Trends in Undergraduate Research (2019) 2(2): c1-7
https://doi.org/10.33736/tur.1981.2019

3D Face Recognition Analysis Using Random Forest

NURUL ATIFAH ROSLAN, HAMIMAH UJIR* & IRWANDI HIPNI MOHAMAD HIPINY

Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia
*Corresponding author: uhamimah@unimas.my

ABSTRACT

Face recognition is an emerging field due to the technological advances in camera hardware and for its application in various fields such as the commercial and security sector. Although the existing works in 3D face recognition perform well, a similar experiment setting across classifiers is hard to find, which includes the Random Forest classifier. The aggregations of the classification from each decision tree are the outcome of Random Forest. This paper presents 3D facial recognition using the Random Forest method using the BU-3DFE database, which consists of basic facial expressions. The work using other classifiers such as Neural Network (NN) and Support Vector Machine (SVM) using a similar experiment setting also presented. As for the results, the Random Forest approach has yield 94.71% of recognition rate, which is an encouraging result compared to NN and SVM. In addition, the experiment also yields that fear expression is unique to each human due to a high confidence rate (82%) of subjects with fear expression. Therefore, a lower chance to be mistakenly recognized someone with a fear expression.

Keywords: 3D face recognition, Neural Network, Random Forest, Support Vector Machine

INTRODUCTION

3D face recognition is the enhancement of face recognition technology in a few aspects. It has the possibility to overcome feature localization, pose and illumination problems. 3D face recognition is known for its easy and non-intrusive acquisition of face information compared to the other biometrics. Moreover, 3D face recognition has the low privacy of information compared to signatures and fingerprints (Gökberk *et al*., 2008). In recent years, 3D face recognition system has become a popular biometric system due to its higher capability and accuracy, and for its application in various fields such as in commercial, security, health and banking sector.

Existing works in 3D face recognition have been using various kinds of classifiers, for example, Neural Networks (NN) (Hu *et al*., 2013; Kim *et al*., 2017), Support Vector Machine (SVM) (Mousavi *et al*., 2008), Principal Component Analysis (PCA) (Lee & Han, 2006) and Linear Discriminate Analysis (Drira *et al*., 2012; Hiremath, 2013). These classifiers have a few issues when dealing with high dimensional data and facing an overfitting problem. Works using Random Forest also have been discussed in Salhi *et al*. (2012) and, Kremic and Subasi, (2016). However, the comparison with other classifiers using a similar experimental setting has not been presented in any work.

Our objectives are to develop 3D facial recognition using Random Forest and to compare Random Forest with two existing methods, which are Probabilistic Neural Network (PNN) and SVM using a similar experiment setting. In this paper, the analyses of the 3D face recognition experiments conducted are presented. Section 2 describes the existing-related Random Forest works in this study and followed by an explanation of our work in section 3. Section 4 discusses the results and analysis of the experiments. Finally, the conclusions and future works are drawn.

RELATED WORKS

Random Forest

Random forest is one of the most advanced ensembles learning algorithms available, and it runs efficiently on large databases (Kulkarni & Sinha, 2013). In addition, Random Forest is also an effective method to estimate the missing data and capable of maintaining accuracy when a huge proportion of data is missing (Breiman, 2001). Random Forest besides able to generate forest, which can be used on additional data in the future. Despite having those positive traits, Random Forest has its own drawbacks where it can be over fitted for certain data sets, which include data sets used in classification and regression tasks. Plus, data with categorical variables as well as various numbers at a level will make the random forest become more biased towards the attributes which have more levels.

Kremic and Subasi (2016) has proposed a 2D facial recognition with Random Forest method using International Burch University (IBU) database. The dataset consists of face images, training data for training classifier and test data for the evaluation of the final methods. They compared two methods, which are SVM and Random Forest. It begins with reading the image from a database, perform skin colour detection, then convert RGB to greyscale, optimize the histogram value of the methods and used it to classify and retrieve the accuracy of the images. During the testing phase, RF has built 30 trees where each of the trees has considered nine random features. As the results, Random Forest obtained 97.17% inaccuracy. The number of cross-validation folds used in Random Forest evaluation is k = 10. The F-measure and receiver operating characteristic (ROC) area of Random Forest is 97% and 0.998, respectively; which are better than SVM with F-measure of 95% and ROC of 0.95. Therefore, it shows that Random Forest has good performance and gave high accuracy as well as had less time taken to build the model compared to the SVM approach.

Bayramoglu *et al*. (2013) have proven that Random Forest performed better in 3D Facial Action Unit detection by combining the person independent geometric features and descriptor based on Local Binary Pattern (LBP) approach. They proposed the combining descriptors for detecting 3D facial action unit (AU) by using the Random Forest as the classifier. The reason they used person independent geometric feature is that they want to overcome the diversity between different persons and different in ages by distance-based features. Therefore, they proposed in using ratios of distances and areas as well as angles on a single 3D face data. Besides, the Bosphorus database is used in their experiment where they tested their 3D AU detection methods. This database consists of 105 subjects with 24 facial AUs which result in 4666 face scans. Plus, it also has six basic expressions (anger, disgust, fear, happy, sad and surprised), occlusions, fixed rotation and image's intensity. A Gaussian filter is used to smooth the raw 3D data that contains spikes and noises. The performance of the 3D facial AU detection using Random Forest is evaluated by obtaining the mean percentage of ROC of 97.7%.

On the other hand, Drira *et al*. (2012) propose 3D Dynamic expressions' recognition based on novel deformation vector field and Random Forest. In this work, they used BU-4DFE database and Deformation Vector Field (DVF) approach where it is basically a Riemannian facial shape analysis dynamic information of the whole face where the resulting temporary vector field will be used to create a feature vector for recognition of expression from 3D dynamic faces. Besides, a multi-class Random Forest is used with LDA-based feature space transformation to achieve the average recognition rate. The experiment was conducted with 60 subjects, including six basic expressions. Based on this experiment, they obtained 93.21% accuracy for the average recognition rate where it has 19% difference from Wang *et al*. (2006) where they only achieved 73.61% with their proposed method.

An iterative Multi-Output Random Forest (iMORF) algorithm for analysis of the unified face which estimating the head poses, facial expression and landmark position is proposed by Zhao *et al*. (2014). Three databases were used, which were 300-W, Bosphorus and CK+. For 300-W, it is used to automatically detect facial landmarks where 6193 images re-annotated with 68 landmark points and three basic expressions (neutral, happy and others), while Bosphorus is used for face image processing with 105 subjects and 4666 faces with various expressions as well as a variety of poses. The CK+ database was used for research in automatic recognition of expressions where it has eight facial expressions (neutral, anger, contempt, disgust, fear, happiness, sadness and surprised) with 68 landmark points. The performance is evaluated by obtaining an accuracy of a head pose estimation and expression recognition. For head pose estimation using 300-W database, their method achieved the highest accuracy with 86.40%. Meanwhile, for the accuracy of expression recognition, the method yet achieved another highest percentage compared to the other method with 90.04%.

Probabilistic Neural Network

Generally, Probabilistic Neural Network (PNN) is a multilayer feedforward network which consists of four layers, which are input, hidden, summation and output layers. The reason why PNN is chosen is that it is suitable and often used for classification and pattern recognition problems. The process flow is close with the basic NN as the input in the first layer will compute the distance from input vector to training input vector which then produces a vector that the elements will indicate how close the input with training input. For the second layer, the contribution for each class of inputs will be summed together and produce output as a vector of probabilities. Finally, the maximum of probabilities is picked from the complete for transfer function on the output of the second layer which results in producing 1 for positive identification for the class and 0 for negative identification for the non-targeted class (Rutkowski, 2004).

Support Vector Machine

Support Vector Machine (SVM) can be defined as a dimensional hyperplane that separates a set of positive examples from a set of negative examples with maximum margin. Tang and Huang (2008) proposed a 3D 25 facial expression recognition based on the properties of line segments connecting with facial feature points. In their work, they used multi-class SVM and perform expression recognition for person and gender independent using the properties of line segments that connect with certain 3D facial feature points. About 96-dimensional features comprised of the normalized distances and slopes of the line segments are used to recognize the six basic expressions which are anger, disgust, fear, happiness, sadness, and surprised. The result of the recognition rate obtained from this work is 87.1% with the highest average recognition rate which is 99.2% for recognition of surprise expression. The results in this work have outperformed the other related work which they compared.

OUR PROPOSED METHOD

One of the standard phases in the face processing field is to perform facial feature extraction. Between facial features used in face processing are Local Binary Pattern (Bayramoglu *et al*., 2013; Sandbach *et al*., 2012), surface normal (Ujir, 2014), facial feature distances (Soyel & Demirel (2008); Ying *et al*., 2017) et cetera. Certain fiducial points are used as the reference for the facial feature in the recognition process. Facial feature distance is the characteristic of distances extracted from the facial feature points. Facial feature distance can be obtained by various types of distance metric measurements such as Euclidean distance, Geodesic, and others. In our work, 83 facial feature points are used to calculate the facial feature Euclidean distance in the facial extraction phase, refer to Figure 1. The Euclidean distance equation is as in equation 1.

The fundamental element of the Random Forest is to build a small decision tree with few other features where it can lead to a computationally cheap process. In addition, the main principle of Random Forest is a few groups of weak trees (learner) are combined to form a strong learner by obtaining the prediction data and result through averaging of all reached terminal nodes which are known as regression or taking the majority vote where it is based on categorical variables, which are called as classification. Furthermore, in a decision tree of the random forest, the input data will be entered from the top and while it is transverse down the tree, the data will be bucketed into smaller sets.

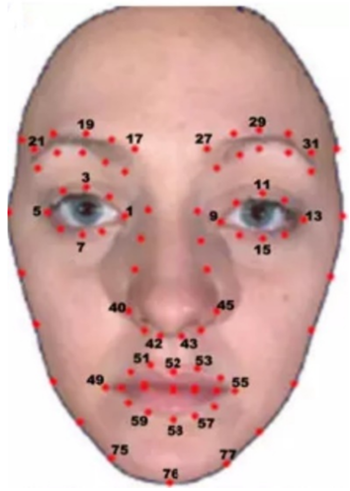

Figure 1: 83 facial feature points on BU-3DFE data (Yin *et al*. 2006)

$$d(x, y, z) = \sqrt{|x_1 - x_2|^2 + |y_1 - y_2|^2 + |z_1 - z|^2} \qquad \text{(Equation 1)}$$

The framework of the Random Forest used is as illustrated in Figure 2. According to Radenkovic (2015), in order to create a training model for the random forest, there are many decision trees required. Therefore, based on the figure above, for each of the decision trees, the number of features and N samples are chosen. Then, m will be extracted from the M features and should be much less than. A training set for this tree will be randomly chosen by using the estimated prediction error of the tree. Later, to choose the split feature, we will choose only among the features, and the best split is calculated based on the in the training set. Now, the decision tree will be created until its fully grown, and the grown tree is split as a maximum as it can, which means until the stopping criterion is achieved and no pruning of the tree. Therefore, when many trees have grown, it will create a random forest model where it is a combination of all the decision trees.

During the training phase for the Random Forest method, the data was divided into two parts, which were training and validation data set. In order to partition the data; hold-out validation was used due to its suitability for a very large data set while preventing the model from overfitting. A hold-out method was used to predict the output values for the unseen data (test/validation data). Then, it will estimate the error for the test set which was used to evaluate the model. In this project, 50% of a hold-out method was used.

Seven basic expressions with four expression levels of intensities (excluding neutral expression, which is only one intensity) were used. During the holdout validation, the data set was divided equally into training and testing data set. After partitioning the data sets, training was proceeded using classification by the Random Forest method. During the classification, the data sets were trained with grown trees, which were equal to 100 trees. According to Breiman (2001), the more trees used in Random Forest, the better the accuracy will be obtained for prediction.

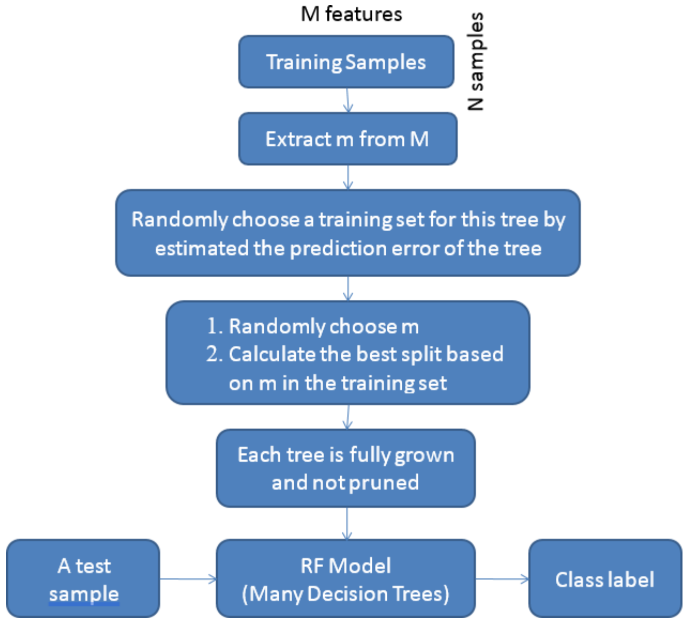

Figure 2: Random Forest process (Radenkovic, 2015).

For the validation test, the validation data sets were used to determine the performance of the Random Forest method in terms of predicting the true class of the subjects. One of the parameters to know whether Random Forest can make a better prediction is by measuring the out-of-bag (OOB) error, which also known as a probability of misclassification. Each tree has its own OOB data set, which is used for error estimation of the individual trees in the forest (Kulkarni & Sinha, 2013). We obtained a 7.14% OOB error. The lower the percentage of error rate, the better the performance of the method used for classification technique. In addition, OOB error for each grown tree also has been calculated and the graph for those errors for each of the grown trees is shown in Figure 3. Based on figure 3, 100 number of grown trees will give the lower value of OOB and that justifies the 100 grown trees used.

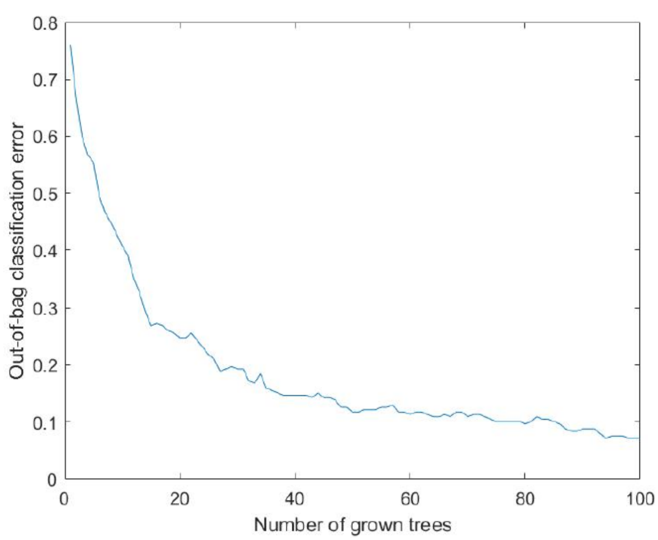

Figure 3: OOB error against number of grown trees

RESULTS AND DISCUSSION

Table 1 shows the comparison between Random Forest, PNN and SVM based on their recognition rate. The result for Random Forest obtained 94.71% of recognition rate and followed by work by Soyel and Demirel (2008) yielded 87.88% of the recognition rate. For the implementation of SVM in work from Tang and Huang (2008), they produced 87.81% of the recognition rate. In this comparison, BU-3DFE database is used, which contains basic expressions.

Table 1. Comparison of error rate.

Work	Method	Database	Emotions	Recognition rate (%)
Our work	Random Forest	BU-3DFE	7 basic emotions	94.7
Soyel and Demirel (2008)	Probabilistic Neural Network (PNN)	BU-3DFE	7 basic emotions	87.88
Tang and Huang (2008)	Support Vector Machine (SVM)	BU-3DFE	6 basics emotions	87.81

Table 2 shows the average confidence rate for the subject with expressions using Random Forest. The confidence rate is the percentage of probability of a subject is recognized based on their facial expressions. The higher the confidence rate obtained, the higher the probability of the expressions belongs to the subject tested, which results in the correct recognition of the subjects. For instance, on average, 75.4% of the subjects with anger expression could be accurately recognized. There were 24.6% of them are mistakenly recognized as other subjects' anger face. The lowest recognition rate is neutral, 58.2%. In other word, when a subject shows a neutral/expressionless face, it can be easily mistaken as other persons. The highest recognition is for fear with 82%. For fear, the facial features for each of the subjects are significant. Therefore, a subject with fear expression is easy to recognize as their fear faces are different from one another.

Table 2. Average of confidence rate (%) for subject with expressions.

Anger	Disgust	Fear	Happy	Neutral	Sad	Surprised
75.4	76.1	82	65.8	58.2	76.7	79

CONCLUSION

Our goal is to compare the Random Forest method with PNN and SVM. Based on the results obtained, Random Forest has given a good performance where it has obtained a 94.71% recognition rate, which is better than PNN and SVM obtained 18.18% of the misclassification rate. The experiment setting used in this work is similar to the existing works in Soyel and Demirel (2008) and, Tang and Huang (2008).

With these promising preliminary results, our future work will be on using Random Forest on real-time 3D face data, specifically on the face-recognition mobile application developed. In addition, a further test of the classification abilities of Random Forest for 3D facial expression classifications in consumer setting as well as data that consist of the non-frontal pose will be carried out.

ACKNOWLEDGEMENTS

This research is fully supported by Malaysian Ministry of Higher Education (MOHE) through Race Acculturation Collaborative Efforts RACE/1331/2016(4). The authors fully acknowledged MOHE and Universiti Malaysia Sarawak for the approved fund which makes this important research viable and effective.

REFERENCES

Bayramoglu, N., Zhao, G., & Pietikäinen, M. (2013). CS-3DLBP and geometry-based person independent 3D facial action unit detection. *Proceedings of International Conference on Biometrics* (ICB), pp.1-6.

Breiman, L. (2001). Random Forests, *Machine Learning*, *45*(1), 5-32.

Drira, H., Ben Amor, B., Srivastava, A., & Berretti, S. (2012). 3D dynamic expression recognition based on a novel deformation vector field and Random Forest. *Proceedings of the 21st International Conference on Pattern Recognition* (ICPR2012), pp. 1104-1107.

Gökberk, B. Salah, A.A., Akarun, L., Etheve, R., Riccio, D., & Dugelay, J.L. (2008). 3D face recognition. In D. Petrovska-Delacretaz, G. Chollet, B. Dorizzi (Eds.), *Guide to biometric reference systems and performance evaluation*, London: Springer-Verlag.

Hiremath, P.S., & Hiremath, M. (2013). Linear discriminant analysis for 3D face recognition using radon transform. In: Swamy P., & Guru D. (Eds.), *Multimedia processing, communication and computing applications*. Lecture Notes in Electrical Engineering, Vol 213. New Delhi: Springer.

Hu, H., Shah, S.A.A., Bennamoun, M., & Molton, M. (2013). 2D and 3D face recognition using Convolutional Neural Network, *TENCON 2017 - 2017 IEEE Region 10 Conference*, pp. 133-132.

Kim, D. Hernandez, M., Choi, J., & Medioni, G. (2017). Deep 3D face identification, *IEEE International Joint Conference on Biometrics (IJCB)* pp. 133-142.

Kremic, E., & Subasi, A. (2016). Performance of Random Forest and SVM in face recognition. *International Arab Journal of Information Technology*, *13* (2), 287-293.

Kulkarni, V.Y., & Sinha, P.K. (2013). Efficient learning of Random Forest classifier using disjoint partitioning approach. *Proceedings of the World Congress on Engineering* 2013 Vol II.

Lee, Y.H., & Han, C.W. (2006). 3D facial recognition using Eigenface and Cascade Fuzzy Neural Networks: normalized facial image approach. In: Grigoriev D., Harrison J., Hirsch E.A. (Eds.), *Computer Science – Theory and Applications*. CSR 2006. Lecture Notes in Computer Science, Vol 3967. Berlin, Heidelberg: Springer.

Mousavi, M.H., Faez K., & Asghar, A. (2008). Three-dimensional face recognition using SVM classifier, *Seventh IEEE/ACIS International Conference on Computer and Information Science* (ICIS 2008), pp. 208-213.

Radenkovic, P. (n.d.). RandomForest. Retrieved from home.etf.rs/~vm/os/dmsw/Random%20Forest.pptx

Rutkowski, L. (2004). Introduction to Probabilistic Neural Networks. In: *New Soft Computing Techniques for System Modeling, Pattern Classification and Image Processing*. Studies in Fuzziness and Soft Computing, Vol 143. , Berlin, Heidelberg: Springer.

Sandbach, G., Zafeiriou, S., & Pantic, M. (2012). Local normal binary patterns for 3D facial action unit detection. *Proceedings of* 19th *IEEE International Conference on Image Processing*, pp. 1813-1816.

Salhi, A.I., Kardouchi, M., & Belacel, N. (2012). Fast and efficient face recognition system using Random Forest and histograms of oriented gradients. *Proceedings of the International Conference of Biometrics Special Interest Group*.

Soyel, H., & Demirel, H. (2008). 3D facial expression recognition with geometrically localized facial features. *23rd International Symposium on Computer and Information Sciences*, pp. 1-4.

Tang, H., & Huang, T.S. (2008). 3D facial expression recognition based on properties of line segments connecting facial feature points. *Proceedings of 8th IEEE International Conference on Automatic Face Gesture Recognition*, pp. 1-6.

Ujir, H., Spann, M., & Hipiny, I. (2014). 3D facial expression classification using 3D facial surface normals. *In* H.A. Mat Sakim & M.T. Mustaffa (Eds.), *The 8th International Conference on Robotic, Vision, Signal Processing & Power Applications*, Lecture Notes in Electrical Engineering 291, pp. 245-253.

Wang, J., Yin, L., Wei, X., & Sun, Y. (2006). 3D facial expression recognition based on primitive surface feature distribution. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1399-1406.

Yin, L., Wei, X., Sun, Y., Wang, J., & Rosato, M.J. (2006). A 3D facial expression database for facial behavior research, *7th International Conference on Automatic Face and Gesture Recognition*, pp. 211 – 216.

Ying, A.C.C., Ujir, H., & Hipiny, I. (2017). 3D facial expression intensity measurement analysis. *In* Zulikha, J. & N.H. Zakaria (Eds.), *Proceedings of the 6th International Conference of Computing & Informatics*, pp. 43-48. Sintok: School of Computing.

Zhao, X., Kim, T., & Luo, W. (2014). Unified face analysis by Iterative Multi-Output Random Forests. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,* pp. 1765-1772.