Mapping 2D to 3D Forensic Facial Recognition via Bio-Inspired Active Appearance Model

Siti Zaharah Abd. Rahman[a], Siti Norul Huda Sheikh Abdullah[a*], Lim Eng Hao[a], Mohammed Hasan Abdulameer[a], Nazri Ahmad Zamani[b], Mohammad Zaharudin A. Darus[b]

[a]Pattern Recognition Research Group, Center of Artificial Intelligent Technology, Faculty of Information Science and Technology, National University of Malaysia, Bangi, Selangor, Malaysia
[b]Digital Forensic Department, CyberSecurity Malaysia, Seri Kembangan, Selangor, Malaysia

Graphical abstract
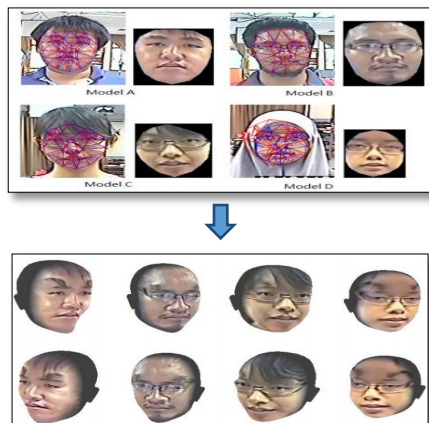
Abstract

This research done is to solve the problems faced by digital forensic analysts in identifying a suspect captured on their CCTV. Identifying the suspect through the CCTV video footage is a very challenging task for them as it involves tedious rounds of processes to match the facial information in the video footage to a set of suspect's images. The biggest problem faced by digital forensic analysis is modeling 2D model extracted from CCTV video as the model does not provide enough information to carry out the identification process. Problems occur when a suspect in the video is not facing the camera, the image extracted is the side image of the suspect and it is difficult to make a matching with portrait image in the database. There are also many factors that contribute to the process of extracting facial information from a video to be difficult, such as low-quality video. Through 2D to 3D image model mapping, any partial face information that is incomplete can be matched more efficiently with 3D data by rotating it to matched position. The first methodology in this research is data collection; any data obtained through video recorder. Then, the video will be converted into an image. Images are used to develop the Active Appearance Model (the 2D face model is AAM) 2D and AAM 3D. AAM is used as an input for learning and testing process involving three classifiers, which are Random Forest, Support Vector Machine (SVM), and Neural Networks classifier. The experimental results show that the 3D model is more suitable for use in face recognition as the percentage of the recognition is higher compared with the 2D model.

Keywords: Mapping, forensic facial recognition, AAM

Abstrak

Pemetaan 2D kepada 3D pengecaman muka dikaji untuk menyelesaikan masalah yang dihadapi oleh penganalisis digital forensik dalam mengenal pasti suspek melalui rakaman video CCTV. Pengecaman seseorang suspek dalam rakaman video CCTV adalah satu tugas yang amat mencabar bagi penganalisis forensik kerana melibatkan proses yang kompleks iaitu dengan memadankan maklumat muka dalam rakaman video kepada satu set imej suspek. Masalah utama bagi analisis ini adalah 2D model yang diekstrak dari video CCTV tidak mempunyai maklumat yang cukup untuk menjalankan

pengecaman. Sebagai contoh, dalam rakaman video CCTV, suspek tidak menghadap ke arah kamera, oleh itu, imej yang dapat diekstrakk ialah imej sisi bagi suspek di mana imej tersebut adalah tidak sesuai untuk membuat pemadanan dengan imej dalam pangkalan data iaitu imej potret. Oleh itu, kajian meneruskan dalam pemetaan 2D kepada 3D bagi pemadanan muka. Melalui kajian ini, sebarang maklumat muka yang tidak lengkap boleh dipadankan dengan lebih cekap dengan data 3D. Kaedah yang terlibat alam projek ini bermula dengam pengumpulan data, iaitu merakam pergerakan muka individu dan selepas itu, video tersebut akan diproses untuk menukarkan video kepada imej. Imej yang diperolehi daripada proses ini digunakan untuk membangunkan *Active Appearance Model* (AAM) 2D dan juga AAM 3D dan digunakan sebagai input bagi tiga klasifier iaitu *Random Forest*, *Support Vector Machine*, dan Rangkaian Neural diguna untuk membuat pembelajaran dan pengujian. Keputusan yang diperolehi menunjukkan model 3D lebih sesuai digunakan dalam pengecaman muka berbanding dengan model 2D kerana ia memberi kejituan yang tinggi.

Kata kunci: Pemetaan, pengecaman muka forensic, AAM

1.0 INTRODUCTION

Forensic Science is a very huge research area. The examples of forensic science are firearms calculation, license plate recognition, suspicious behavior recognition, motion detection whereas digital forensics only emphasizes on digital imaging. It covers various aspects of determining criminal or suspect evidences through digital imaging [1] [2]. Face recognition system is a computer program that works to identify an individual automatically. It works using some facial features in the image of an images in the database. Face recognition system is used as a reliable security protection. Face recognition systems can be compared with biometric security systems available such as retina scanners, scanner fingerprints and more [3]. In Digital Forensics, face recognition is important because it can help in providing new evidence in court. In the court, this process is very tedious and requires detail process explanation once the source image has been enhanced or modified inside out.

At regular cases, 2D images which have been captured in advance or kept unattended are being used and manipulated legally to justify any serious cases in order to find the truth of judgment. These procedures are normally done manually and taking ages to complete and verify the evidence. It also plays an important role in the daily life of people, specifically in the identification and authentication of individual identity. In this paper, we used artificial bee colony (ABC) algorithm for solving the fitting problem in AAM [4] use in mapping process. The ABC algorithm simulates the intelligent foraging behavior of a honey bee swarm. In spite of that, we have not directly used the conventional ABC algorithm for the research, due to its limitation in its neighborhood search for generating new food sources. However, this issue has been addressed by introducing an adaptive ABC algorithm to solve the fitting problem.

This research aims to mapping the 2D to 3D face recognition. The first step of this research is the collection of data recorded by the video. Next, the recorded video will go through the process of extracting the image to get facial image of an individual. After successfully extracting the images, the images will be used to develop 2D AAM. Next, the developed 2D AAM will be used to develop 3D AAM and the data model is going through the process scaling before classification. In the process of classification, the algorithms used are the Random Forest, Neural Network and Support Vector Machine (SVM).

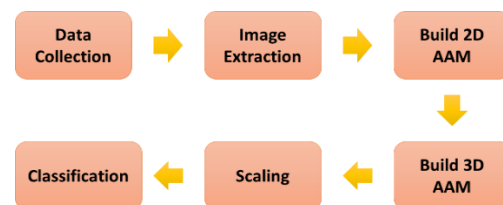Figure 1 below show the overall step conducted in this research.

Figure 1 Step on mapping 2D to 3D facial recognition

1.1 Developing 2D Interface Model and Algorithm AAM

The Type of 2D model used in this research is Active Appearance Model (AAM). For selected images, the image will be given a landmark point manually. Algorithms to develop AAM are shown below.

Given a set of training images, I_i (x,y) ; $x = 0,1, …, M-1$, $y=0,1,…,N-1$ and $i=0,1,…,N_I – 1$ where I_i (x,y) is of size $M \times N$. In the training images, the active portions are manually labeled to extract the shape model parameters and the appearance model parameters. A vector X_{ij} and Y_{ij}: $j= 1,2,…,N_P$ are generated by x and y co-ordinates . Let N_P be active portion of an image so that the vectors X and Y would follow the below criteria:

$$|X|ij = |Y|ik : j = k \qquad (1)$$

$$|X|_j = (or) \neq |Y|_k : j \neq k \text{ and } k = 1,2,\ldots, Np \qquad (2)$$

The grey portions of I_i are extracted using the X_i and Y_i as follows:

$$G_{ij} = \begin{cases} G_i(x,y); & if \; x \leq X_{ij}(k) \;\; y \leq Y_{ij}(k) \\ 0; & otherwise \end{cases} \qquad (3)$$

Where

$$g_i(x,y) = 0.3I_i^R(x,y) + 0.6I_i^G(x,y) + 0.1I_i^B(X,y) \qquad (4)$$

In Eq. (4) R, G and B indicate color spaces (R-Red, G-Green, B-Blue). Then, appropriate normalization is applied over X, Y, and G as follows:

$$\overline{X_j} = \frac{1}{N_I} \sum_{i=0}^{N_I-1} X_{ij} \qquad (5)$$

$$\overline{Y_j} = \frac{1}{N_I} \sum_{i=0}^{N_I-1} Y_{ij} \qquad (6)$$

$$\overline{X_j} = \frac{1}{N_I} \sum_{i=0}^{N_I-1} X_{ij} \qquad (7)$$

Such that,

$$\overline{X_j} = \overline{Y_j} \qquad (8)$$

From X, Y and G, the shape and grey parameters are determined as follows:

$$A_{ij} = \begin{bmatrix} S_p^{(ij)} w_{ij} \\ G_p^{(ij)} \end{bmatrix} \qquad (9)$$

$$S_{(p)}^{(ij)} = \begin{bmatrix} (\overline{X_J} - X_{ij})\xi_j^x \\ (\overline{Y_j} - Y_{ij})\xi_j^y \end{bmatrix} \qquad (10)$$

$$G_p^{(ij)} = (\overline{G_j} - G_{ij})\xi_j^g \qquad (11)$$

The obtained shape and grey parameters are already subjected to decomposition to generate a vector of the appearance model parameter as follows:

$$A_{ij} = Q_{ij}a_{ij} \qquad (12)$$

Where Q_{ij}, is eigenvectors and a_{ij} is vector of appearance parameters, respectively. In the process of iterations for development of 2D AAM, AAM will be created based on the point of the original features. For the first iteration, AAM is created and no matching is made. Starting with the second iteration, AAM start matching the face image. In this iteration, the created model moves to the right, to get the best position for the face in the image. AAM gets the matching face on the eighth iteration. After obtaining a good position on the model, AAM begins to change the contour to get the best shape for the face by using Adaptive Bees Algorithm [5]. In iteration 16th shown in Figure 2, the contour of the model is for landmark features in the model. The Matching of AAM is done on the 20th iterations.

Figure 2 Iteration for AAM Development

1.2 Adaptive Bees Algorithm

Swarm intelligence is a research field that models the collective intelligence in swarms of insects or animals. Many algorithms that simulate these models have been proposed in order to solve a wide range of problems. The Artificial Bee Colony algorithm is one of the most recent swarm intelligence based algorithms that simulates the foraging behavior of honeybee colonies. It is also applied on multimodal and multi-dimensional numerical optimization problems [6]. In ABC algorithm, an artificial bee colony consists of employed bees, onlookers and scouts. A bee waiting on the dance area to obtain the information about food sources is called an onlooker, a bee going to the food source is named as an employed bee, and a bee carrying out random search is called a scout. The position of a food source denotes a possible solution to the optimization problem, and the nectar amount of a food source represents the quality of the associated solution. Initially, a randomly distributed population is generated. For every food source, there is only one employed bee. So the number of employed bees is equal to the number of food sources. Thereafter, the positions (solutions) will be updated repeatedly with the following cycles until the maximum iteration is reached or stop conditions are satisfied. Each employed bee always remembers its previous best position, and produces a new position within its neighborhood in its memory. According to the greedy criterion, the employed bee updates its food source. In other words, when the new food source is better, the old food source position is updated with the new one. After all employed bees finish their search process, they share the information about the direction and distance to the food sources and the nectar amounts with onlookers via a so-called waggle dance in the dancing area.

From the observation on the waggle dance, each onlooker chooses a food source depending on the probability value associated with the food source, and searches the area within its neighborhood to generate a new candidate solution. And then, the greedy criterion is applied again just as in the employed bees. If a position cannot be improved after a predetermined number of cycles, the position should be abandoned; meanwhile, the corresponding employed bee becomes a scout. The abandoned position will be replaced with a new randomly generated food source. As mentioned above, the fitness function is a key component of ABC algorithm, which evaluates the foraging quality of the colony, i.e. the accuracy of possible solutions. Besides, some control parameters, such as the number of employed

bees or onlooker bees, the limit times for abandonment, the maximum number of iterations or stop conditions, need to be assigned. They would have a direct influence on the speed and stability of convergence.

1.3 Developing 3D Interface Model

A set of images of a face or more will use I^t (∪) where $t = 0, ... , N$ matched to 2D AAM for each image and shape parameters for 2D AAM shape in the image I^t (∪) is $P^t = (P_1^t, ..., P_m^t)^T$. This formula, $s = s_0 + \sum_i^m = 1 p_i s_i$ is used to calculate a 2D AAM vector shape of S^t :

$$S^t = \begin{pmatrix} u_1^t\, u_2^t\, ...\, u_n^t \\ v_1^t\, v_1^t\, ...\, v_n^t \end{pmatrix}, t = 0, ..., N \qquad (13)$$

$$W = \begin{pmatrix} u_1^0\, u_2^0\, ...\, u_n^0 \\ v_1^0\, v_2^0\, ...\, v_n^0 \\ \vdots \\ u_1^N\, u_2^N\, ...\, u_n^N \\ v_1^N\, v_2^N\, ...\, v_n^N \end{pmatrix} \qquad (14)$$

$$W\ -MB-\ \begin{pmatrix} \mathbf{P}^0 P_1^0\, ...\, P_{\frac{0}{m}}\mathbf{P}^0 \\ \mathbf{P}^1 P_1^1\, ...\, P_{\frac{1}{m}}\mathbf{P}^1 \\ \vdots \\ \mathbf{P}^N P_1^N\, ...\, P_{\frac{N}{m}}\mathbf{P}^N \end{pmatrix} \begin{pmatrix} \bar{S}_0 \\ \vdots \\ \bar{S}_{\bar{m}} \end{pmatrix} \qquad (15)$$

Where $P^{i,}$ I =0,...., N is 2x3 projection matrix,

$$p^i = \begin{pmatrix} i_x & i_y & i_z \\ j_x & j_y & j_z \end{pmatrix}$$

Stated that M is 2(N +1) x 3(m+1) matrix, B is 3(m+1) x n, and m is the number of 2D vector form any $3(\bar{m} + 1)$ x $3(\bar{m} + 1)$ nonsingular matrix G with inverse G^{-1} can be inserted between $\tilde{M}$ and $\tilde{B}$ product yield remained W. Scaled projection matrix M and the matrix model form is:

$$M = \bar{M} . G, B = G^{-1}.\bar{B} \qquad (16)$$

The algorithm below shows on how to develop a 3D model from a 2D model in programming language. Eq. 13 and Eq. 14 use the vector shape model constructed by 2D AAM and compiled with scaled back projection matrix. Some calculations were made before finally able to develop a 3D face model [7]. Figure 3 shows the created 3D AAM model and Figure 4 shows the reconstructed 3D AAM face model.
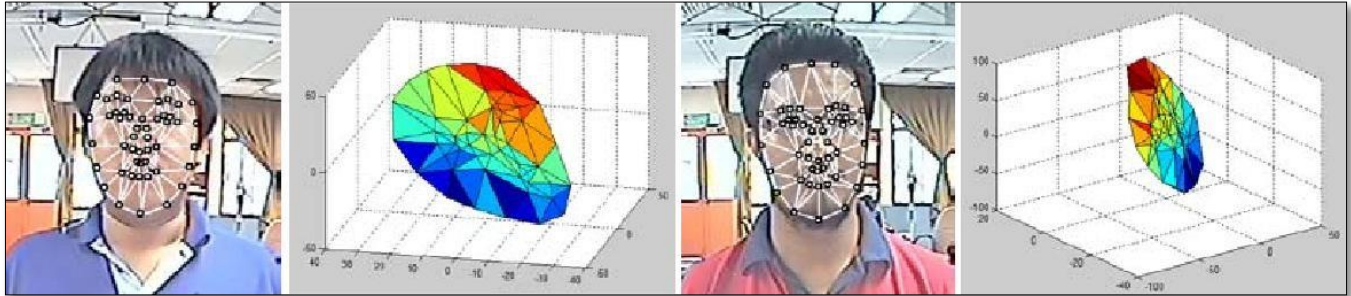
Figure 3 3D AAM

Figure 4 The reconstructed face using 3D AAM

1.4 Scaling and Classification

Scaling is a process of normalizing existing data. For a developed 3D face model, it stores the information for each of the three points, x, y and z. Min max scaling will normalize the information. After getting the new data for the x, y and z, the data will be saved and used in the classification process. Classification is a process that identifies the patterns of the data. In this research, the classification will be carried out using software, WEKA. In WEKA there are many classification algorithms such as Random Forest decision trees, neural networks, SVM, *bayesnet* and other. The type of classification will focus on Random Forest. Random Forest algorithm is one of the popular, well known and often used algorithms in research. In addition, other classifications are used in this research to compare the results of the Random Forest with other classifiers. Other classifiers used are SVM and neural networks. There are two sets of data used, 2D and 3D data. Both of these data will be used in WEKA to perform classification and testing. Results for 2D and 3D data are compared and the results will be recorded. Testing is conducted 10 times for both data.

1.5 Random Forest Algorithm

Random forest is an ensemble learning methods for classification and regression. Random forest classification tree grows substantially. To classify a new object from the input vector, the input vector has to be put down on every tree in the forest. Each tree gives a classification decision as a "vote", Random forest will choose a class that has a "vote" for the day (Leo Breiman, 2001). Random Forest algorithm is as follows:

Given a set of training data:

$$\mathcal{D}_{n=} \{X_i, Y_i\}_{i=1}^n \tag{17}$$

The weighted neighborhood scheme makes a prediction for a query point X, by computing:

$$\hat{Y} = \sum_{i=1}^n W_i(X)Y_i \tag{18}$$

For some set of non-negative weights $\{W_i(X)\}_{i=1}^n$ which sums to 1.
Random Forests with constant leaf predictors can be interpreted as a weighted neighborhood scheme in the following way. Given a forest of tree M, prediction made for the X in the tree M is

$$T_m(X) = \sum_{i=1}^n W_{im}(X)Y_i \tag{19}$$

Where $W_{im} = 1/k_m$ if X and X_i are on the same leaf and on the same tree, the m. W_{im} = 0 if X and X_i are not on the same leaf and tree. K_m is the number of training data which falls on the same leaf and tree X. The prediction of the whole forest is:

$$F(X) = \frac{1}{M}\sum_{m=1}^{M} T_m(X) \quad = \quad \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{n} W_{im}(X)Y_i \quad =$$
$$\sum_{i=1}^{n}\left(\frac{1}{M}\sum_{m=1}^{M} W_{im}(X)\right)Y_i \quad\quad\quad\quad (20)$$

Which shows that the Random Forest prediction is weighted average of Y_t with weights:

$$W_i(X) = \frac{1}{M}\sum_{m=1}^{M} W_{im}(X) \quad\quad\quad\quad (25)$$

2.0 RELATED WORK

2D face recognition began in the mid-1960s. In the study of face recognition, 2D is the most traditional method since it is easily accessible. Generally the system use images to extract the information from the images to create a face model in 2D. After the 2D model is created, the model is used to be compared with the existing database [5]. Although using 2D face model in face recognition shows good results, still very limited for identifying a person recorded in CCTV. 2D model is only suitable when using portrait images with an angle of 35 ° degrees. It is said that the comparison between the two models is difficult if the source image is a portrait image and a side face image because both models do not have the same characteristics. Furthermore, since the facial expressions are not the same for both images, it may lower recognition result. [3]

Face recognition in 3D is to solve problems that cannot be resolved by using 2D face model. In theory, the 3D face recognition can identify the face of the image better than the angle of 35 ° degrees. It is said that the 3D face model not only record the points x and y of the facial features, but also the point z (depth) on a face. The advantage of using three points in the model enables the model to be identified even when the model is rotated with an angle of 90 ° degrees [3]. 3D models are usually created by two or more 2D models. After the 2D model is created for a set of images, 2D model is used to calculate the new information depth for a 2D face model. Calculating the depth of the model is complex and takes a long time; the calculation is based on information from the 2D model. Therefore, a high quality 2D model is easier to create a 3D model [11]. Not all 2D models can converted to 3D model as current algorithms still have a certain limit and will be improved in the future [12].

3.0 METHODOLOGY

The data used were video recorder, extracted images from video recorder, 2D AAM and 3D AAM. Figure 5 shows a print screen of video recorded in the process of data collection. DVR is used to support four video footages at the same time. One of the videos is dark; only three video will operate for one DVR.

Figure 5 Print-screen for video footage

Figure 6 shows some extracted images from video footage. A total of 240 images from video are extracted for each model. 10 model x 240 image, the 2400 images are saved in TIF file format so that the image quality can be maintained with the video source.
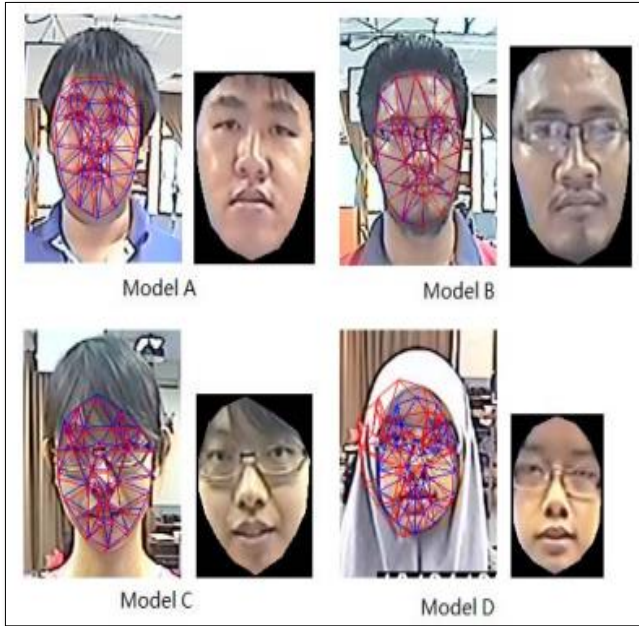

Figure 6 2D AAM

Figure 6 shows a developed 2D AAM for the four models. The left image is the original image and the 2D AAM has been plotted on the original image, the blue model is a model of the original shape model while the red model is the optimal AAM after using Artificial Bee Colony (ABC) algorithm. The right image is reconstructed facial images based on created AAM. Model A, Model B and Model C show a good AAM. AAM Model D is not a good model because the lid affects model form.
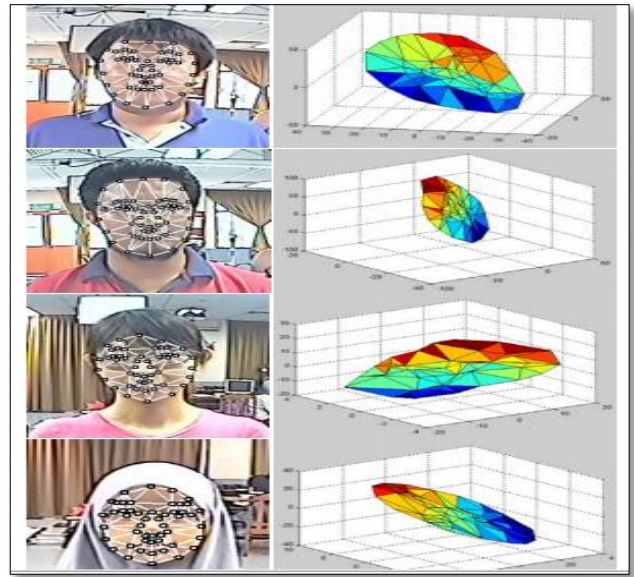

Figure 7 2D AAM

Figure 7 shows the reconstructed 3D AAM face model. 3D AAM is built based on 2D AAM, however this model has a feature that is not available in a 2D model which can support spin from various angles, as shown in Figure 8 below:

Figure 8 The reconstructed 3D AAM faces model

4.0 RESULTS AND DISCUSSION

Testing was conducted using WEKA software. Several tests had been carried out which were the experimental data for 2D and 3D using Random Forest neural networks and SVM classifiers. All tests were conducted in a 10 fold cross validity. The results were recorded and analyzed in the form of tables and graphs so that the accuracy, sensitivity called the true positive rate (TPR) and specificity called the false positive rate (FPR) of the classifier can be compared

Table 1 The Accuracy Value of 2D and 3D Data Based on Three Classifier

Data	Classifier	Classification correct	Classification Incorrect	Accuracy %	TPR %	FPR %
2D	Random Forest	69.5	0.4	99.429	99.400	0.080
3D	Random Forest	69.5	0.4	99.429	99.440	0.080
2D	Neural Network	68.0	2.0	97.143	97.100	0.300
3D	Neural Network	70.0	0.1	99.857	99.860	0.020
2D	SVM	49.5	20.5	70.714	70.720	3.260
3D	SVM	68.6	1.4	98.000	97.980	0.220

TPR means if an image is proven valid, the system also indicates the validation of the image. Similarly, for FPR means that if an image is proven invalid, the system also must indicate invalid person
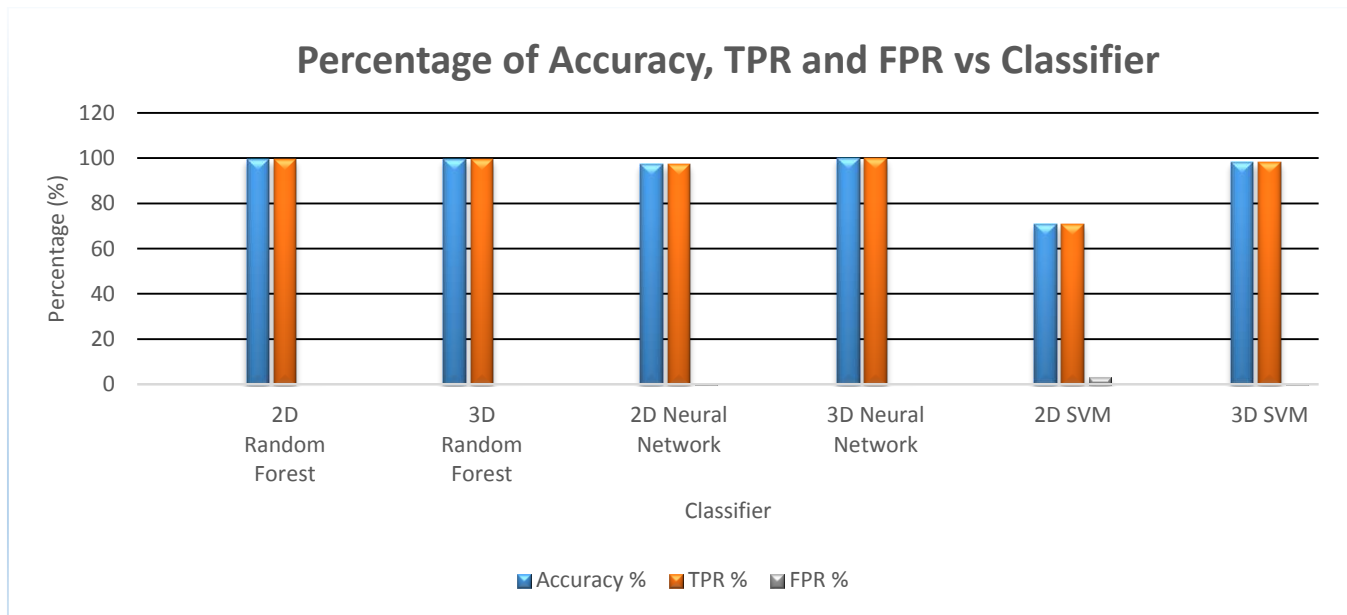

Figure 9 Performance accuracy of classification for 2D and 3D data

Table 1 shows the accuracy of the test for the 2D and 3D data on different classifier for 70 data. In the figure, neural network and SVM classifier has a better accuracy when using 3D data compared to 2D data. Random Forest has shown that it is more appropriate to use the 2D data rather than 3D data. Therefore, the 3D data is more suitable to be used in face recognition as the classifier gives a better performance compared to using 2D data. Figure 9 shows a graph of accuracy and performance versus classifier. This graph shows the classifier that uses 3D data gives a better performance compared to the classifier that uses 2D data.

5.0 CONCLUSION

Face detection is important for digital forensic. This research proves that the 3D model has a better performance of the comparison with 2D models for face recognition. However, knowledge of science and the methods of manipulating this facial recognition technology are expected to be one of the key elements in the process of computer science students' learning and provide opportunities for the students to learn the knowledge of the latest technology. This is due to the diversity of skills and knowledge so thoroughly practical encountered during the development of the project. For the conclusion we conclude that an average accuracy of 3D is higher than the average accuracy for 2D

data. Therefore, a 3D face model is arguably more suitable for use in face recognition compared with a 2D face model.

Acknowledgement

References

[1] Md Ghani, N. A., Liong, C.Y. and Jemain, A. A. 2010. Analysis Of Geometric Moments As Features For Firearm Identification. *Forensic Sci Int*. 198(1-3):143-149.

[2] Abdullah, S. N. H. S., Omar, K., Shahnorbanum, S., Pirahansiah, F. 2010. Multi Threshold Approach For License Plate Recognition System. *WASET International Conference on Signal and Image Processing (ICSIP), 2010*. Singapore. 25-27 August 2010. 1046-1050.

[3] Spencer, W. 2012. Face Recognition: Facial Recognition System–TechFaq

[4] Mohammed Hasan Abdulameer, Siti Norul Huda Sheikh Abdullah, and Zulaiha Ali Othman. 2014. A Modified Active Appearance Model Based on an Adaptive Artificial Bee Colony. *The Scientific World Journal*. 2014, Article ID 879031, 16 page.

[5] M. Hassan, S. H. S. Abdullah, and Z. A. Othman. 2013. Face Recognition Based On Opposition Particle Swarm Optimization And Support Vector Machine. *Proceedings IEEE International Conference on Signal and Image Processing Application (ICSIPA '13). Melaka, Malaysia*. 417-424.

[6] D. Karaboga. 2005. An Idea Based On Honey Bee Swarm for Numerical Optimization, Technical Report-TR06, Erciyes University, Engineering Faculty, Computer Engineering Department.

[7] Iain Matthews, Jing Xiao, and Simon Baker. 2007. 2D vs. 3D Deformable Face Models: Representational Power, Construction, and Real-TimeFitting [Online]. From: http://www.pitt.edu/~jeffcohn/biblio/Matthews%20Xiao%2 0Baker%202007.pdf on 26 January 2013.

[8] Wang, Jun, Yen, Gary G., Polycarpou, Marios M. 2012. *Advances in Neural Networks.* Springer.

[9] Neapolitan, Richard; Jiang, Xia. 2012. *Contemporary Artificial Intelligence*. Chapman & Hall/CRC. ISBN 978-1-4398-4469-4.

[10] Sherk. 2011. Talking More 2D to 3D Conversion with Barry Sandrew from Legend3D [Online]. From: http://3dvision-blog.com/tag/shrek-3d-conversion/ on 26 January 2013.

[11] David Sontag. 2012. Support Vector Machines & Kernels. New York University. [Online]. From: http://cs.nyu.edu/~dsontag/courses/ml12/slides/lecture6.pdf on 01 February 2013.

[12] Edwards, G. J., Taylor, C. J., Cootes, T. F. 1998. Interpreting Face Images Using Active Appearance Models. *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*. ISBN 0-8186-8344-9.