Lower-quarter-based face verification using correlation filter

A Sleit*, R Abu-Hurra and **W AlMobaideen**
University of Jordan, Amman, Jordan

Abstract: Correlation filters have been applied successfully in pattern matching needed in face verification. In this paper, we propose a face verification technique based on half the lower face using the minimum average correlation energy (MACE) filter which has the ability to satisfy the correlation peak at the origin, and minimise the average correlation energy. Our experiments show that the application of the MACE filter for face verification using the lower-left or -right parts of the face generates authentication with at least 80% accuracy. Authentication accuracy using the MACE filter based on the whole face increases to at least 88%, but requires approximately twice as much response time.

Keywords: MACE filter, correlation filters, face verification, peak-to-sidelobe ratio, lower-face verification, authentication

1 INTRODUCTION

Biometric systems are automated methods of verifying or recognising the identity of a person on the basis of some physiological characteristics, such as a fingerprint, or face pattern and/or some aspect of behavior, such as spoken voice, or handwriting. Because most physiological features are virtually non-alterable without sever damage to the person, physiological-based biometric authentication methods are more popular than behavioral-based methods.[1] An important usage of the biometric verification is fighting crime, where evidence from the crime scene is compared with the data recorded earlier. Also, biometric verification is used in physical access, rights management, identity access management, surveillance and authentication for computer systems in the security field. There are two possible recognition tasks that may be considered:[1]

- person identification — an image of an unknown individual is collected ('probe') and the identity is found searching a large set of images (gallery). Matching becomes especially difficult when the probe is a duplicate rather than same (counterpart) image from the gallery
- identity verification — the system checks if a given probe belongs to a relatively small gallery labelled as a set of intruders. The system is usually flooded with thousands of faces (e.g. airport security).

The face-based recognition process consists of four main steps. The first step is image acquisition, which is achieved by using a camera to capture the images in grey or colour scale. Image processing is the second step which aims to reduce the noise and enhance the important features by applying filtering algorithms. The third step is face detection, where the desired face is detected, segmented and localised from the given image. The final step is face identification/verification, which decides if the test image is a client or an impostor.[2]

Identity verification using face information has taken a lot of attention over the past decades.[3–5] Correlation filters are successful solutions for automatic target recognition problems in addition to biometric verification.[6,7] Researchers argue that due to imperfect image acquisition, it is an absolute necessity to develop recognition techniques based on

The MS was accepted for publication on 18 July 2010.
** Corresponding author: Azzam Sleit, University of Jordan, Amman, Jordan; email: azzam.sleit@ju.edu.jo*

partial face verification.[8,9] This article proposes a face verification technique based on the availability of the lower face of the person. The proposed technique operates on 25% of the face in order to reduce storage requirements and increase the speed of computations.

Neo *et al.* used Faces-94 database, which consists of 153 individuals, each one with 20 images.[10] Face images of size 180×200 are normalised to 73×61, and cropped to several sizes (25, 50 and 75% of the face). This face verification technique is conducted by using different part-based linear subspace feature extractors to estimate the optimal face ratio, such as non-negative matrix factorisation, local non-negative matrix factorisation and spatially confined non-negative matrix factorisation. During the verification stage, the norm distance metric is applied on an input test image after being cropped.

Since thermal images are invariant to visible illumination variations, visual and thermal infrared face images are usually combined together by face recognition systems. The optimal tradeoff synthetic discriminant function (OTSDF) filter gives a recognition rate with a value equal to 96.5%[11] when using three training thermal face images. However, thermal images suffer from challenges, one of which is that some information may be lost around the eyes in thermal images of a person wearing eyeglasses. The other challenge is that body temperatures can be changed due to ambient temperatures or physical exercise, which leads to change in thermal signatures. Also, research has shown that the performance of infrared degraded when indoor and outdoor data are incorporated together.[12]

The minimum average correlation energy (MACE) as well as the unconstrained minimum average correlation energy (UMACE) filters are used to specify if a user access to the system is authentic or imposter. For each user, facial images in various positions must be collected in the database. Savvides *et al.* applied the MACE filter on illumination variation to perform face authentication and identification, and demonstrated the ability to suppress impostor faces using a universal threshold and shift invariance.[6]

This paper is organised as follows. The next section reviews some correlation filters and their application in face verification. Section 3 proposes a lower-face verification technique using the MACE filter. Section 4 presents experimental results obtained for applying the proposed technique using the lower-left and -right parts for face verification. Section 5 concludes this paper.

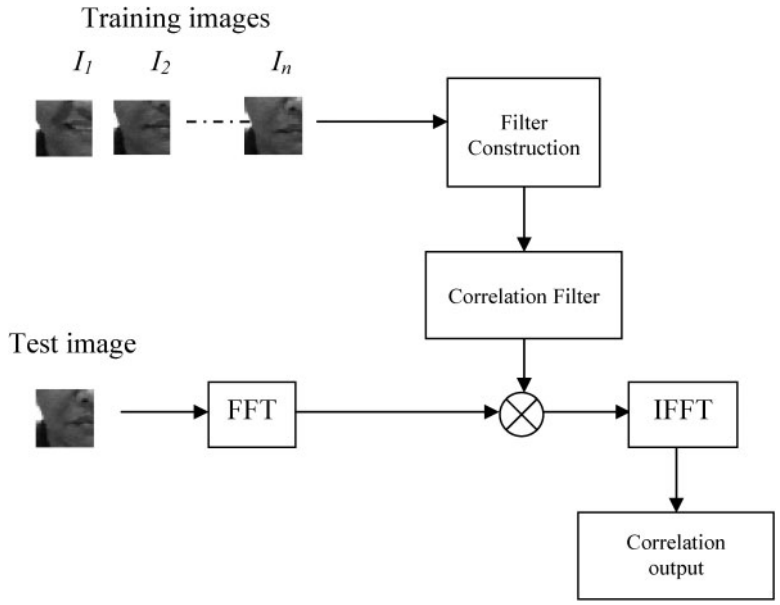
2 BACKGROUND AND RELATED WORKS

Correlation filter techniques are used to detect, locate and classify targets in scenes. They have been applied successfully in pattern matching needed in face verification. The advanced correlation filters can provide an excellent matching performance as they require minimal amount of memory storage with less complexity.[13]

Matched Spatial Filter (MSF) is the most basic correlation filter, which tends to generate optimal solution for recognition of known targets in the presence of additive white Gaussian noise. The MSF suffers from two major limitations, namely, the performance of the MSF decreases significantly with geometric image distortions such as scaling, and it cannot be applied for multiclass pattern recognition.[14] These limitations were addressed by developing a linear combination of MSFs called synthetic discriminate function (SDF) filter. In SDF filter, a single correlation filter is synthesised using several training images. The correlation outputs yield pre-specified value at the origin (known as peak constraints) according to the training images when an image is tested using the SDF filter. When the peak values are set to 1, they belong to authentic class. In this case, the SDF filter is called the equal correlation peak (ECP) SDF filter.[6,7,15] The ECP SDF filter controls the correlation values at the origin and nowhere else, which results with large sidelobes besides pre-specified correlation peak values. Moreover, the pre-specified peak values sometimes lead to misclassification because the sidelobes are larger than this pre-specified value of the peak.[16]

Advanced correlation filters such as MACE, UMACE, OTSDF and UOTSDF filters reduce the large sidelobes which are noticed from the ECP SDF filter.[14,16–18] In this research, we are interested in the MACE filter because of its ability to satisfy the correlation peak at the origin, and minimise the average correlation energy due to the training images of the correlation outputs. The resulting correlation planes achieve values close to zero everywhere except at the location of a trained object.[15]

The MACE filter first synthesises a single filter using multiple of training images. Then, the test image is cross-correlated with the filter causing the correlation output to exhibit sharp peaks if the test image belongs to the authentic class. Otherwise, the test image belongs to an imposter class.[17] The optimised MACE filter equation is as follows:



1 Block diagram of the correlation schema

$$\mathbf{h}_{\mathrm{MACE}} = \mathbf{D}^{-1}X(\mathbf{X}^{+}\mathbf{D}^{-1}X)^{-1}c \tag{1}$$

Suppose that we have from the authentic class N training images, each image containing d pixels. These images convert to the two-dimensional fast Fourier transforms (2-D FFTs), and we perform lexicographic ordering to convert 2-D FFT to one-dimensional (1-D) column vectors, the column vectors of the dXN matrix $\mathbf{X}$ in equation (1) contain these 1-D column vectors and c is a column vector contains the pre-specified correlation peak values of the training images which has length N. The matrix $\mathbf{D}$ is of length $d \times d$ and conations along its diagonal the average power spectrum of the training images. The 2-D $\mathbf{h}_{\mathbf{MACE}}$ is derived by reordering the 1-D $\mathbf{h}_{\mathbf{MACE}}$. The $+$ symbol in equation (1) refers to the complex conjugate transpose.[14]

For an authentic class of images we expect to see in the correlation output prominent peaks, while for an imposter class the correlation output has no noticeable peak. The peak-to-sidelobe ratio (PSR) is considered to measure the sharpness of the peak.[15,19] The PSR is computed as follows:

$$\mathrm{PSR} = (\mathrm{peak} - \mathrm{mean})/\sigma \tag{2}$$

The peak which is the brightness pixel in the centre is located, and then from the sidelobe region (usually 20×20), we exclude a central mask (usually 5×5) and calculate the mean and standard deviation σ.

Table 1 EER values for all clients of the AMP database

Client	Whole face EER	Lower-left EER	Lower-right EER
A	0	0	0
B	0.024	0.07	0.09
C	0.0129	0.026	0.08
D	0.02	0	0
E	0.052	0.03	0.052
F	0.02	0	0.01
G	0	0	0
H	0.03	0.07	0
I	0.029	0.018	0.13
J	0	0	0
K	0	0	0
L	0	0	0.021
M	0	0	0

2 An example illustrating the cropping step: (a) A0 image; (b) A0 lower; (c) A0L; (d) A0R

3 LOWER-QUARTER-BASED FACE VERIFICATION USING MACE FILTER

Identity verification systems have the main objective of deciding whether a claiming person is a client or impostor. This section proposes a lower-face verification approach using adaptive correlation filter. The proposed technique utilises half the size of the lower face in order to increase the speed of computations and decrease storage requirements. The lower part of the face contains parts like the mouth, nose and jaw. If the original size of the image is $L \times L$, our region of interest will be $(L/2) \times (L/2)$. Starting with a database of authentic face images with each person having several images with different facial expressions, we divide the lower face images vertically from the middle to form two parts left and right. Both parts will be studied in the verification process to determine which part of the lower face preserves its personal characteristics, and is more tolerant to facial expression in the verification process. Extracting features is constructed in the training stage by using MACE filter as per equation (1). For each person, two MACE filters which are the left MACE filter and right MACE filter are constructed. The left MACE filter is synthesised using the lower-left face training images. Among the images of each person, the training images are chosen depending on the largest variations from the lower-left face set. Similarly, the right MACE filter is synthesised using training images from the lower-right face set. Consequently, each person has two MACE filters: left and right MACE filters.

During the verification process, the test image will be cropped to form lower-left and -right face parts. The two parts are transformed by FFT, and then cross-correlation is performed between the lower-left face test image and each left MACE filter. A sharp peak is obtained in the correlation output in the case of authentication. Otherwise, the correlation output does not generate a strong peak. Figure 1 illustrates the schema of the correlation process.

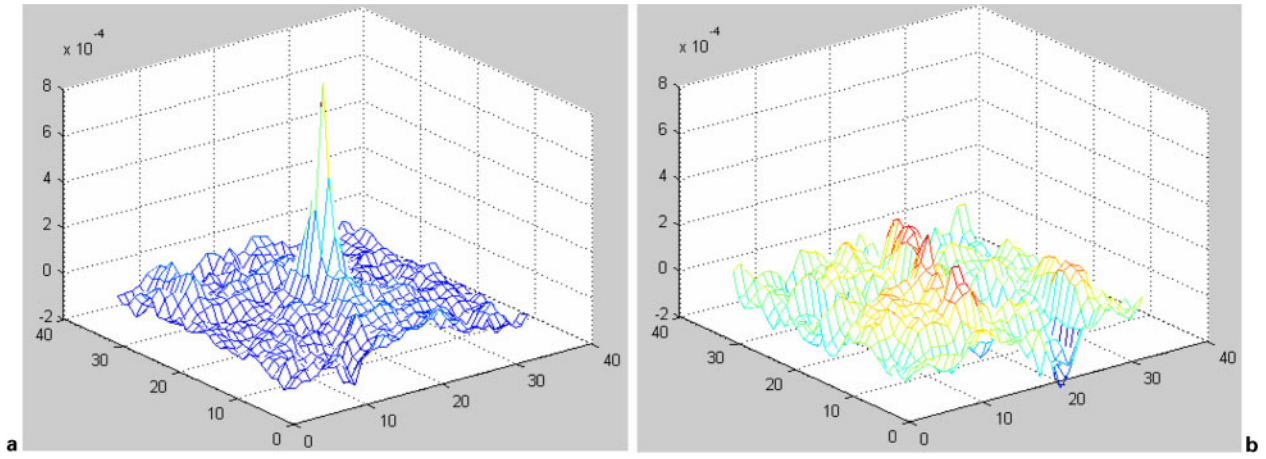
As mentioned earlier, the sharpness of a peak is measured by calculating the PSR from the left correlation output. If the PSR value obtained by equation (2) is greater than a specific threshold, the test image is a client. Otherwise, the test image is an impostor.

In the case of the right test image, cross-correlation is performed between the Fourier transform of the right test image and each right MACE filter. Similarly, the correlation output produces a sharp peak if the test image is from an authentic class.

Assume that we have a database of n authentic persons, each person having m images of different facial expressions. Figure 2a displays an example of facial expression A0 of person A. The A0 image is

Table 2 Performance of the verification process for all clients

	Whole face	Lower-left	Lower-right
A	100%	100%	100%
B	93.31–100%	83.4–100%	80.9–100%
C	93.31–100%	83.4–100%	80.9–100%
D	87.9–100%	95.2–100%	80.8–100%
E	87.9–100%	95.2–100%	80.8–100%
F	87.9–100%	95.2–100%	80.8–100%
G	100%	100%	100%
H	93.31–100%	83.4–100%	100%
I	87.9–100%	95.2–100%	80.8–100%
J	100%	100%	100%
K	100%	100%	100%
L	100%	100%	80.9–100%
M	100%	100%	100%



3 Left correlation planes for A0L: (a) correlation plane for class A; (b) Correlation plane for class B

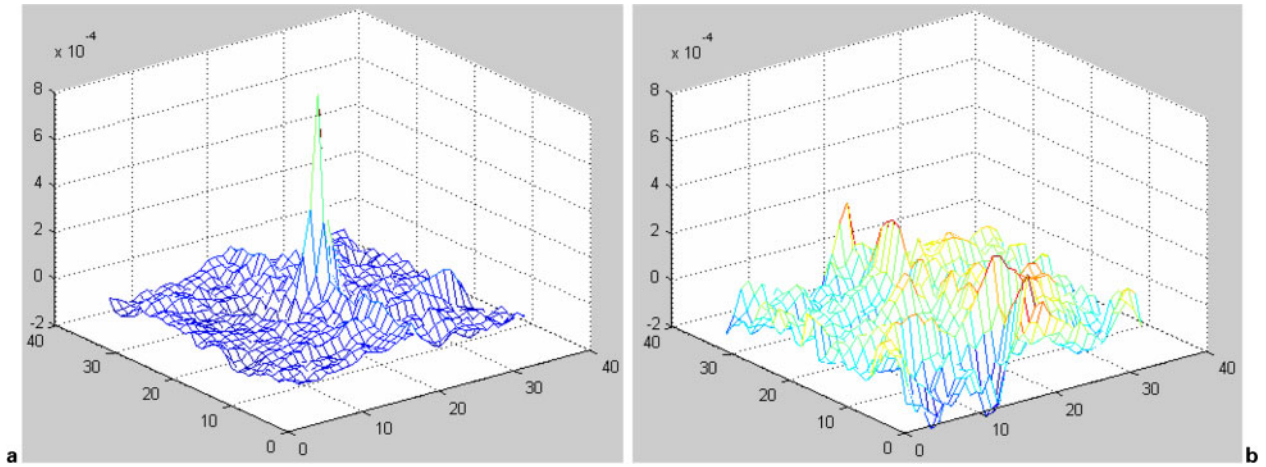
divided horizontally from the middle to obtain A0 lower face as per Fig. 2b. The lower face image is divided vertically to produce A0L for the left part and A0R for the right part as shown in Fig. 2c and d, respectively. Applying our identification approach on A0L, produces *m* correlation outputs; two of these correlation planes are shown in Fig. 3. The A correlation plane in Fig. 3a contains a strong peak, and all other correlation planes do not contain a peak. Therefore, the test image belongs to client A, and PSR will be calculated from this correlation output.

During the verification stage for A0R, m correlation outputs will be generated. Figure 4 illustrates two of the right correlation planes. The PSR is calculated from correlation output for class A because it has the maximum peak.

4 EXPERIMENTAL RESULTS

We utilised a facial expression database collected at the Advanced Multimedia Processing (AMP) Lab at the Electrical and Computer Engineering Department of Cornell University.[20] The database consists of 975 images representing 13 subjects such that each one has 75 images of different facial expressions. The faces were captured from a video sequence, where the movement of the user's head was tracked by a face tracker, and registered face images of size 64×64 pixels which were extracted depending on an eye localisation routine. Figure 5 shows example images from the database.

As mentioned before, the MACE filter requires a numbers of images for synchronisation. During the

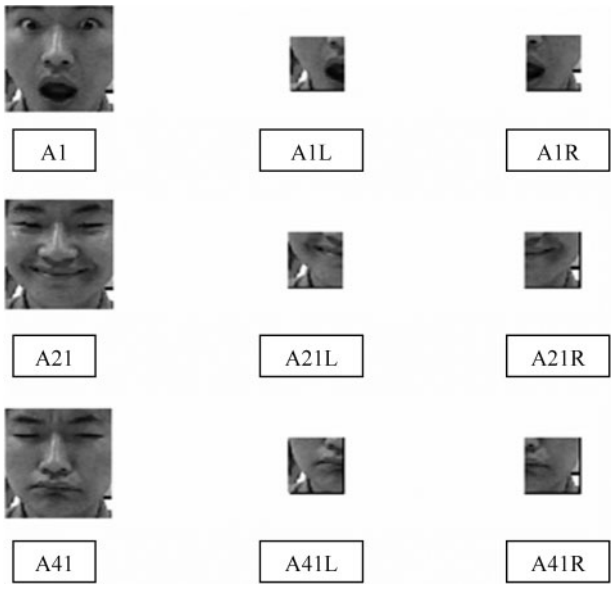


4 Right correlation planes for A0R: (a) correlation plane for class A; (b) correlation plane for class B

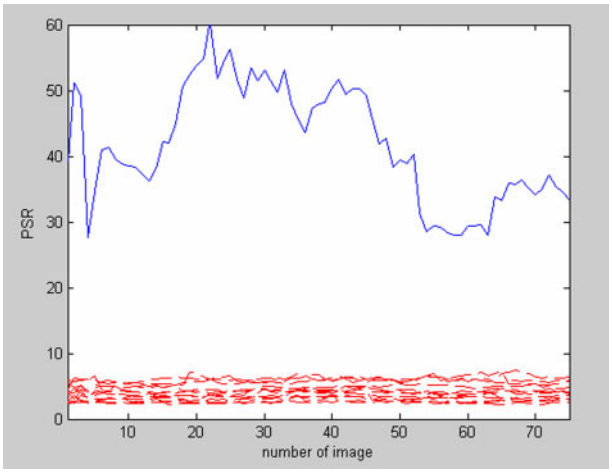
5 Sample images from the AMP Lab facial expression database

training stage, we used three training images of various facial expressions to build 13 left and right MACE filters. The 13 clients in the dataset are named as A, B, C, …, M. For client A, for example, there are 75 images named as A1, A2, …, A75. The lower-left and -right images of image A1 are named as A1L and A1R, respectively. Figure 6 displays the relevant images of A1, A21 and A41 which are used in the training stage to generate the left and right MACE filters of client A.

We evaluated the performance of the proposed technique using the 975 correlation outputs which result by cross-correlating each client's left MACE filter with the 975 images in the database. The PSRs were measured and recoded. The performance of the technique with respect to the right part was evaluated in the same way. Figure 7 displays the PSR values for the lower-left part of client L. The PSRs for all authentic images are displayed by the solid line, while those for imposter images are displayed by the dotted lines. The PSR values for the authentic images are above 27, whereas those for all imposter images are below 9. A wide margin of separation between the



6 Training images for client A



7 PSR plot of lower-left part of client L

minimum PSR value for the authentic images and the maximum PSR value for imposter images presents a good potential for the verification process.

We measured the performance of the verification by setting a PSR threshold value equal to T_0. Then, we calculated the false acceptance rate (FAR) and false rejection rate (FRR) as illustrated in equations (3) and (4), respectively:[9,19]

$$\text{FAR} = (\text{No. of imposter images with PSR} > T_0)/\text{No. of imposter images} \quad (3)$$

$$\text{FAR} = (\text{No. of authentic images with PSR} > T_0)/\text{No. of authentic images} \quad (4)$$

The performance of biometric systems is typically determined by measuring the equal error rate (EER) values. The FAR and FRR probabilities are plotted on a chart at different threshold values for each client, and the intersection between the two curves specify the EER values.[19] Table 1 displays the EER values which are computed for the whole face, lower-left and lower-right parts for each person.

We evaluated the performance of the verification process using the whole face, lower-left, and lower-right parts and the results are shown in Table 2. Using the MACE filters for all clients generated performance of 100% for clients A, G, J, K and M when using any of the three options. The performance of the technique ranges between 83 and 100% in the case of using the lower-left part, between 80 and 100% using the lower-right part, and between 88 to 100% using the whole face.

It is obvious that using the lower-left or lower-right parts of the face as a basis for authentication causes a 75% reduction in space requirements which can be of importance in the case of large databases. Furthermore, we measured the average authentication response time for 100 queries using half the lower part of the face versus that required for the whole face. We observed that an average 55% improvement in response time is obtained. We believe that the verification accuracy along with the enhancement in storage requirements and authentication response time justify the proposed approach.

5 CONCLUSION

This article investigates the application of the MACE correlation filter for face verification using the lower-left or -right part which represents 25% of the face. The MACE filter is known for its ability to satisfy the correlation peak at the origin, and minimise the average correlation energy due to the training images of the correlation outputs. The resulting correlation planes achieve values close to zero everywhere except at the location of a trained object. The proposed technique is motivated by the objective of reducing space and increasing computation speed. Our experiments show that the percentage of improvement for the speed of computation reaches 55% when using the left or right lower-quarter of the face during the verification process.

REFERENCES

1 Gi, P. N., Byung, J. K. and Kang, R. P. Robustness of face recognition to variations of illumination on mobile devices based on SVM. *KSII Trans. Internet Inf. Syst.*, 2010, **4**, 25–44.

2 Marcel, S. and Bengio, S. Improving face verification using skin color information, Proc. 16th Int. Conf. on *Pattern recognition: ICPR 2002*, Quebec City, Que., Canada, August 2002, IEEE Computer Society, Vol. 2, pp. 378–381

3 Wark, T.J., Sridharan, S. and Chandran, V. The use of speech and lip modalities for robust speaker verification under adverse conditions, Proc. IEEE Int. Conf. on *Multimedia computing and systems: ICMCS '99*, Florence, Italy, June 1999, IEEE Computer Society, Vol. 1, pp. 812–816.

4 Kittler, J., Li, Y., Mastas, J. and Sánchez, M. U. Lip-shape dependent face verification. *Lect. Notes Comput. Sci.*, 1997, **1206**, 61–68.

5 Zhu, X., Liao, S., Lei, Z., Liu, R. and Li, S. Z. Feature correlation filter for face recognition. *Lect. Notes Comput. Sci.*, 2007, **4642**, 77–86.

6 Savvides, M., Kumar, B. V. K. V. and Khosla, P. Face verification using correlation filters, Proc. 3rd IEEE Workshop on *Automatic identification advanced technologies*, Tarrytown, NY, USA, March 2002, IEEE Computer Society, pp. 56–61.

7 Samad, S. A., Ramli, D. A. and Hussain A. Lower face verification centered on lips using correlation filters. *Inf. Technol. J.*, 2007, **7**, 1146–1151.

8 Gutta, S. and Wechsler, H. Partial faces for face recognition: left vs right half. *Lect. Notes Comput. Sci.*, 2003, **2756**, 630–637.

9 Samad, S. A., Ramli, D. A. and Hussain, A. Region selection for robust face verification using UMACE filters, Proc. Int. Conf. on *Electrical engineering and informatics: ICEEI 2007*, Bandung, Indonesia, June 2007, Institute Teknologi Bandung, pp. 611–614.

10 Neo, H. F., Teo, C. C. and Teoh, A. B. J. A study on optimal face ratio for recognition using part-based feature extractor, Proc. 3rd Int. IEEE Conf. on *Signal-image technologies and internet-based system: SITIS 2007*, Shanghai, China, December 2007, IEEE Computer Society, pp. 735–741.

11 Heo, J., Savvides, M. and Vijayakumar, B. V. K. Performance evaluation of face recognition using visual and thermal imagery with advanced correlation filters, Proc. IEEE Computer Society Conf. on *Computer vision and pattern recognition: CVPR 2005*, San Diego, CA, USA, June 2005, IEEE Computer Society, pp. 9–19.

12 Jung, S.-W., Kim, Y., Teoh, A. B. J. and Toh, K.-A. Robust identity verification based on infrared face images, Proc. Int. Conf. on *Convergence information technology: ICCIT 2007*, Gyongju, Korea, November 2007, IEEE Computer Society, pp. 2066–2071.

13 Mahalanobis, A., Carlson, D. W. and Kumar, B. V. K. Evaluation of MACH and DCCF correlation filters for SAR ATR using the MSTAR public database. *Proc. SPIE*, 1998, **3370**, 460–468.

14 Casasent, D. and Hester, C. F. Multivariate technique for multiclass pattern recognition. *J. Appl. Opt.*, 1980, **19**, 1758–1761.

15 Mahalanobis, A., Kumar, B. V. K. and Casasent, D. Minimum average correlation energy filters. *J. Appl. Opt.*, 1987, **26**, 3633–3640.

16 Mahalanobis, A., Kumar, B. V. K., Song, S., Sims, S. R. F. and Epperson, J. F. Unconstrained correlation filters. *J. Appl. Opt.*, 1994, **33**, 3751–3759.

17 Kumar, B. V. K. Tutorial survey of composite filter designs for optical correlators. *J. Appl. Opt.*, 1992, **31**, 4773–4801.

18 Kerekes, R. A., Savvides, M., Kumar, B. V. K. and Sims S. R. F. Fractional power scale-tolerant correlation filters for enhanced automatic target recognition performance. *Proc. SPIE*, 2005, **5807**, 317–328.

19 Tahir, N. M., Hussain, A., Samad, A. and Husain, H. Posture recognition using correlation filter classifier. *J. Theor. Appl. Inf. Technol.*, 767–773.

20 Advanced Multimedia Processing Lab, Electrical and Computer Engineering Department at CMU. Available at: <http://amp.ece.cmu.edu> [accessed 10 March 2010].