Active shape model-based user identification for an intelligent wheelchair

P. Jia* and H. Hu

School of Computer Science and Electronic Engineering,
University of Essex,
Colchester, CO4 3SQ, UK
E-mail: jp4work@gmail.com E-mail: hhu@essex.ac.uk
*Corresponding author

Abstract: Recently, some novel human-machine interfaces (HMI) have been created for disabled and elderly people to control intelligent wheelchairs (IW) using facial and head gestures. To operate a wheelchair in this new visual-based control mode, user identification should be conducted beforehand. Rather than traditional user identification that requires the user to input his/her username and password by typing, the state-of-the-art biometric-based user identification provides a more convenient way for the disabled users. This paper first elaborates active shape model in details; then, video-based user identification using Mahalanobis distance is presented. As an extension, an adaptive learning module is designed to append or update the user's face record in the constructed face database. Experimental results show that our login subsystem is able to function well for a comparatively small face database.

Keywords: intelligent wheelchair; IW; user identification; human-machine interface; HMI; active shape model; ASM.

Reference to this paper should be made as follows: Jia, P. and Hu, H. (2009) 'Active shape model-based user identification for an intelligent wheelchair', *Int. J. Advanced Mechatronic Systems*, Vol. 1, No. 4, pp.299–307.

Biographical notes: Pei Jia received his BEng from Wuhan Technical University of Surveying and Mapping (now part of Wuhan University), China in 1999 and MEng in Huazhong University of Science and Technology, China in 2003. He is currently a PhD student in School of Computer Science and Electronic Engineering, University of Essex, UK. His research interests include image processing and image analysis, pattern recognition, computer vision, robotics, etc.

Huosheng Hu received his MSc from Central South University in China and his PhD from Oxford University in the UK. He is a Professor in School of Computer Science and Electronic Engineering at the University of Essex, UK, leading the Human Centred Robotics Group. His research interests include biologically inspired robotics, service robots, human-robot interaction, evolutionary robotics, sensor integration, data fusion, artificial intelligence, embedded systems, mechatronics and pervasive computing. He has published over 280 papers in journals, books and conference proceedings in these areas. He is one of Founding Members of IEEE Robotics & Automation Society Technical Committee on Networked Robots, a fellow of IET and a Senior Member of IEEE and ACM. He has been a Chair or Committee Member for many international conferences such as IEEE ICMA, IEEE ROBIO, IEEE IROS, RoboCup Symposiums and IASTED RA, CA and CI conferences. He currently serves as one of Editors-in-Chief for *International Journal of Automation and Computing*.

1 Introduction

Advanced mechatronic systems and intelligent robots have played a key role to improve the life quality of the disabled and elderly people. In many practical applications of wheelchairs, the severely disabled users can not use their hands freely and may show clear intention by their heads instead. Therefore, head gestures have been used as a novel visual human-machine interface (HMI) in our previous work (Jia et al., 2007), in which four face directions (left, right, up and down) are recognised to control the wheelchair in a switching mode. Recently, more flexible head gesture-based HMI using 2D active appearance model (AAM) was presented in Jia and Hu (2007), which is able to tell face directions by the fitted mesh details and control the wheelchair more smoothly. Systematically, before starting controlling the wheelchair using his/her head gestures, a user needs to be authenticated in advance in order to ensure the system security.

It has been noticed for several decades that human biological information can be applied to verify different users and various kinds of biometric methods for user identification have been developed. Jain et al. (2004) summarises and compares the existing popular biometric technologies, which is modified and published online (Wikipedia). Since 2D statistical face models are adopted in Essex IW to control the wheelchair, it is natural for us to adopt the same model to identify the users in the biometrical manner. Therefore, among the existing biometric technologies, face recognition is our concern in this paper.

In the last two decades, lots of face recognition methods have been proposed, among which, eigenface has been paid tremendous attention due to its universality and favourable performance. Eigenface was first proposed by Turk and Pentland (1991a, 1991b), which is based on canonical principal component analysis (PCA). Derived from PCA-based eigenface, independent component analysis (ICA)-based eigenfaces (Bartlett et al., 2002) and linear discriminant analysis (LDA)-based eigenfaces (Etemad and Chellappa, 1997) are also proposed for face recognition. All the above mentioned methods based on eigenface could be extended to non-linear kernel methods (Yang, 2002). Moreover, elastic bunch graph matching (EBGM) was proposed to represent faces in topological graphs (Wiskott et al., 1997). By defining a 'jet' as a series of Gabor coefficients in different scales and orientations at one node, i.e., graph intersections, the face is recognised using all 'jets'.

In addition, as one of the current hottest topics in computer vision, active shape model (ASM) (Cootes and Taylor, 1992) successfully constructs face shape statistics. After ASM fitting, the fitted point coordinates can be used to recognise different persons. Unlike ASM that seeks to match only the positions of the model points, AAM tries to match both shape and texture representations of an object simultaneously (Edwards et al., 1998; Cootes and Taylor, 2004; Matthews and Baker, 2004).

As extensions, several 3D statistical models have been developed and outperformed 2D statistical models for face recognition. Blanz and Vetter (2003) proposed 3D morphable models (MM), in which up to eight points on a 2D face are manually labelled, upon which the corresponding 3D face can be synthesised for later recognition. Furthermore, Park and Jain (2007) concluded that view synthesis method is more appealing than the view-based method and declares 40% improvement in matching performance after synthesising 3D faces by 3D AAM. In Cootes and Taylor (2004), it was concluded that ASM outperforms basic AAM in terms of both accuracy and efficiency.

In this paper, ASM is revisited in detail. Shape and texture model parameters from image sequences are used to identify different users. In addition, an adaptive learning module is proposed to keep the user's face record in the face database up to date. Our demonstration shows that the designed login subsystem works well for a comparatively small face database.

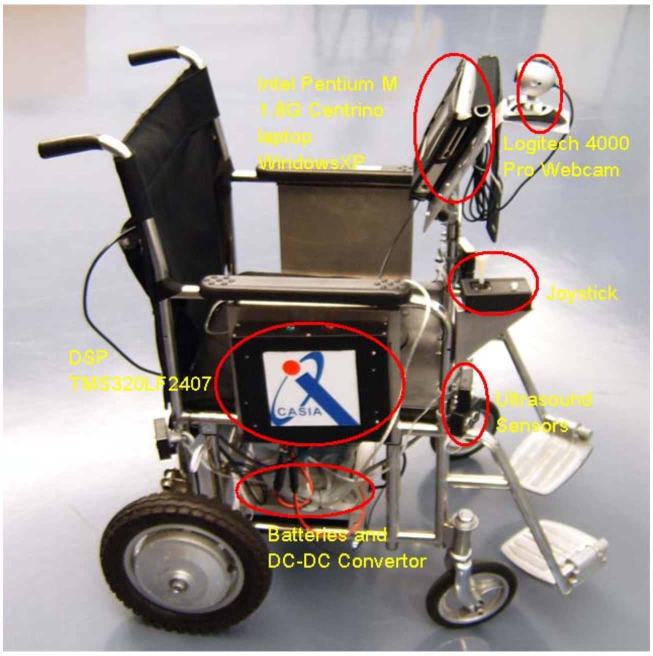
The rest of the paper is organised as follows. Section 2 introduces the system architecture used in our intelligent wheelchair (IW), as well as the user identification subsystem. In Section 3, 2D statistical models are briefly explained, including shape and texture models. Section 4 describes ASMs and the canonical fitting algorithm using 1D profile normal of the shape. In Section 5, video-based user identification using Mahalannobis distance with the adaptive learning module is addressed. Experimental results are given in Section 6 to show the performance of the proposed algorithm. Finally, a brief conclusion and future work are outlined in Section 7.

2 User identification for Essex IW

Figure 1 shows the picture of an IW at Essex, which has the following major components:

- six ultrasonic sensors at a height of about 50 centimetres used for obstacle avoidance (four at the front and two at the back)
- a joystick used to control the wheelchair manually if necessary
- a Logitech 4000 Pro Webcam used for recognising the head gesture of the user
- an Intel Pentium-M 1.6G Centrino laptop with Windows XP Operating System installed to analyse the head gesture
- a DSP TMS320LF2407 microprocessor to control two differentially-driven wheels.

Figure 1 Photo of an IW at Essex (see online version for colours)



Instead of the traditional way to login a system by using a keypad, we are developing a vision-based user identification module to be integrated into this IW, as shown in Figure 2. Essex IW will be able to identify the users by face biometrics. The same idea has already been applied in a very successful and popular product, i.e., IBM Thinkpad notebook whose login subsystem is based on finger print recognition. Figure 3 shows the framework of our face recognition-based login subsystem. Super users can only login IW using the account name and corresponding password and only administrators may know the super users' names and passwords. Ordinary users may login IW by their faces or by the traditional typing way, which means there exists a more convenient alternative for them if they prefer login without typing.

Figure 2 Essex IW control architecture (see online version for colours)

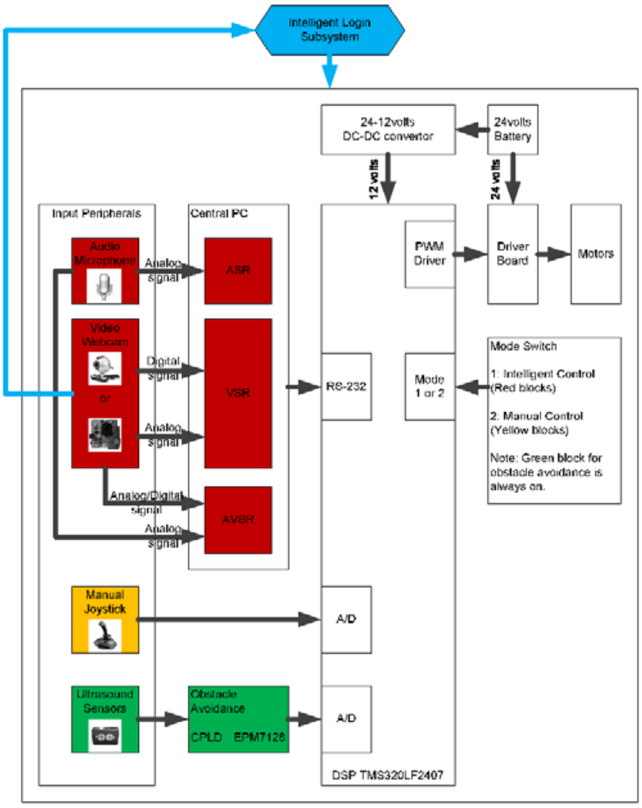
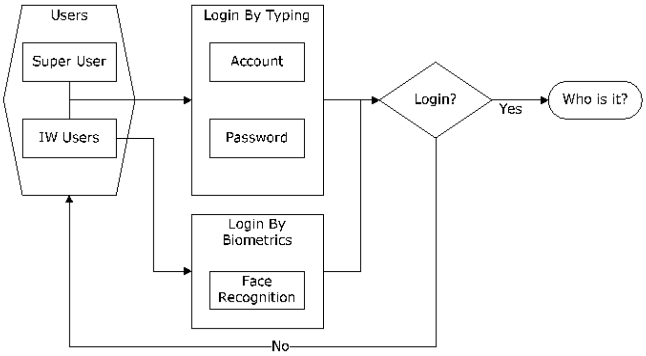


Figure 3 Login subsystem



3 2D statistical models

3.1 Shape model

The statistical model for a *shape* s is defined by the coordinates (x, y) of v manually labelled points, which actually compose a triangle mesh.

$$\mathrm{s} = [x_1, y_1, x_2, y_2, \cdots, x_v, y_v]^T \tag{1}$$

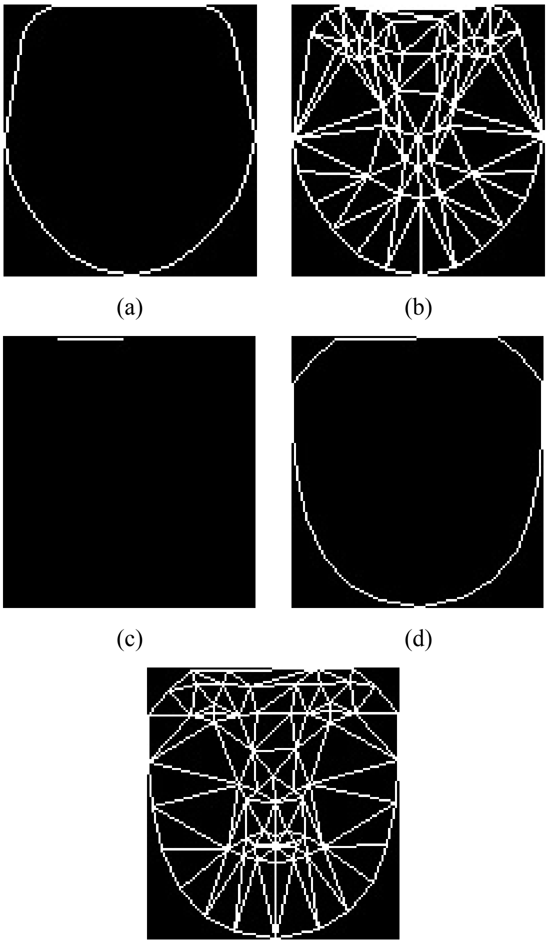
The shape statistical model allows linear shape variation. This means an arbitrary shape s is considered as a base shape s_0 plus a linear combination of N shape vectors s_i :

$$\mathrm{s} = \mathrm{s}_0 + \sum_{i=1}^{N} p_i \mathrm{s}_i = \mathrm{s}_0 + \mathbf{P}_s \mathbf{p} \tag{2}$$

where p_i are the shape parameters and s_i are orthonormal eigenvectors.

An arbitrary shape s might have various triangulation structures due to different rules. Here in our application, Delaunay triangulation (Wikipedia) is implemented on the mean shape s_0 to define the shape's inner triangulation structure, which is absolutely existing and unique.

Figure 4 Mean shape for IMM and XM2VTS (a) IMM convex hull (b) IMM triangulation (c) XM2VTS concave part (d) XM2VTS concave hull (e) XM2VTS triangulation


(a) (b)

(c) (d)

(e)

Figures 4(a) and 4(b) generate the convex hull and Delaunay triangulation of the mean shape computed from IMM face database (Nordstrøm et al., 2004). However, it is not guaranteed that the training face database will always show a convex hull, rather than a concave hull, such as Surrey XM2VTS database (Messer et al., 2004). Figures 4(c), 4(d) and 4(e) show its concave part, concave hull and

Delaunay triangulation respectively. For different face databases, the numbers of labelled points are different. There are 58 and 68 annotated points for IMM and XM2VTS respectively.

3.2 Texture model

The statistical model for a *texture* is defined in Stegmann (2000) as the pixel intensities across the object (here, the face) in question (if necessary after a suitable normalisation). Then, the texture statistical model is a vector of coordinate-related intensities $\mathbf{A}(\mathbf{x})$ defined over all u pixels inside the base mesh s_0, i.e., $\mathbf{x} \in s_0$, where $\mathbf{x} = (x, y)^T$. Equation (3) shows a texture vector when RGB channels or a single channel is considered respectively:

$$\begin{aligned}\mathbf{A}(\mathbf{x}) &= [b_1, g_1, r_1, b_2, g_2, r_2, \cdots, b_u, g_u, r_u]^T \\ \mathbf{A}(\mathbf{x}) &= [gray_1, gray_2, \cdots, gray_u]^T\end{aligned} \tag{3}$$

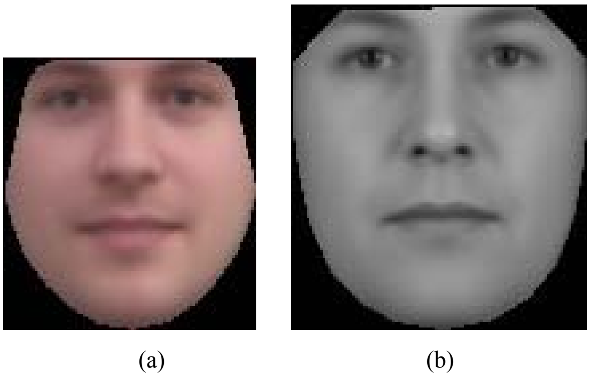
The texture statistical model allows linear texture variation. This means an arbitrary texture $\mathbf{A}(\mathbf{x})$ is looked on as a base texture $\mathbf{A}_0(\mathbf{x})$ plus a linear combination of M textures $\mathbf{A}_i(\mathbf{x})$:

$$\mathbf{A}(\mathbf{x}) = \mathbf{A}_0 + \sum_{i=1}^{M} \lambda_i \mathbf{A}_i(\mathbf{x}) = \mathbf{A}_0(\mathbf{x}) + \mathbf{P}_\mathbf{t}\lambda, \forall \mathbf{x} \in s_0 \tag{4}$$

where λ_i are the texture parameters and $\mathbf{A}_i(\mathbf{x})$ are orthonormal eigenvectors.

Figure 5 shows the mean face textures synthesised from IMM and XM2VTS face databases. For IMM, three channels are used; for XM2VTS, the texture is synthesised from grey-level images.

Figure 5 Mean texture (a) IMM (b) XM2VTS (see online version for colours)



(a) (b)

4 Active shape model

4.1 Canonical 1D profile model of ASM

Canonical ASM fitting algorithm computes the normal to the profile at each model point. At every given point, the tangential line could be defined by two line segments:

1. the segment connecting this point and the point it connected from
2. the segment connecting this point and the point it connected to.

The normal direction to the profile is then defined by normalising the sum of the two unit normal vectors of these two line segments. Figure 6 shows all the profile normals at the manually labelled 87 points of one sample face from our self-labelled face database.

Figure 6 Profile normal vector

Along the normal direction, at each side of the model point i, k pixels' derivatives $\mathbf{g}_i$ (rather than the original grey-level values) are sampled and then normalised as the feature set for this specific key point.

$$\mathbf{g}_i \rightarrow \frac{1}{\sum_{j=-k}^{k} \left| g_{ij} \right|} \mathbf{g}_i \tag{5}$$

For every given model point of each training image, a normalised unit vector of derivatives can be extracted. Assuming a multivariate Gaussian, the quality of fitting a new face to the model at this specific point could be evaluated by Mahalanobis distance:

$$f(\mathbf{g}_{new}) = \sqrt{(\mathbf{g}_{new} - \overline{\mathbf{g}})^T \mathbf{S}_g^{-1} (\mathbf{g}_{new} - \overline{\mathbf{g}})} \tag{6}$$

where $\overline{\mathbf{g}}$ and $\mathbf{S}_g^{-1}$ are respectively the trained mean vector and the inverse of covariance matrix in terms of this specific point.

4.2 ASM fitting algorithm

An intuitive idea to find those key points on an arbitrary face based on ASM 1D profile model could be summarised as:

1. Approximately locate the face by Adaboost face detection (Viola and Jones, 2001, 2004) first.

2 At each side of the model point i, along its profile normal, m pixels' derivatives $m > k$ are sampled. We then calculate the fit quality at each of the $2(m-k)+1$ possible slots along the normal direction and choose the best match as the updated position for this model point. This process repeats for every model point and gives out a modification on it for the next iteration.

3 After all points have been updated, constrain the shape by shape model parameters. Apparently, the current constructed shape is corresponding to a vector of shape model parameters $\mathbf{p}_i, 0 \le i < N.$ With the built statistical shape model, each of above model parameters is restricted within 3σ around its mean value.

4 Repeat 2 and 3 until convergence.

To speed up the search process and improve the robustness of this algorithm, multiscale ASM is adopted in our experiments. This involves searching for the face in a coarse image first and then refining the face key points' locations in a series of finer resolution images.

5 User identification

ASM can fit a face accurately and generate a triangulation mesh, through which the face texture could also be extracted. Edwards et al. (1998) successfully carried out face recognition through the shape and texture model parameters by calculating Mahalanobis distance. For Essex IW, the video-based user identification algorithm using Mahalanobis distance is revisited and an adaptive learning module is designed to update the user face record in the constructed face database.

5.1 Construction of prior user face database

The user face database is constructed with the prior knowledge that we can obtain from all available users, including the usernames, passwords and the face biometric data – here, 2D shape and texture statistics. For each user, N frames of frontal faces are automatically selected from an image sequence to compute his/her personal face statistic information.

For each frame in the image sequence, ASM 1D profile model is applied to calculate the best fitted face shape first. After two PCA transforms on the fitted shape and the corresponding extracted texture, the statistical shape and texture parameters can be obtained for this single image frame. Mean shape and texture parameters and the covariance matrixes can be calculated over all N images.

In summary, each user record in the face database can be constructed in the following steps.

1 Start image capturing and track the user's face by ASM 1D profile model.

2 If the face is well fitted (during database construction process, this could be judged by the user himself/herself), hold the same posture with limited movements for around one second (In our experiments, $N = 16.$ If the webcam captures 30 frames per second, the time of capturing 16 images will be 16/30 = 0.53 second. Considering the fitting time for our algorithm, the time of fitting 16 images will be at least one second).

3 Once the button 'add a user' is activated, N dynamic (real-time updating) images are available to construct the user's face record.

4 With N fitted shapes and extracted textures, N pairs of shape parameters and texture parameters can be calculated and respectively denoted as $\mathbf{p}_i, 0 \le i < N$ and $\lambda_i, 0 \le i < N.$

5 Then, the mean shape parameters and texture parameters for this person and the respective covariance matrixes are computed as:

$$\mathrm{E}[\mathbf{p}] = \frac{\sum_{i=1}^{N} \mathbf{p}_i}{N} \tag{7}$$

$$\mathrm{E}[\lambda] = \frac{\sum_{i=1}^{N} \lambda_i}{N} \tag{8}$$

$$\Sigma[\mathbf{p}] = \mathrm{E}\left[(\mathbf{p} - \mathrm{E}[\mathbf{p}])(\mathbf{p} - \mathrm{E}[\mathbf{p}])^{\mathrm{T}}\right] \tag{9}$$

$$\Sigma[\lambda] = \mathrm{E}\left[(\lambda - \mathrm{E}[\lambda])(\lambda - \mathrm{E}[\lambda])^{\mathrm{T}}\right] \tag{10}$$

6 Store all the above statistical data calculated in Step 5 into the database as a record field of the user's face.

It is possible to add as many face records as possible into the face database. However, one IW will be normally used by a single person. Besides, in order to increase the user identification accuracy, it is better to restrict the number of subjects in the face database. Here, three users, including one female and two males participated in our experiments.

5.2 Video-based user identification using Mahalanobis distance

After the face database is constructed, it is the time to verify whether the IW face identification subsystem is applicable. It is expected that the login face could be matched to the corresponding user's record in the face database. By looking on the user identification problem as a classification one, canonical Mahalanobis distance is able to be used to classify different users. Theoretically, to classify a testing sample to the class with the minimal Mahalanobis distance is equivalent to selecting the class with the highest probability (Wikipedia). Meanwhile, Mahalanobis distance could favourably deal with 'outliers' (Wikipedia) if the testing sample does not obey the distribution, which means, it can handle the 'belong to none' situation.

Algorithm 1 Voting based on Mahalanobis distance

```
Pre-computation during training
  • With the training $N$ frames for each person, calculate the mean shape and texture parameters and respective covariance matrixes for all $C$ available users and store them in database: $\mathrm{E}_j[\mathbf{p}], \mathrm{E}_j[\lambda]$, $\Sigma_j[\mathbf{p}]$ and $\Sigma_j[\lambda]$, where $0 \le j < C$
Initialisation
  • $WeakFitting = false$
  • $PGood = 0.75$
for $j = 0$ to $C$ do
    $Votes[j] = 0$
end for
for $i = 0$ to $N$ do
    Obtain current fitted face shape parameters $\mathbf{p}_i$ and texture parameters $\lambda_i$
    for $j = 0$ to $C$ do
        Calculate the Mahalanobis distances $sDist[j]$ from the current fitted face shape parameters $\mathbf{p}_i$ to the stored shape record of person $j$ in the face database
        $sDist[j] = \sqrt{(\mathbf{p}_i - \mathrm{E}^j[\mathbf{p}])^T \Sigma^j[\mathbf{p}]^{-1} (\mathbf{p}_i - \mathrm{E}^j[\mathbf{p}])}$
    end for
    Pick up the record index $k$ with the shortest Mahalanobis distance in terms of shape $sDist[k]$, where $0 \le k < C$
    if $sDist[k]$ is not outlier then
        if $(!WeakFitting)$ then
            Calculate the Mahalanobis distances $tDist$ from the current fitted face texture parameters $\lambda_i$ to the stored texture records of person $k$ in the face database
            $tDist = \sqrt{(\lambda_i - \mathrm{E}^k[\lambda])^T \Sigma^k[\lambda]^{-1} (\lambda_i - \mathrm{E}^k[\lambda])}$
            if $tDist$ is not outlier then
                $Votes[k]++;$
            end if
        else
            $Votes[k]++;$
        end if
    end if
end for
Pick up the record index $j$ with the most votes, that is, the biggest value of $Votes[j]$, where $0 \le j < C$
if $(Votes[j] \ge PGood * N)$ then
    return $j$;
else
    return –1;
end if
```

During user login procedure, N pairs of shape and texture parameters for N qualified images are calculated and respectively denoted as $\mathbf{p}_i, 0 \le i < N$ and $\lambda_i, 0 \le i < N$. N Mahalanobis distances in terms of the shape and texture parameters are then calculated as:

$$D_M(\mathbf{p}_i) = \sqrt{(\mathbf{p}_i - \mathrm{E}[\mathbf{p}])^T \Sigma[\mathbf{p}]^{-1} (\mathbf{p}_i - \mathrm{E}[\mathbf{p}])} \tag{11}$$

$$D_M(\lambda_i) = \sqrt{(\lambda_i - \mathrm{E}[\lambda])^T \Sigma[\lambda]^{-1} (\lambda_i - \mathrm{E}[\lambda])} \tag{12}$$

For simplicity, a voting strategy for the above image sequence is adopted to identify the users as shown in Algorithm 1.

5.3 Adaptive learning for user records' updating

Every time when a user succeeds in the system login based on his or her face, additional biometric knowledge is deployed at the same time. In order to keep the user's face statistics up to date, the newly input knowledge should be added into the face database whenever a user succeeds in login. Suppose that the newly collected statistics from N qualified frames in an image sequence are denoted as $\mathrm{E}_n[\mathbf{p}], \mathrm{E}_n[\lambda], \Sigma_n[\mathbf{p}], \Sigma_n[\lambda]$, then, the updated statistics $\mathrm{E}_u[\mathbf{p}], \mathrm{E}_u[\lambda], \Sigma_u[\mathbf{p}], \Sigma_u[\lambda]$ could be calculated as follows:

$$\mathrm{E}_u[\mathbf{p}] = \frac{\mathrm{E}[\mathbf{p}] + \mathrm{E}_n[\mathbf{p}]}{2} \tag{13}$$

$$\mathrm{E}_u[\lambda] = \frac{\mathrm{E}[\lambda] + \mathrm{E}_n[\lambda]}{2} \tag{14}$$

$$\Sigma_u[\mathbf{p}] = \frac{\Sigma[\mathbf{p}] + \Sigma_n[\mathbf{p}] + \Sigma_m[\mathbf{p}] + \Sigma_{nm}[\mathbf{p}]}{2} \tag{15}$$

$$\Sigma_u[\lambda] = \frac{\Sigma[\lambda] + \Sigma_n[\lambda] + \Sigma_m[\lambda] + \Sigma_{nm}[\lambda]}{2} \tag{16}$$

where

$$\Sigma_m[\mathbf{p}] = \mathrm{E}\left[(\mathrm{E}[\mathbf{p}] - \mathrm{E}_u[\mathbf{p}])(\mathrm{E}[\mathbf{p}] - \mathrm{E}_u[\mathbf{p}])^T\right] \tag{17}$$

$$\Sigma_{nm}[\mathbf{p}] = \mathrm{E}\left[(\mathrm{E}_n[\mathbf{p}] - \mathrm{E}_u[\mathbf{p}])(\mathrm{E}_n[\mathbf{p}] - \mathrm{E}_u[\mathbf{p}])^T\right] \tag{18}$$

$$\Sigma_m[\lambda] = \mathrm{E}\left[(\mathrm{E}[\lambda] - \mathrm{E}_u[\lambda])(\mathrm{E}[\lambda] - \mathrm{E}_u[\lambda])^T\right] \tag{19}$$

$$\Sigma_{nm}[\lambda] = \mathrm{E}\left[(\mathrm{E}_n[\lambda] - \mathrm{E}_u[\lambda])(\mathrm{E}_n[\lambda] - \mathrm{E}_u[\lambda])^T\right] \tag{20}$$

6 Experimental results

6.1 Fitting results for single frontal images

We first conduct the ASM fitting experiments on all images of the two mentioned face databases IMM and XM2VTS. In Figure 7, the first row contains ASM fitting results for three image samples from IMM database; the second row shows ASM fitting results for two images from XM2VTS database. From the samples here, we can see that ASM can fit female and male faces from various races, even can deal with faces with glasses or beard to some extend.

To evaluate the fitting results numerically, two criteria are established as:

1 if more than 80% of the fitted points are within a coordinate tolerance t to the manually labelled points in one image, this face is well fitted

2 if criterion 1 is not satisfied, but the total deviation for all the key points is less than $(t+1)/\sqrt{N}$, this face is well fitted.

Figure 7 ASM fitting effects (see online version for colours)

We can see from Table 1 that no matter whether evaluated from numerical aspect or evaluated by human eyes, ASM fitting results are reasonably well. Full fitting results can be found from the personal website of the first author at http://www.visionopen.com/cv/aam.php

Table 1 Fitting results

Database		*IMM*	*XM2VTS*
Number of images		37*	2,360
ASM	Numerical human	94.59%	85.89%
	Eye	100%	88.81%

Note: *A subset of all 37 frontal faces from IMM is tested here, rather than the full set of 240 frontal and profile faces.

6.2 Fitting results for an image sequence

Six images out of an image sequence fitted by ASM are given in Figure 8, which demonstrates that ASM is good at tracking faces and telling face details in real-time. We did obtain a real-time tracking, more than 20 FPS using the ASM fitting algorithm. In this experiment, a self-labelled face database with 87 annotated points for each face is trained and adopted.

Figure 8 Face tracking sequence fitted by ASM (see online version for colours)

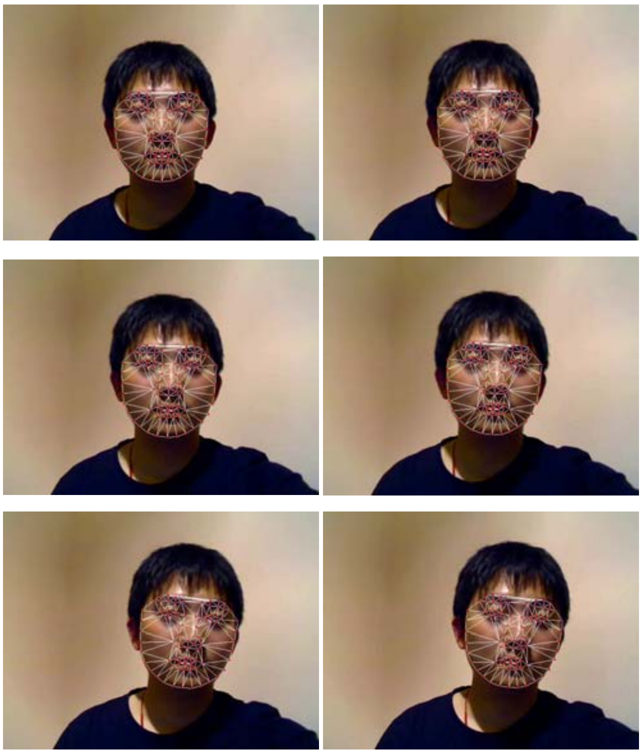


Figure 9 User face identification, (a) login dialog (b) status bar of IW application dialog with user's account (see online version for colours)

AAMFitting 1.0

Account: Empty

Password: Empty

User Image

(a)

suzhen,Thank you for choosing AAMFitting GUI Application!

(b)

6.3 GUI application for user face identification

Figure 9 shows the self-designed GUI application for the login subsystem of Essex IW. It is clear that the proposed user identification method works well.

7 Conclusions and future work

Traditional electric-powered wheelchairs are normally controlled by users via joysticks, which cannot satisfy the needs of elderly and disabled users who have restricted limb movements caused by diseases such as Parkinson and quadriplegia. Therefore, we have recently created some novel HMIs for hands-free control of our IW. In this paper, a video-based user identification algorithm using Mahalanobis distance is proposed to ensure the integrity and security of our IW after 1D profile ASM is thoroughly revisited. This user identification mechanism is to be used in our IW for user login. Some experimental results show that the video-based user identification algorithm performs reasonably well if the stored user face records always keep updated by the adaptive learning algorithm that we have proposed.

Our future work will be using ASM to accurately recognise the face directions and test it on the real IW for the control purpose.

Acknowledgements

We would like to thank Emount Ltd., China (http://www.emount.com.cn/) for their generous contribution to the Chinese face database with 87 labelled points for each face.

References

Bartlett, M.S., Movellan, J.R. and Sejnowski, T.J. (2002) 'Face recognition by independent component analysis', *IEEE Transactions on Neural Networks*, Vol. 13, No. 6, pp.1450–1464.

Blanz, V. and Vetter, T. (2003) 'Face recognition based on fitting a 3D morphable model', *IEEE Transcations on Pattern Analysis and Machine Intelligence*, September, Vol. 25, No. 9, pp.1–12.

Cootes, T. and Taylor, C. (1992) 'Active shape models – smart snakes', *Proceedings of British Machine Vision Conference (BMVC 1992)*, Leeds, UK, pp.266–275.

Cootes, T. and Taylor, C. (2004) 'Statistical models of appearance for computer vision', Technical Report on Active Shape Models and Active Appearance Models, University of Manchester, available at http://www.isbe.man.ac.uk/bim/Models/app models.pdf.

Edwards, G., Cootes, T. and Taylor, C. (1998) 'Face recognition using active appearance models', *Proceedings of 5th European Conference on Computer Vision*, LNCS-Series 1406–1607, pp.581–595.

Etemad, K. and Chellappa, R. (1997) 'Discriminant analysis for recognition of human face images', *Journal of the Optical Society of America A*, August, Vol. 14, No. 8, pp.1724–1733.

Jain, A.K., Ross, A. and Prabhakar, S. (2004) 'An introduction to biometric recognition', *IEEE Transactions on Circuits and Systems for Video Technology*, January, Vol. 14, No. 1, pp.4–20.

Jia, P. and Hu, H. (2007) 'AAM-based HCI for an intelligent wheelchair', *MIPPR 2007: Remote Sensing and GIS Data Processing and Applications; and Innovative Multispectral Technology and Applications, Proceedings of The International Society for Optical Engineering (SPIE)*, Wuhan, Hubei, China, November, Vol. 6790, 67904F.

Jia, P., Hu, H., Lu, T. and Yuan, K. (2007) 'Head gesture recognition for hands-free control of an intelligent wheelchair', *Journal of Industrial Robot*, April, Vol. 34, No. 1, pp.60–68.

Matthews, I. and Baker, S. (2004) 'Active appearance models revisited', *International Journal of Computer Vision*, November, Vol. 60, No. 2, pp.135–164.

Messer, K., Matas, J., Kittler, J., LÄuttin, J. and Maitre, G. (2004) 'XM2VTSDB: the extended M2VTS database', *Audio- and Video-based Biometric Person Authentication, AVBPA'99*, Washington DC, USA, March, pp.72–77.

Nordstrøm, M.M., Larsen, M., Sierakowski, J. and Stegmann, M.B. (2004) 'The IMM face database – an annotated dataset of 240 face images', Technical Report of Informatics and Mathematical Modeling, Technical University of Denmark, May.

Park, U. and Jain, A.K. (2007) '3D model-based face recognition in video', *The 2nd International Conference on Biometrics*, Seoul, South Korea, August, pp.1085–10941.

Stegmann, M.B. (2000) 'Active appearance models: theory, extensions and cases', Technical Report of Informatics and Mathematical Modeling, Technical University of Denmark, September.

Turk, M. and Pentland, A. (1991a) 'Eigenfaces for recognition', *Journal of Cognitive Neurosicence*, Winter, Vol. 3, No. 1, pp.71–86.

Turk, M. and Pentland, A. (1991b) 'Face recognition using eigenfaces', *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, June, pp.586–591.

Viola, P. and Jones, M.J. (2001) 'Rapid object detection using a boosted cascade of simple features', *Proceedings of IEEE Computer Society's Computer Vision and Pattern Recognition*, Vol. 1, pp.511–518.

Viola, P. and Jones, M.J. (2004) 'Robust real-time face detection', *International Journal of Computer Vision*, Vol. 57, No. 2, pp.137–154.

Wikipedia, *Biometrics*, Wikipedia, The Free Encyclopedia, available at http://en.wikipedia.org/wiki/Biometrics.

Wikipedia, *Delaunay Triangulation*, Wikipedia, The Free Encyclopedia, available at http://en.wikipedia.org/wiki/Delaunay triangulation.

Wikipedia, *Mahalanobis Distance*, Wikipedia, The Free Encyclopedia, available at http://en.wikipedia.org/wiki/Mahalanobis distance.

Wikipedia, *Outlier*, Wikipedia, The Free Encyclopedia, available at http://en.wikipedia.org/wiki/Outlier.

Wiskott, L., Fellous, J-M., KrÄuger, N. and von der Malsburg, C. (1997) 'Face recognition by elastic bunch graph matching', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, July, Vol. 19, No. 7, pp.775–779.

Yang, M-H. (2002) 'Kernel eigenfaces vs. kernel fisherfaces: face recognition using kernel methods', *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, Washington DC, USA, May, pp.215–220.