



Journal of Applied Sciences

ISSN 1812-5654

science
alert

ANSI*net*
an open access publisher
<http://ansinet.com>

On The Use of Advanced Correlation Filters for Human Posture Recognition

¹Nooritawati Md. Tahir, ¹Aini Hussain, ¹Salina Abdul Samad, ¹Hafizah Husain and ²Andrew Teoh Beng Jin

¹Signal Processing Research Group, Department of Electrical, Electronics and Systems,
Faculty of Engineering, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor DE, Malaysia

²Faculty of Information Science and Technology, Multimedia University,
Jalan Ayer Keroh Lama, Bukit Beruang, 75450, Melaka, Malaysia

Abstract: This study affords the method of using advanced correlation filters in human posture recognition task. Two types of correlation filters were implemented and their efficacy evaluated. The correlation filters under consideration are Minimum Average Correlation Energy (MACE) and Unconstrained Minimum Average Correlation Energy (UMACE). Initial results prove that correlation filters offer significant potential used in posture recognition task with UMACE outperforming the MACE filter. In this research, both filters were subjected to a challenging task to recognize human posture without any restriction on the gender, clothing and posture variations. The UMACE filter performs remarkably well with an average accuracy of 89% compared to MACE filter which attained 42%.

Key words: Advanced correlation filters, posture, recognition

INTRODUCTION

One of the central topics in computer vision is the recognition of human postures. The recognition of posture is one step in the global process of analyzing human behaviour. Behaviour analysis is a vital field in dealing with many applications such as video surveillance that detect and track a person and recognize special actions such as theft, activity recognition and pedestrian detection. Accordingly, the major goal is to accomplish a more natural means of interaction with computers. In this paper we describe a novel methodology of recognizing four main postures namely standing, sitting, bending and lying position using advanced correlation filters. Previous methods done by Haritaoglu *et al.* (1998, 2000), Wren *et al.* (1997), Bobick and Davis (2001) and Iwasawa *et al.* (1997) that are based on 2D appearances used similar approach. Firstly, the principal parts of the body such as head, hands and feet are detected and based on these detections the secondary parts of the body such as shoulders, elbows and knees are revealed. Haritaoglu *et al.* (1998, 2000) segmented the silhouette from the background and computed the vertical and horizontal projections of the silhouette to determine the global posture of the person either standing, sitting, crawling-bending and laying.

To perform posture recognition, the system computed the projections of the current silhouette and compared to the model of projection realized for a set of predefined posture. Alternatively, Iwasawa *et al.* (1997) proposed a method that determines the center of gravity of the human silhouette. Next, they compute the orientation of the upper half of the body. Finally, the significant points such as feet, hands, elbows and knees are estimated by using a heuristic contour analysis of the human silhouette. The principal drawback of these methods is the dependency on the point of view. However these methods are not resource demanding since only one camera is used and therefore are adapted in real time applications.

Recently, advanced correlation filters have been investigated and they have evolved into very effective algorithms for pattern recognition applications specifically biometrics verification (Hennings and Vijaya Kumar, 2004; Jingu *et al.*, 2005; Savvides and Vijaya Kumar, 2003; Venkataramani and Vijaya Kumar, 2004; Chong *et al.*, 2006). For instance, Savvides and Vijaya Kumar (2003) examined the performance of advanced correlation filters for face authentication. The results are based on the illumination subsets of the CMU PIE (Carnegie Mellon University Pose, Illumination and Expression) database. They also presented methods that reduce the memory

requirements of these filters to run on limited computational resources including computationally efficient methods of synthesizing these filters. Further, they described an online training algorithm implemented on a face verification system for synthesizing correlation filters from a video stream to handle pose and scale variations. Their system also uses an efficient scheme to perform face localization within the current framework during the authentication stage. Another related study is by Jingu *et al.* (2005) on the evaluation of face recognition performance based on visual and thermal infrared (IR) face images using advanced correlation filter methods. They used MACE and OTSDF (Optimal Tradeoff Synthetic Discriminant Function) filters and achieved better performance over commercial face recognition algorithms such as FaceIt®. Their study proved that correlation filters performed very well when the face images are of significantly low resolution and could be applied in human identification at a distance (HID). They also described in detail a fully automated way of eyeglass detection and removal in thermal images resulting in a significant increase in thermal face recognition performance. Venkataramani and Vijaya Kumar (2004) also evaluated the performance of composite correlation filters in fingerprint verification for access control applications. They focused in obtaining digital live-scan fingerprints from sensors, rather than the inked fingerprints usually used in criminal identification. The NIST Special Database 24, obtained from an optical fingerprint sensor, was used to evaluate the performance of fingerprint verification in the presence of distortion. Their results showed that unconstrained filter was a good choice since it can be incrementally updated to reduce complexity demonstrating the distortion tolerance potential of correlation filters and performed reasonably well even with low resolution images. On the other hand, Hennings and Vijaya Kumar (2004) introduced the application of correlation filter classifiers for palmprint identification and verification. They described how the extraction of an appropriate region of interest in the palmprint surface could be used to design correlation filters that accomplish 100% recognition on a database of 50 persons. Recently, Chong *et al.* (2006) developed a private biometrics formulation that is based on the concealment of random kernel and the iris images to synthesize a MACE filter for iris authentication. The purpose was to provide private biometrics realization in iris authentication in which biometric template can be reissued once it was compromised. The proposed method was able to decrease the computational load by reducing the filter size and thus, improved the authentication rate significantly.

In this study, in an attempt to further demonstrate the efficiency and robustness of correlation filters, we investigate the possible use of advanced correlation filters in our posture recognition system. The main advantage of this method over previous approaches of posture recognition (Haritaoglu *et al.*, 1998, 2000; Wren *et al.*, 1997; Bobick and Davis, 2001; Iwasawa *et al.*, 1997) is that no feature extraction stage is involved since the image pixels are the input domain for training the filters. This particular property of correlation filters has been proven effective in recognition and classification (Sao and Yegnanarayana, 2004; Henning and Vijaya Kumar, 2004). As such, in this study we aim to illustrate the additional abilities of correlation filters in recognizing posture images.

MATERIALS AND METHODS

Firstly, four categories of human postures namely standing, bending, sitting and lying position were chosen for classification purpose. The database of images is recorded under normal office lighting system in an indoor environment. The system implementation includes background subtraction and silhouette extraction. In this work, we assume a static background and the background subtraction are achieved by thresholding the difference between the current frame and the static background image. In doing so, a human silhouette is extracted. Further, advanced correlation filters will feat as the posture classification system. Advanced correlation filters also known as composite filters, are the family of correlation-based classifiers. A template is used to correlate a test image and look for a sharp peak in the output plane that ideally resembles an impulse. This will be the evident characteristic of correlation outputs corresponding to images of the correct posture. In contrast, when a different posture image is correlated, the output shape does not contain a well-defined peak and it indicates that the image is a false or incorrect posture. This concept is shown in Fig. 1. The performance of such correlation templates depends highly on the set of images used for computing the template. Indeed, advanced correlation filters have been thoroughly studied in the last two decades and they have evolved into very effective algorithms for pattern recognition applications (Gader *et al.*, 2004; Sao and Yegnanarayana, 2004; Hennings and Vijaya Kumar, 2004). Advanced correlation filters can be designed to accommodate the intrinsic amplitude variability of the images in the training set while being tolerant to noise pervading the images (Jingu *et al.*, 2005; Savvides and Vijaya Kumar, 2003).

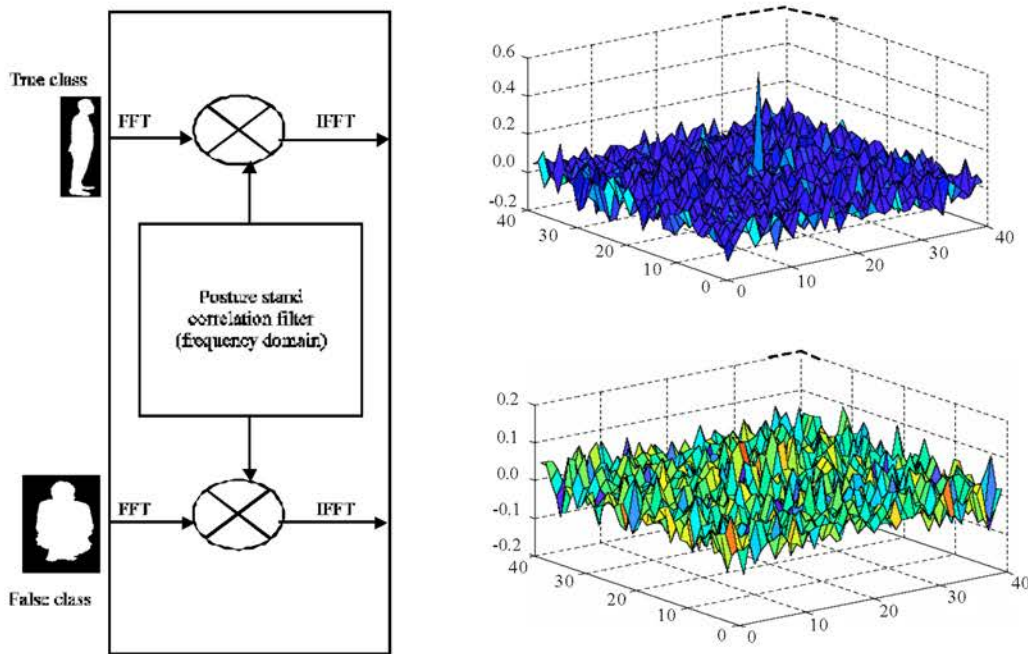


Fig. 1: Application of advanced correlation filter in posture recognition

In present experiments, 150 images of size 64×64 , for each category of the four main postures are chosen as our database to evaluate the performance of MACE and UMACE for posture recognition. We first applied 3 training images comprising of the front and side posture views in order to synthesis the MACE filter of posture. This process continues as the number of training images are increased in step of 3 each time until reaching 15 images. To evaluate the performance of each posture, cross correlations of all the images in the dataset were computed using each posture MACE filter resulting in $150 - 15 = 135$ correlation outputs corresponding to true class posture and $150 \times 3 = 450$ false class postures. The corresponding PSRs were then measured, recorded and evaluated.

Overview of composite filters: The fundamental motivation for designing correlation filters was driven by distortion invariant optical pattern recognition. Muller and Herbst (2002) introduced the Synthetic Discriminant Function (SDF) approach for this purpose in 1980. SDF is a linear combination of matched spatial filters whereby the weights are chosen so that the correlation outputs corresponding to the training images will yield equivalent correlation peak values at the origin. Basic SDF is also

referred to as the Equal Correlation Peak (ECP) SDF. The drawback of the ECP SDF is twofold. First, it cannot tolerate significant input noise. Second, ECP SDF has no built-in shift-invariance capability. Thus, shifts in the input target are replicated in the correlation output. Nevertheless, the SDF method has deeply influenced the current design of advanced correlation filters although the idea was actually introduced more than twenty years ago. To achieve robustness to noise, Vijaya Kumar *et al.* (2005) introduced the MVSDF (Minimum Variance Synthetic Discriminant Function). There are two drawbacks to this method. The first is that the MVSDF also controls only one point in the correlation map, just like the ECP SDF. The second is that the variance of the noise matrix must be made known beforehand in order to design the filter. However, even if the latter is known exactly, MVSDF is impractical because it requires inverting a large noise covariance matrix (Vijaya Kumar *et al.*, 2005). The MACE (Minimum Average Correlation Energy) filter was an attempt to control the entire correlation plane. Vijaya Kumar *et al.* (2005) reduced correlation function levels at all points except at the origin of the correlation plane and thereby obtained a very sharp correlation peak. However, MACE filters often suffer from two main drawbacks. Firstly, there is again no built-in immunity to noise. Secondly, the MACE filter is often excessively

sensitive to intra-class variations. Nevertheless, MACE leads to a useful frequency domain design approach for object recognition. Studies have shown that hard constraints on correlation values at the origin are not only unnecessary but can be counterproductive (Vijaya Kumar *et al.*, 2005). Hence, Unconstrained Correlation (UC) filters were introduced.

In this study, we have implemented one of the special UC filters, called UMACE short for Unconstrained Minimum Average Correlation Energy. For notation, matrices are denoted by upper case and vectors by bold italic characters. Upper case symbols refer to the frequency domain, while lower case symbols represent quantities in the spatial domain. The symbols T , $*$ and $+$ denote the conjugate transpose, complex transpose and complex conjugate transpose, respectively. Let us consider a set of images, each of which has a size $d_1 \times d_2 = d$. By assuming that each image is represented in a d -dimensional image space and given by a column vector of the form:

$$x = x(n) = [x(1), x(2), \dots, x(n), \dots, x(d)]^T \quad (1)$$

Next, let x_i denotes the i th training image, where $i = 1$ to N and X_i is the Fourier Transform of x_i .

Thus

$$X_i = X_i(k) = [X_i(1), X_i(2), \dots, X_i(k), \dots, X_i(d)]^T \quad (2)$$

and

$$X = [X_1, X_2, \dots, X_i, \dots, X_N]$$

Where:

$X_i(k)$ = The k th element in the frequency domain vector X_i ,
 $k = 1, \dots, d$.

Note that X is a $d \times N$ matrix with N column vectors X_i . Let the column vector h be the impulse response of the correlation filter and H its frequency response. We then define c_i as the correlation function of the i th image x_i with h , such that:

$$c_i = x_i \otimes h = \text{IFFT} \{X_i(k) \cdot H(k)\}$$

and
$$C_i = \text{FFT} \{c_i\} \quad (3)$$

MACE (Minimum Average Correlation Energy) filters: Vijaya Kumar *et al.* (2005) has defined several quantities that are useful for formulating the design approach in terms of an optimization problem. The overall objective is

to determine H so that the average energy is minimized in the correlation plane and a sharp peak is obtained at the origin. The quantities are given as follows:

(a) The power spectrum of X_i :

$$D_i = D_i(k,k) = |X_i(k)|^2 \quad (4)$$

where, D_i is a diagonal matrix of size $d \times d$ whose diagonal components are the magnitude squared of the associated component of X_i . It is also the power spectrum of x_i .

(b) The energy of the i th correlation plane:

$$E_i = \sum_{n=1}^d |c_i(n)|^2 = (1/d) \sum_{k=1}^d |C_i(k)|^2 = (1/d) \sum_{k=1}^d |H(k)|^2 |X_i(k)|^2 = H^* D_i H \quad (5)$$

(c) The correlation peak amplitude at the origin:

$$c_i(0) = X_i^* H = u_i \quad (6)$$

In general, u_i is user-defined where $i = 1, \dots, N$. All u_i belonging to the same postures class are set to 1; otherwise they are set to 0.

(d) The average correlation plane energy:

$$E_{av} = (1/N) \sum_{i=1}^N E_i = (1/N) \sum_{i=1}^N H^* D_i H = (1/N) H^* \left(\sum_{i=1}^N D_i \right) H \quad (7)$$

The optimal design problem solved by the MACE filter H can be expressed as:

$$\min_H E_{av} = \min_H (1/N) H^* D H$$

Where:

$$D = \sum_{i=1}^N D_i \quad (8)$$

subject to $c_i(0) = X_i^* H = u_i$

Where:

$$i = 1, \dots, N.$$

It can be shown Eq. 6 that the optimal solution of the MACE filter is given by:

$$H = D^{-1} X (X^* D^{-1} X)^{-1} u$$

Where:

$$u = [u_1, \dots, u_N]. \quad (9)$$

If we have N training images each of size $d_1 \times d_2$ pixels, then X in Eq. 9 is a matrix of size $L \times N$, where L is the total number of pixels in a single training image ($L = d_1 \times d_2$). X is a matrix that contains along its columns lexicographically re-ordered versions of the 2-D Fourier transforms of the N training images. D is a diagonal matrix of dimension $L \times L$ containing the average power spectrum of the training images along its diagonal. The column vector u contains N entries, corresponding to the desired values at the origin of the correlation plane of the training images.

UMACE (unconstrained minimum average correlation energy) filters: UMACE is another variant of the MACE filter (Savvides and Vijaya Kumar, 2003) and in this case, instead of imposing a hard constraint on the origin of the correlation plane, its height at the origin is free to increase according to the test data. The optimized filter equation is given as $H = D^{-1} m$ where m is a column vector containing the average of the 2D Fourier transforms of the training images. UMACE filters are computationally more attractive as they require the inversion of only a diagonal matrix. Noise tolerance can be built in to the filters as described in (Savvides and Vijaya Kumar, 2003). This is done by substituting D with D and $D = \alpha D + \text{sqrt}(1 - \alpha^2)C$,

where, C is a diagonal matrix containing the noise power spectral density. For white noise, C is the identity matrix and α ranges from 0 to 1 and is chosen to trade-off noise tolerance for discrimination. Note that an α of 1 yields a MACE filter.

Peak-to-Sidelobe Ratio (PSR): The MACE and UMACE methodologies determine a different filter for each member of the same class. In other words, given a test image, we wish to ascertain how similar it is to a MACE or UMACE filter. However, the design procedures discussed above output a complete correlation plane. Savvides and Vijaya Kumar (2003) Ventakaramani and Vijaya Kumar (2004) have suggested the Peak-to-Sidelobe Ratio (PSR) as a summary of the information in each correlation plane. Thus, the PSR is used to evaluate the degree of similarity of correlation planes. The Peak-to-Sidelobe Ratio (PSR) is a metric for measuring the overall correlation function for a test image. The significance of the PSR is that it measures the sharpness of the correlation function.

RESULTS AND DISCUSSION

The recognition performances of the MACE filter to discriminate all postures are as shown in Fig. 2. Clearly,

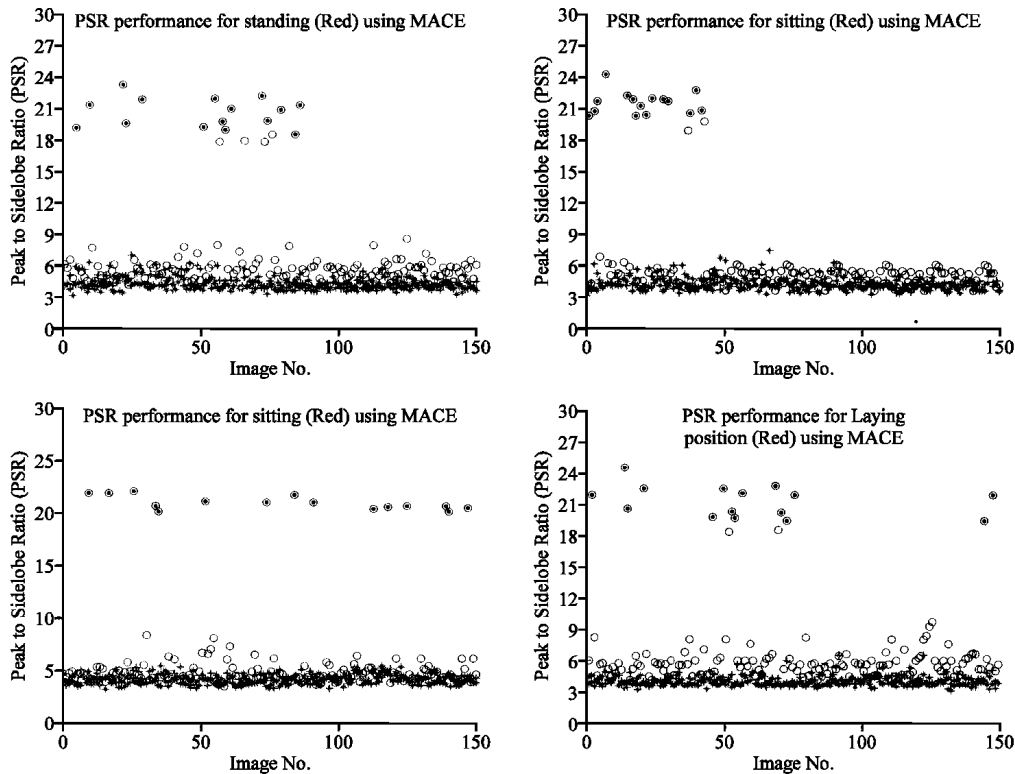


Fig. 2: PSR score for each class posture using MACE filter. The black dots indicate the correlation peaks for the training images

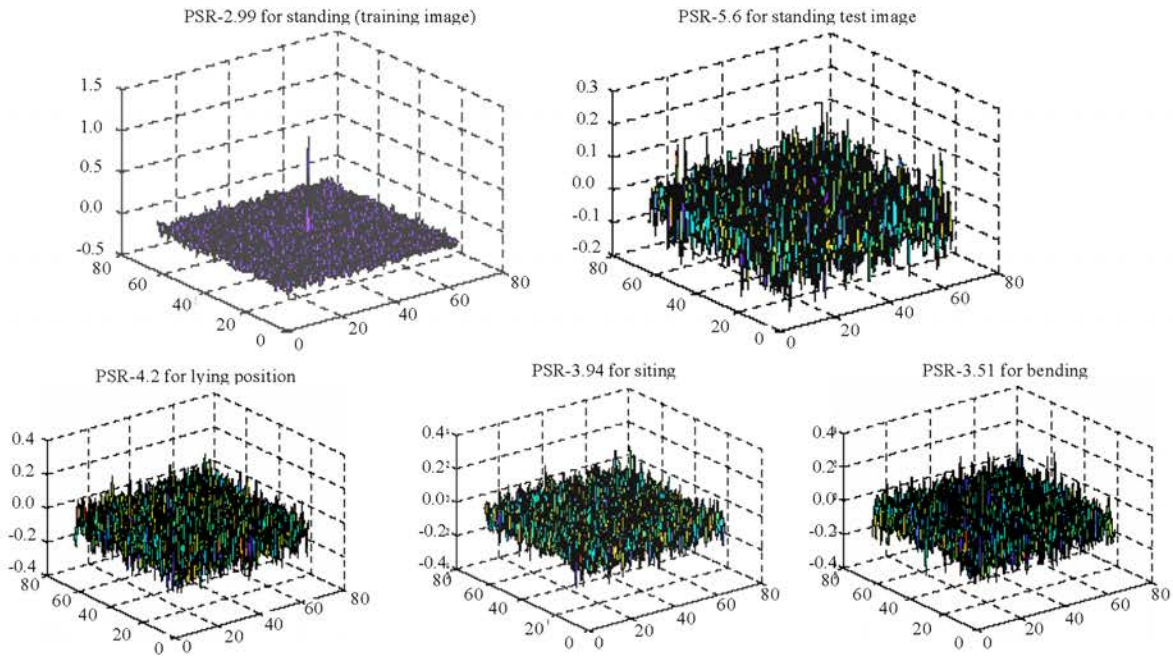


Fig. 3: Sample Correlation outputs using MACE filter designed for Standing Posture. Top (right): Input is a standing posture belonging to training set. Top (left): Input is a standing posture from test data. Bottom: Inputs are posture images from false class

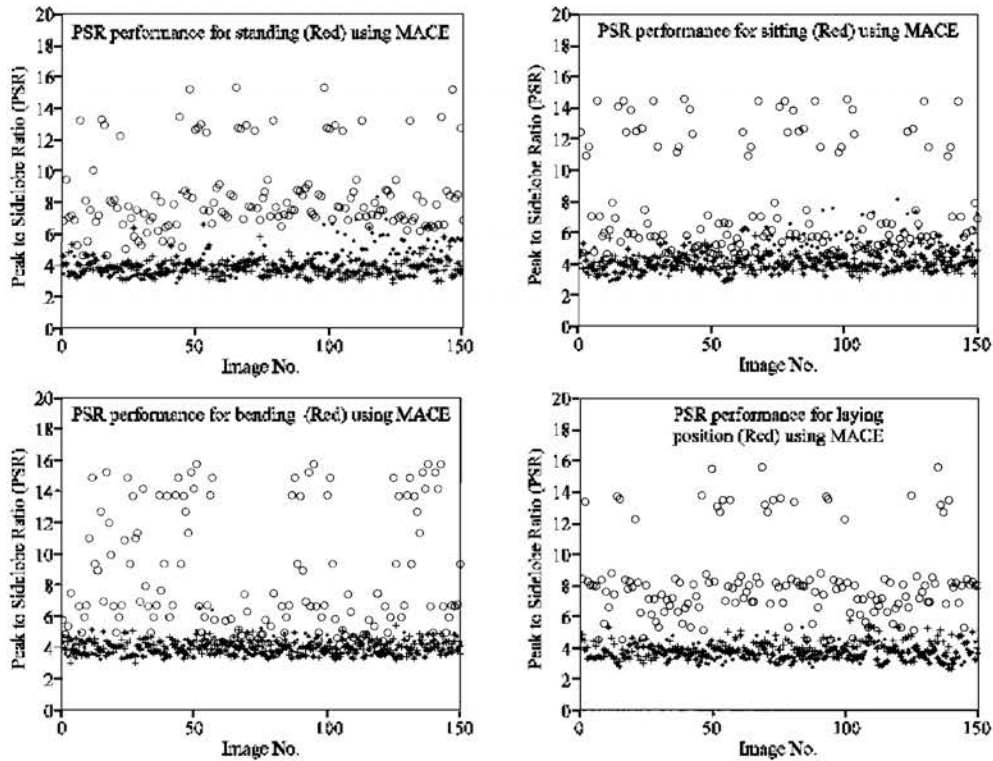


Fig. 4: PSR score for each class posture using UMACE filter

MACE filter fails to distinguish between the true class posture and the false group postures for all categories of postures. The black dots indicate the correlation peak for the training images used to synthesize the MACE filter for all posture and as expected they yield high PSR values. Subsequently, Fig. 3 shows the correlation output for standing posture using MACE filter. The sharp correlation peak resulting in a large PSR value is the training image. The left correlation output shows the response to a test image of the true class posture followed by the false class postures namely sitting, bending and lying position. The poor performance of MACE is expected. It is due to variations in the test dataset that were not adequately captured by the training set since human postures have very flexible and, not so well defined, complex shapes. In face recognition (Sao and Yegnanarayana, 2004), fingerprint verification (Venkataramani and Vijaya Kumar, 2004), palmprint recognition (Hennings and Vijaya Kumar, 2004) and iris recognition (Chong *et al.*, 2006) applications, increasing the number of training images help improve authentication, but in our study, adding more training images do not necessarily enhance the PSR values. These are due to the downside of MACE filter that has no in

built immunity to noise and therefore, could not accommodate the intra-class variations. Next, UMACE filter is applied as an alternative to MACE filter in order to attain better posture recognition. We repeat the same procedures using UMACE in order to determine the PSR performances in each posture class. As can be seen, UMACE filter produces an acceptable significant margin that discriminates the true class posture and the false category (Fig. 4).

Present experimental results for the UMACE filter showed that by increasing the number of training images has helped increase the margin between the true class and the false class posture and thus, giving better discriminating ability. It was also observed that the variance of the false class is reduced as the number of training images is increased. In the verification stage of using UMACE filter, the following three quantities were used to select the proper threshold namely False Acceptance Rate (FAR), False Rejection Rate (FRR) and Equal Error Rate (EER). For a given verification PSR threshold θ , for a class, the performance can be measured by the False Acceptance Rate (FAR) and False Rejection Rate (FRR), defined as follows:

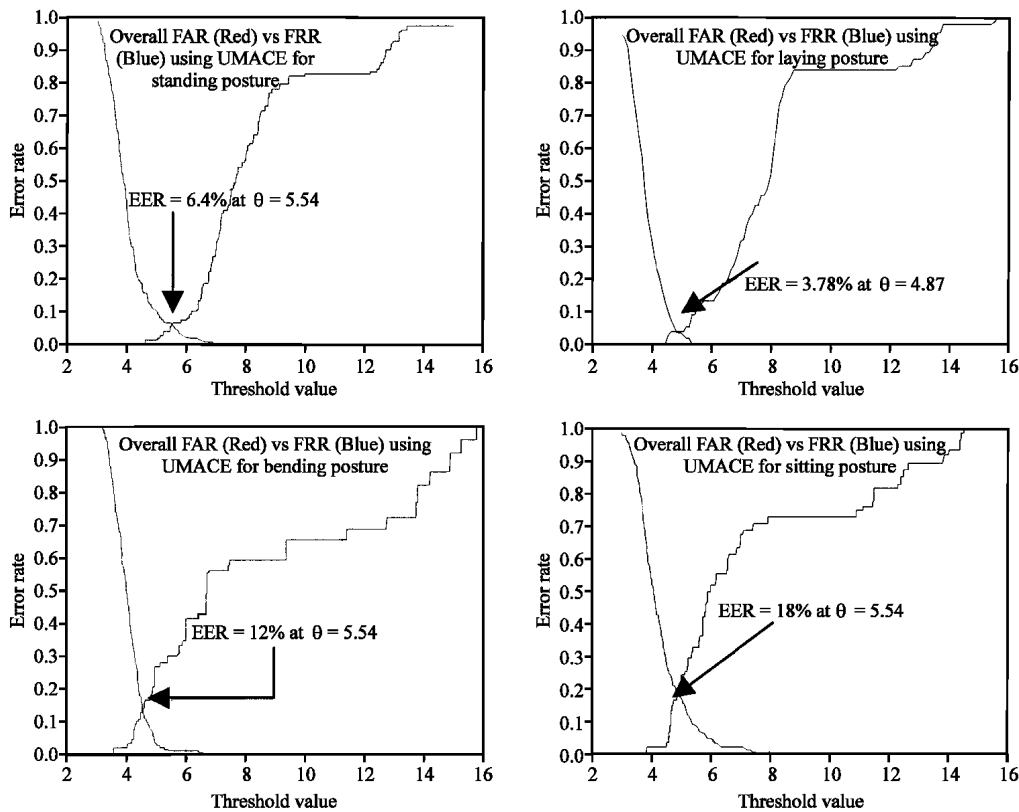


Fig. 5: FAR vs threshold and FRR vs threshold. ERR is obtained by adjusting FAR vs FRR

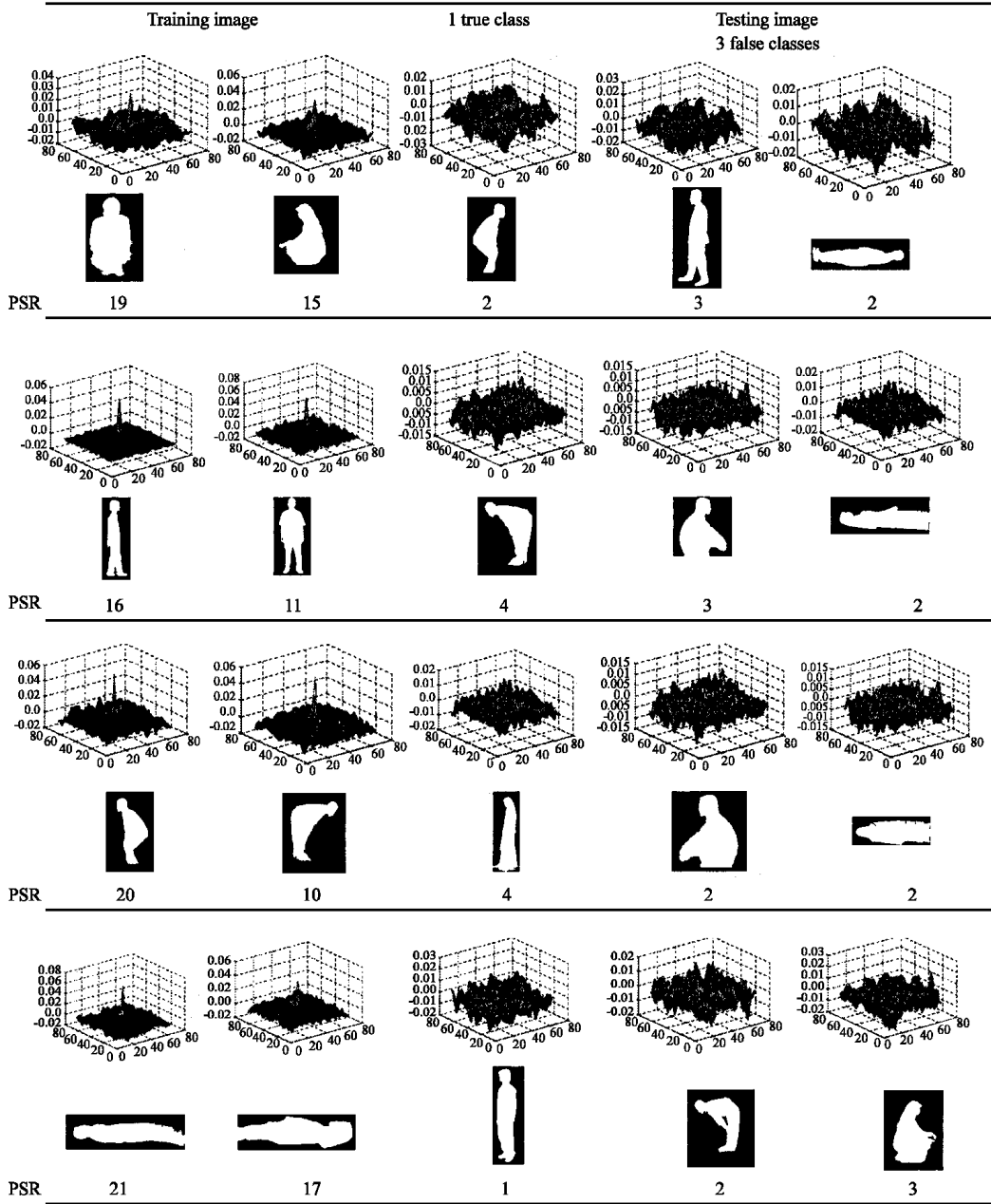


Fig. 6: Correlation outputs for each posture using UMACe filter with 12 training images

$$FAR = \frac{\text{No. of false posture images having } PSR > \theta_y}{\text{Total No. of false posture images}}$$

$$FAR = \frac{\text{No. of false posture images having } PSR < \theta_y}{\text{Total No. of true posture images}}$$

When, FAR and FRR are equal, the common value is referred to as Equal Error Rate (EER). Next, each posture in the database was verified against every other

posture. For a particular UMACe filter, all of the false class postures (150×3 = 450) and true class posture (150- number of training images used) criterion scores were obtained. Table 1 shows the average EER for each posture using 3 images and then being increased in step of 3 until 21 images. It is noted that at 12 training images offered almost equal EER performances even if the number of training images is increased. Therefore we conclude that 12 training images are acceptable as our templates.

Table 1: EER for all 4 UMACE filters synthesized using different number of training images

| No. of images | EER | | | |
|---------------|--------|-------|-------|--------|
| | Stand | Sit | Bend | Lying |
| 3 | 0.1200 | 0.270 | 0.230 | 0.0900 |
| 6 | 0.0960 | 0.254 | 0.210 | 0.0860 |
| 9 | 0.0920 | 0.230 | 0.180 | 0.0520 |
| 12 | 0.0643 | 0.200 | 0.144 | 0.0390 |
| 15 | 0.0640 | 0.196 | 0.133 | 0.0382 |
| 18 | 0.0639 | 0.180 | 0.129 | 0.0370 |
| 21 | 0.0635 | 0.175 | 0.120 | 0.0363 |

Subsequently, the threshold value to be selected for each posture is determined from the performance curves for all postures using UMACE filter as shown in Fig. 5. Thus, the lower the EER, the superior the overall performance of the verification system. As expected, it is obvious from the results that UMACE filter performs better and could separate the true class posture from the false category particularly for standing and lying down postures with recognition results of 96.2 and 93.6% respectively. Some samples correlation outputs of the postures are presented in Fig. 6. The best testing model that matches the training image will have a large PSR value. The PSR value reflects the UMACE filters ability to recognize and verify similarity between postures. It can be realized in Fig. 6 that the true class posture has higher PSR than the false category.

CONCLUSIONS

The performance of advanced correlation filters specifically MACE and UMACE for posture verification is presented. Based on the results obtained, the UMACE filter is found to be more robust compared to MACE with respect to variation in postures. It was determined that using 12 training images, the UMACE filters are capable of discriminating over 90% of the true class posture from the false profile namely for the standing and lying position posture where as for sitting and bending categories it obtained above 80% correct classification. It should be noted that the postures database comprises of various position for both gender without clothing restriction. This recognition rate outperformed the results obtained by Goldmann *et al.* (2004) which obtained average classification rate of 79.77% for classifying the same four main postures. Therefore, the recognition outcomes are considered remarkable. These results are promising and demonstrated the potential use of advanced correlation filters as an interesting option for posture recognition. Further work includes evaluating the

effect and performance of different image regions and using multiple filters for training images. Accordingly, incremental updating of filters could also be considered for memory reduction as space is a constraint in most recognition system. Hence only a few training images are involved at a time. The performance of other correlation filters such as Maximum Average Correlation Height (MACH), Distance Classification Correlation Filter (DCCF) filters could also be experimented for posture verification system.

ACKNOWLEDGMENTS

This research was supported by MOSTI under the IRPA Grant No: 03-02-02-0017-SR0003/07-03. The authors also acknowledge Prof. Dr. Burhanuddin Yeop Majlis as the Program Head and Universiti Teknologi Mara (UiTM) for the UiTM-JPA SLAB scholarship award.

REFERENCES

Bobick, A.F. and J.W. Davis, 2001. The recognition of human movement using temporal templates. In: Proceeding of IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23.

Chong, S.C.A., T.B. Jin and D.N.C. Ling, 2006. Iris authentication using privatized advanced correlation filter. Lecture Notes on CS, 3832, pp: 382-388.

Gader, P., J.M. Keller, T. Jones, J. Miramonti and G. Hobson, 2004. MACE Prefiltering for neural network based automatic target recognition neural networks. In: Proceeding of IEEE World Congress On Computational Intelligence, 6: 4006-4011.

Goldmann, L., M. Karaman and T. Sikora, 2004. Human body posture recognition using MPEG-7 descriptors. Proceeding of IS&T/SPIE Electronic Imaging.

Haritaoglu, I., D. Harwood and L. Davis, 1998. W4: Real-time surveillance of people and their activities. In: Proceeding of 3rd Face and Gesture Recognition Conf., pp: 222-227.

Haritaoglu, I., D. Harwood and L. Davis, 2000. W4: Who, when, where, what: A real time system for detecting and tracking people. In: Proc. IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22.

Hennings, P. and B.V.K. Vijaya Kumar, 2004. Palmprint recognition using correlation filter classifiers. In: Proceeding of Signals. Syst. Comput., 1: 567-571.

Iwasawa, S., K. Ebihara, J. Ohya and S. Morishima, 1997. Real-time estimation of human body posture from monocular thermal images. In: Proceeding of Computer Vision and Pattern Recognition.

- Jingu, H., M. Savvides and B.V.K. Vijaya Kumar, 2005. Performance evaluation of face recognition using visual and thermal imagery with advanced correlation filters. In: Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 3: 9-14.
- Muller, N. and B.M. Herbst, 2002. On the use of sdf-type filters for distortion parameter estimation. In: Proceeding of IEEE Transaction on Pattern Analysis and Machine Intelligence, 24: 1521-1528.
- Sao, A.K. and B. Yegnanarayana, 2004. Face verification using correlation filters and autoassociative neural networks. In: Proceeding of Intelligent Sensing and Information Processing, pp: 364-367.
- Savvides, M. and B.V.K. Vijaya Kumar, 2003. Efficient design of advanced correlation filters for robust distortion-tolerant face identification. In: Proceeding of IEEE International Conference on Advanced Video and Signal Based Surveillance (AVVS), pp: 45-52.
- Venkataramani, K. and B.V.K. Vijaya Kumar, 2004. Performance of composite correlation filters for fingerprint verification. *J. Optic. Eng.*, 24: 1820-1827.
- Vijaya Kumar, B.V.K., A. Mahalanobis and R.D. Juday, 2005. *Correlation Pattern Recognition*. 1st Edn. Cambridge University Press, Cambridge, UK., pp: 130-243.
- Wren, C., A. Azarbayejani, T. Darrell and A. Pentland, 1997. Pfänder: Real-time tracking of the human body. In: Proceeding of IEEE Transactions on Pattern Analysis and Machine Intelligence, 19: 780-785.