speaking are summarized in Tables 2 and 3 together with the accuracies of LHD on single views. It is found that the LHD method on single view matching only achieved an average accuracy of 77.42% for smiling faces and 88.17% for speaking faces. The proposed fusion method significantly improved the recognition rate up to 96.77% for smiling faces and 100% for speaking faces. 100% identification was achieved for smiling faces when the top 2 identification scheme was used. These are very encouraging identification accuracies considering the fact that faces are very similar from person to person with small inter-class variations and the non-rigid distortions produce large intra-class variations. 19.35% and 11.83% of increases in accuracy for distorted faces due to smiling

and speaking are significant and promising. The experimental results reveal that the MVLHD approach effectively provides a new similarity measure for image fusion at feature-level, which could be used for non-rigid object matching. These results are obtained by using an independent image database, which is publicly available. Thus it can be used as a benchmark for direct performance comparison by other researchers.

Table 1. The system performances of the MVLHD fusion approach on faces and palmprints together with the LHD single view matching method.

| | LHD (on single view) | | | | MVLHD (fusing frontal face, profile face and palmprint) | Improvement |
| | Face | | Palm print | Average | | |
	Frontal view	Profile view				
Accuracy	82.86%	77.14%	82.86%	80.95%	97.14%	20.00%

Table 2. The identification accuracies (%) of MVLHD and LHD on smiling faces.

| | LHD (on single view) | | | | MVLHD (fusing 3 views) | Improvement |
	Frontal view	¾ view	Side view	Average		
Accuracy	80.65%	70.97%	80.65%	77.42%	96.77%	19.35%
Top 3 identification	83.87	87.10	93.55	--	100	--

Table 3 The identification accuracies (%) of MVLHD and LHD on speaking faces.

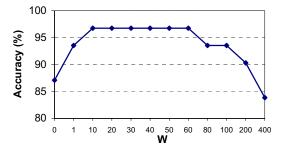| | LHD (on single view) | | | | MVLHD (fusing 3 views) | Improvement |
	Frontal view	¾ view	Side view	Average		
Accuracy	90.32	80.65	93.55	88.17%	100	11.83%
Top 3 identification	90.32	90.32	93.55	--	100	--

Figure 2. The effect of W.

4.3 Effect of *W*

The effects of *W* are illustrated in Figure 2. It is found that the system performance improved by introducing the LEM orientation features into the MVLHD calculation and reached the optimal value of 96.77% for smiling faces and 100% for speaking faces when *W* ranged from 20 to 60.

5. Conclusion

Traditionally, the personal identification decision is drawn on a single modality of biometric information. Recent studies show that the use of multiple modalities and views to solve automatic personal identification problem leads to tangible benefits in terms of accuracy and robustness. Multi-modal and multi-view systems can achieve better performance and are more foolproof than single modal and single view systems because different modals and views provide complementary information. The current studies of multi-modal and multi-view personal identification focused on combining the outputs of multiple classifiers at the decision level. The feature-level fusion still remains an unexplored area for personal identification.

In this paper, we have shown that it is possible for a personal identification system to combine multiple views and modals at the feature-level. A new similarity measure is proposed to integrate multiple view features from different viewpoints and from different sensors. Three

experiments had been conducted on human faces and palmprints to evaluate the feasibility and capability of the proposed technique for personal identification. The experimental results demonstrated that our technique consistently improved the system performance by fusing multiple views and/or modals at the feature level, which is in accordance with the belief in [8] mentioned in the introduction. The improvements were significant and accuracies were encouraging. 20% improvement in identification accuracy was achieved by combining frontal faces, profile faces and palmprints. 19% and 12% of increases in accuracy for identifying distorted faces due to smiling and speaking were obtained using frontal, three-quarter and profile view faces. Human faces are very similar in structure from person to person and the deformations caused by expression and speaking significantly increased the intra-face variations. The experimental results obtained in this research are thus very encouraging due to the difficulty of face matching under non-rigid deformations. Actually, handling the variability in appearance due to varying expression /deformation is one of the key problems in face recognition. This research demonstrates that multi-modal and/or multi-view feature-level fusion provides a new way for personal identification.

Furthermore, the proposed MVLHD is a generic similarity measure. We believe that it may also prove useful in other applications, in which multiple views and modals can be represented by line features.

Acknowledgements

6. References

[1] H. Lin and A. Jain, "Integrating Faces and Fingerprints for Personal Identification", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, pp. 1295-1307, 1998

[2] R. Brunelli and D. Falavigna, "Person Identification Using Multiple Cues", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, pp. 955-966, 1995

[3] J. Kittler, M. Hatef, R. P. W. Duin and J. Matas, "On Combining Classifiers", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, pp. 226-239, 1998

[4] J. Kitter, J. Matas, K. Jonsson and M. U. R. Sánchez, "Combining evidence in personal identity verification systems", Pattern Recognition Letters, Vol 18, pp.845-852, 1997.

[5] S. Ben-Yacoub, Y. Abdeljaoued and E. Mayoraz, "Fusion of Face and Speech Data for Person Identity Verification", IEEE Transactions on Neural Networks, Vol. 10, pp.1065-1074, 1999.

[6] U. Dieckmann, P. Plankensteiner and T. Wagner, "SESAM: A Biometric Person Identification System Using Sensor Fusion", Pattern Recognition Letters, Vol 18, pp.827-833, 1997.

[7] V. Chatzis, A. G. Borş and I. Pitas, "Multimodal Decision-Level Fusion for Person Authentication", IEEE Transactions on Systems, Man, And Cybernetics – Part A, Vol. 29, pp. 674-680, 1999.

[8]] A. K. Jain and C. Dorai, "Practicing Vision: Integration, Evaluation and Applications", Pattern Recognition, Vol. 30, pp.183-196, 1997.

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

[10] 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.

[11] I. Biederman and J. Gu, "Surface versus Edge-based Determinants of Visual Recognition", Cognitive Psychology, Vol. 20, pp. 38-64, 1988.

[12] V. Bruce et.al., "The importance of 'Mass' in Line Drawings of Faces", Applied Cognitive Psychology, Vol.6, pp. 619-628, 1992.

[13] J. H. Mcintish and K. M. Mutch, "Matching straight lines", Computer Vision, Graphics and Image Processing, Vol. 43, pp. 386-408, 1988.

[14] D. Zhang and W. Shu, "Two novel characteristics in palmprint verification: datum point invariance and line feature matching", pattern Recognition, Vol. 32, pp.691-702, 1999.

[15] Y. Gao and M. K. H. Leung, "Face Recognition Using Line Edge Map", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No.6, pp.764-779, 2002.

[16] J, You, W. Li and D. Zhang, "Hierarchical palmprint identification via multiple feature extraction", Pattern Recognition, Vol. 35, pp.847-859, 2002.

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

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

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

[20] University of Stirling face database. http://pics.psych.stir.ac.uk/

[21] M. H. Yang, D. J. Kriegman and N. Ahuja, "Detecting Faces in Images: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No.1, pp.34-58, 2002.

[22] E. Hjelmås and B. K. Low, "Face detection: a survey", Computer Vision and Image Understanding, Vol. 83, pp. 236-271, 2001.

[23] H. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, pp. 23-38, 1998.