Hybrid Feature-Based Teeth Recognition System

Pritam R.Dungarwal

***Abstract*— Biometric technology is used in several security applications at present. The biometric information and classification have large impact in efficiency of the application. Teeth are chosen as a characteristic for biometric system in this research since teeth are difficult to be reshaped by intent surgery like well-known features such as face, palm print and finger print. In this paper, appropriate features including machine learning model for teeth recognition are proposed. The features of our system can be classified into two categories: global and local features. Global features are composed of singular values and color histogram of teeth image whilst local features are obtained from normalized teeth width. For the experiments, our proposed features are applied into a system of the multilayer perceptrons network with Levenberg-Marquart back propagation training algorithm. The performance of our system is better than other existing techniques in terms of accuracy and false acceptance.**

***Index Terms*— Teeth identification, biometric, neural network, local features, global features, levenberg-marquardt algorithm**

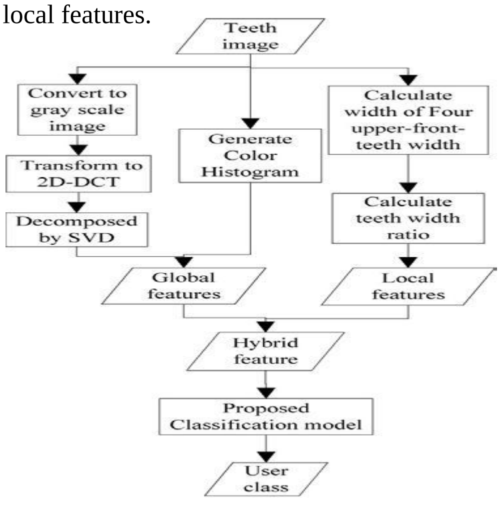
I. INTRODUCTION

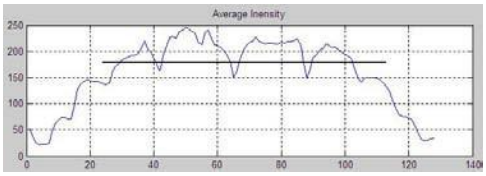
People teeth, among the personal features, had been over-looked. We cannot outwardly distinct the person by teeth pattern. In fact, both pattern and color of people teeth are different according to the personal behavior. However, there is a case that their behaviors are the same; their teeth pattern might be different. In this research, teeth pattern is selected as main feature to distinguish people. The image sample was taken from subjects with the same range of age and same environment. In other hand, this feature can be used in the forensic science. To identify the clay, in some case that face and fingerprint are unusable, the teeth might be usable instead. Apart from that, neural network with back propagation learning is selected as a main tool to identify pattern because of the advantage in time and space consumption. Refer to related works about teeth recognition, Kiattisin etal. Proposed improved PCA and LDA-Based personal identification method [1], which uses PCA as identifier and improved the algorithm, for filtering the unclassified images. Besides, PCA is also used in face recognition as PCA based face recognition and testing criteria [2] proposed by B. Poon et al. and enhanced face recognition through variation of principle component analysis by D.A. Meedeniya et al.[3]. Those techniques require complicated preprocessing. The other technique used for teeth recognition is applying Difference Image Entropy (DIE) to identify individual person. For this technique, Kim et al. proposed teeth recognition based on multiple attempts in mobile device [4], while Joen et al. proposed performance evaluation of DIE-based teeth image recognition system[5]. Their method includes the calculation of average entropy from each image. This causes high complexity. In addition, to improve the accuracy of identification, the mix of two different physical features is proposed by Kim et al. Teeth image and voice are combined for multimodal biometric authentication in mobile environment. Although they achieved high accuracy, many facial features were required for the face recognition process. In this paper, mixed features were proposed along with machine learning model for teeth recognition in terms of accuracy and false acceptance. In our recent experiment, global features are defined by the mixture of SVD of teeth image and color histogram. These features were fed into the neural network based on Levenberg-Marquart back propagation training algorithm because this algorithm can handle the complex input within the short learning period. Moreover, in this paper, the local features are further analyzed to reduce the error affected from the usage of only global features illustrating in preliminary result. To evaluate overall system including proposed features and neural network for teeth recognition, the proposed system was compared with other existing models for teeth recognition such as k-Nearest Neighbor, Naive Bayes, and Resilient propagation (R-Prop). In addition, the system with mixed features is also compared with the system with only global features to ensure that using mixed features yields better results. This organization of this paper is demonstrated in five sections. The next section shows the proposed features and model in details. Next, the experiment environment and results are discussed along with the comparison between our system and other models. Subsequently, the limitation is shown in discussion section. Finally, the last section is conclusion.

II. PROPOSED FEATURES AND MODEL

In this part,our proposed features and model are described. The overall system is depicted in flow diagram as shown in Figure 1. It can be separated into two consecutive parts: the proposed features and classification model.

A. Feature Extraction: The extraction process can be divided into two major parts corresponding to global features and local features.





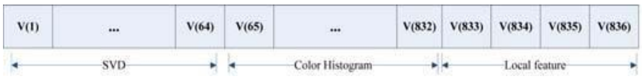
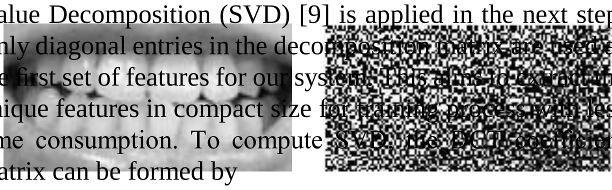


Fig. 1. Flow Diagram

A I. Global Feature: For this part, two Dimensional Discrete Cosine Transform (2D-DCT) is proposed to extract DCT coefficients of teeth image. The 2D-DCT is selected because it has been widely used for extracting a unique set of features for each image in a training set. Due to a property of linear separable transform, 2D-DCT can be processed in term of a traditional DCT (1D-DCT) by performing along a single dimension followed by a one-dimensional DCT in another dimension. The image of transform coefficients can be shown as Figure 2.

Fig. 2. Original gray scale image (left) and its 2D-DCT (right)

From the image of 2D transform coefficients, Singular Value Decomposition (SVD) [9] is applied in the next step. Only diagonal entries in the decomposition matrix are used as the first set of features for our system. This aims to extract the unique features in compact size for training process with less time consumption. To compute SVD, the DCT coefficient matrix can be formed by



$$A = U\,SV^{0} \qquad (1)$$

Where, A :2D-DCT coefficient matrix of size m × n

U :orthogonal columns matrix of size m × r

S: orthogonal columns matrix of size r × r

V: orthogonal columns matrix of size r × n

The composition of 2D-DCT coefficient matrix can be explained in Figure 3

$$\begin{bmatrix} u_{11} & \cdots & u_{1r} \\ \vdots & \ddots & \vdots \\ u_{m1} & \cdots & u_{mr} \end{bmatrix} \begin{bmatrix} s_{11} & 0 & \cdots \\ 0 & \ddots & \\ \vdots & & s_{rr} \end{bmatrix} \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \\ v_{1r} & & v_{rn} \end{bmatrix}$$

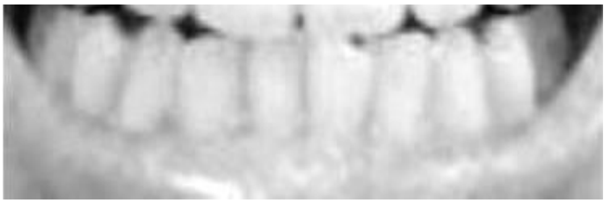
Fig. 3. Sigular Value Decomposition

The second set of global features is obtained from color image histogram. Teeth image are separated into three color channels: red, green, and blue. Then, for each color channel, the frequency of each intensity value is accounted to perform a histogram vector. Combining three histogram vectors with concatenation yields a vector of color histogram. Finally, a set of diagonal entries in singular value matrix and a color histogram will be merged together for creating a complete set of global features.

A II . Local Feature: Local features are the features that can be obtained from some specific locations in the image. For this part, the width of each of four upper front teeth is considered. The reason of selecting a group of upper front teeth is that most of image captured when people smile contains these teeth. To extract local features, suppose that

$$\hat{W}_i = \frac{w_i}{\sum_{i=1}^{4} w_i}. \qquad (2)$$

$\hat{W}_i$ for $i = 1, 2, 3, 4.$



there is a line

Fig. 4. The width of each of four upper front teeth

across four upper front teeth. The intensity values along this line can be projected as shown in Figure 4 and 5. It can be obvious that edge of two adjacent teeth performs local minimum in intensity projection. Then, the width of each of our upper front teeth, wi, is extracted from measuring the distance between two points causing local minimum. After that, the normalized width, Wˆi, of each teeth can be calculated from the equation

After extraction of global and local features, these features are combined into a vector of hybrid features and used as nput of classification model. Figure 6 shows the proposed vector.

Fig. 5. The intensity of the front teeth

Fig. 6. Feature vector

B. Classification Model:

Due to the purpose of reducing time consumption in training step and during the process, multilayer perceptrons back-Propagation with Levenberg-Marquardt training algorithm is selected as the simple model for our system. For the multilayer perceptrons, the output at unit i in layer k + 1 is:

$$a^{k+1}(i) = f^{k+1}(\sum_{j=1}^{k} w^{k+1}(i,j)a^{k}(j) + b^{k+1}(i)). \quad (3)$$

For the Levenberg-Marquardt algorithm, and approximation of the Hessian matrix, the matrix of second derivatives of the updating function which difficult to compute by ordinary analytical methods, so this matrix can be defined using the Jacobian matrix as (4).

$$H = J^{T}J. \quad (4)$$

The Jacobian matrix containing first derivatives of the network errors with respect to parameter vector x in each layer is used to minimize the updating function. The parameters can be updated by:

$$x_{k+1} = x_k - [J^{T}J + \mu I]^{-1}J^{T}e. \quad (5)$$

where e is a network error vector, the μ is a constant to control the speed of parameter learning. In this experiment, we set parameter μ to 0.001 with increasing adaptive value as 5 and decreasing adaptive value is 0.05. There are 2 hidden layer with 3 and 4 neurons respectively in each layer. Each node is linear transfer function. For output layer, the linear transfer function also considered along with target is set to be between -1 and 1. This model architecture is shown Figure 7

III. EXPERIMENT

In this experiment, the images that have been used in the dataset are taken by digital camera, Nikon D3000 with 18-55 VR lens. The original image resolution is 1944 × 2896 Pixels. From the collection, region of teeth are focused and manually located under the normalized size of 64×128 pixels. The example of the teeth images in the dataset is shown in Figure 8. There are 25 subjects in this experiment which are classified into 25 classes and 20 pictures are taken for each person. Therefore, the total number of teeth images is 500

fig. 7. Architecture of Neural network backpropagation model used in this experiment

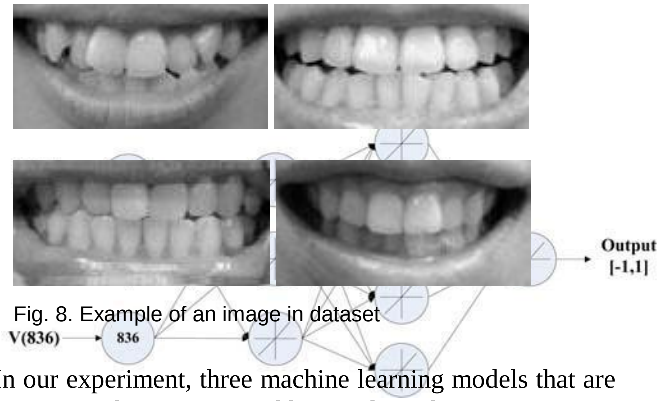


Fig. 8. Example of an image in dataset

In our experiment, three machine learning models that are Naive Bayes, k-Nearest Neighbor, and Resilient propagation training algorithm are selected to compare with multilayer perceptrons with Levenberg-Marquardt training. Moreover, machine learning using only global features and using global and local features are also considered and evaluated in this experiment. For Naive Bayes model, The proportion of training and testing was 75:25. Naive Bayes is a classifier based on probability concepts with normal distribution. In the training process, all related parameters are estimated from the conditional probability of occurrence to form the function of probability distribution for independent features in the given class. In the testing process, the model calculates the posterior probability of the teeth image for an appropriate class. This model yields high false acceptance rate because there is no reject class. In other words, all misclassified images are all images assigned into other class. The False Acceptance Rate (FAR) can be calculated by (6).

$$FAR = \frac{f_n}{f_n + f_p}. \quad (6)$$

where fn: Number of False accepted Images
fp: Number of True accepted Images

For k-Nearest Neighbor, this classifier is based on a majority of k-closest samples in the feature space. In this classification, there are two main steps. First, find the k training samples nearest to the considered input. Then, assign the class of object as the class of most of k training samples. The k-Nearest Neighbor using only global features and that using hybrid features are analyzed. For the Resilient propagation, this model is used to compare with our method. Five-fold cross validation was also considered together with 75:25 proportions of training images and testing images. The learning rate and the maximum epoch are set to be 0.25 and 100, respectively. The values of two classes are set to be ’-1’ and ’1’. The linear activation function is chosen with two hidden layers consisting of three and four neurons, respectively. All parameters are defined to yield the best results as possible. The image is assigned to the class whose output is greater than zero. The R-prop differs to traditional back propagation learning in term of weight tuning based on only the change of direction or sign of the derivative without the considering the magnitude of the derivative. The advantage of resilient propagation is speed for training. The architecture of this model is the same as shown in Figure 7. Besides, only the algorithms for training are different. For our proposed method, neural network with the Levenberg-Marquardt training algorithm, dataset was divided Into training set and testing set with the proportion of 75:25 as well as the five-fold cross validation. The parameters were set as described in section II. The activation function is symmetrical linear function and the targets are set to be -1 and 1.

After the images were classified with provided models, the accuracy rate and False Acceptance Rate (FAR) are calculated to comparison, as shown in Table I.

TABLE I
RESULT TABLE OF BACKPROPAGATION WITH LAVENBERG MARQUARDT AND OTHER MODELS

Model	Feature	Accuracy rate(%)	FAR(%)
Lavenberg Marquardt	**Mixed**	**93.6**	**0**
Lavenberg Marquardt	Global	91.2	0.87
Naive Bayes	Mixed	66.4	50.60241
Naive Bayes	Global	66.4	50.60241
k-Nearest Neighbor	Mixed	76.6316	30.49451
k-Nearest Neighbor	Global	76.6316	30.49451
R-Prop	Mixed	81.6	0
R-Prop	Global	82.4	0.97

From the experimental result, among all models with only global features, it is obvious that multilayer perceptrons with Levenberg-Marquardt training algorithm is better than othermodels in terms of accuracy and FAR.

IV. DISCUSSION

From the experiment , k-Nearest Neighbor and R-Prop are not appropriate to the mixed features. The results prove that although adding local features a input of their systems, the performance of them is not high enough to practically use. In case of Naive Bayes and k-Nearest Neighbor, their accuracy rate is not even changed. This means that the local features does not affect the performance of those models. However, the performance of multilayer perceptrons with Levenberg-Marquardt algorithm is higher after adding than that in recent experiment because the number of person (classes) is increased. However, with the combination of global features and local features, the Levwnberg-Marquardt backpropagation neural network can predict class of input with high accuracy. The FAR of the network is zero as well. Moreover, the failure in classification might be obtained from the physical characteristic of image when taken a shot. For instance, the motion blur causes the misclassified as shown in Figure 9.

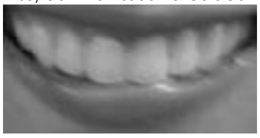


Fig. 9.Example of a motion blur image in dataset

V. CONCLUSION

According to the experiment results, the features based on SVD, color histogram and normalized teeth width with Levenberg-Marquardt training algorithm are suitable to be form teeth recognition system. From the result, my proposed system gave the accuracy rate about 93.6% without false acceptance. This means that the dataset is larger, the more specific features is needed to make all feature vectors easier to classify. However, selecting classification model is also significant. Finally, this proposed system can be used for biometric identification with easiness and low resource consumption.

REFERENCES

[1]S. Kiattisin and A. Nadee, “Improved PCA and LDA-based personal identification method,” Proc. 3rdInt’l Symp. Biomedial Engineering, pp. 461-465, 2008.
[2] B. Poon, M. A. Amin, and H. Yan, “PCA based face recognition and testing criteria,” Proc. 8thInt’l Conf. Machine Learning and Cybernetics,
pp. 2945-2949, Baoding, 2009.
[3] D.A. Meedeniya and D.A.A.C. Ratnaweera, “Enhanced face recognition through variation of principle component analysis(PCA),” Proc. 2ndInt’l Conf. Industrial and Information System, Sri Lanka, pp. 347-351, August
2009
[4] D. Kim, J. Shin, and K. Hong, “Teeth recognition based on multiple

attempts in mobile device," J. Network and Computer Applications, vol. 33, pp. 283-292.

[5] J. Joen, J. Kim, J. Yoon, and K. Hong, "Performance evaluation of teeth image recognition system based on Difference Image Entropy," Proc. 3rd Int'l Conf. Convergence and Hybrid Information Technology, pp.967-972, 2008.

AUTHORS PROFILE

Pritam R. Dungarwal is lecturer at S.H.H.J.B. POLY CHANDWAD and doing Master degree in Computer science and Engg. From MIT COE Aurangabad.