Recognition of Tongueprint Textures for Personal Authentication: A Wavelet Approach

Salim Lahmiri
University of Quebec at Montreal/Department of Computer Science, Montreal, Canada
lahmiri.salim@courrier.uqam.ca

Abstract— In order to verify tongueprint images, three approaches for texture analyses were considered and their performances are compared. They are wavelet transform, Gabor filter, and spectral analysis. In all approaches, six statistical measures are applied to the processed images to extract features. They are the mean, standard deviation, smoothness, third moment, uniformity, and entropy. Finally, k-nearest neighbor algorithm (k-NN) is used to classify tongue textures for verification purposes. The obtained recognition rates show that features extracted from wavelet analysis allow achieving the highest accuracy (92%) among the other approaches. On the other hand, features extracted from spectral images lead to the lowest recognition rate (75%). Features extracted from Gabor filter banks obtained 83%. Therefore, we conclude that wavelet-based features outperform Gabor and spectral-based features employed in the literature.

Index Terms— Tongueprints, wavelets transform, Gabor filter, Spectral analysis, k-NN.

I. INTRODUCTION

Biometric systems seek to automatically recognize persons based on their physiological and/or behavioural characteristic. In particular, "primary focus of biometric technology is verification and identification of humans using their naturally possessed biological (biometric) properties" [1]. Biometrics science is largely employed in security and surveillance applications, forensics, secure control access, and automatic banking, to name a few. The most measurable traits proposed in literature for authentication are fingerprints, palmprints, and face. Other traits and corresponding technologies are described in [2]. Recently, due to the importance of biometrics many comprehensive surveys and works have been published on fingerprint [3], palmprint [4], face [5][6], retina [7], iris [8], speech recognition [9], gait [10], and behavioral human-computer interaction (HCI) [11].

Recently, a novel biometric system for person identification based on the tongueprint was proposed and tested by Zhang et al. [12]. Indeed, the tongue is a unique and distinctive organ because it is characterized by its stable geometric features, crack features and texture features [12]. In addition, unlike fingerprints and palmprints which are exposed to external environment changes, the tongue is well protected since it is contained inside the mouth. Finally, the squirm of the human tongue could be used to identify persons. The authors in [12] employed a fusion approach that makes use of both static (geometric, crack, and texture features) and dynamic features (squirm) to achieve a 95% recognition rate.

In addition, The authors in [12] provided four verification performances: 89.3% using curvature of the contour of the tongue to detect the geometric shape, 79.4% and 72.5% using Gabor filter and spectral analysis respectively to classify texture, and 71% using manifold learning technologies [12] to classify squirm.

According to the authors in [12], tongue contour extraction is difficult since the surface color of the tongue is highly similar to that of the ambient biological tissue. On the other hand, the authors employed spectral analysis to process tongue textures. But, techniques of spectral analysis suffer from artefacts and limited resolution [13]. In addition, the fusion system is not easy to implement and data processing is time consuming. Furthermore, Gabor filters have three major limitations [14][15]. First, the outputs of Gabor filter banks are not mutually orthogonal; then a significant correlation between texture features may occur. Second, there is a need for an optimal tuning of its parameters including the filter central frequency, the filter bandwidth among the x and y-axis and the filter orientation. Third, Gabor filter banks come with high computational costs. The disadvantages of Gabor filter can be avoided if the wavelet transform is used [14][15]. For instance, wavelet transform provides a precise analysis of a signal at different scales. In addition, it uses a low pass and high pass filters that remain the same between two consecutive scales. Therefore, it does not require proper tuning of parameters as with Gabor filters.

The purpose of this study is to design a biometric authentication system based on the analysis of tongueprint texture only. Indeed, the literature argues that texture features are useful in the classification of tongue images [12][16][17]. In addition, tongue contour extraction is avoided since it is difficult to perform and squirm features are ignored because of their poor performance [12]. To overcome limitations of Gabor filter banks and spectral analysis that were employed in [12], discrete wavelet transform is used to process tongueprint images. It is widely recognized as a powerful technique for feature extraction in pattern recognition [18][19]. In addition, relying on texture only and ignoring geometric and crack features allows implementing a simple and fast identification system. In sum, the main hypothesis is that relying only on texture analysis using wavelets transform may lead to higher classification.
performance compared to Gabor filter and spectral analysis that were used in [12].

The paper is organized as follows. In Section 2, the review of literature is presented. Section 3 presents the methods. The data and experimental results are presented in Section 4. Finally, the conclusion and future directions are given in Section 5.

II. RELATED WORKS

A very limited number of papers found in literature examined the tongue texture for classification purposes and mainly for medical diagnosis. This method of diagnosis is important in Traditional Chinese Medicine (TCM) [13]. In this paper, the authors proposed a computerized tongue diagnosis where Bayesian Network classifier based on chromatic and textural measurements to classify healthy and abnormal tongues (13 diseases) from a group of 455 patients. The correct classification rate was 75.8%. Another paper [19] used Gabor Wavelet Opponent Colour Features (GWOCF) to analyze tongue images in order to perform a tongue diagnosis in TCM. In particular, they employed colour information to pre-classify the known texture image before extracting GWOCF to achieve 89% recognition rate using patient’s tongue images captured in Guangzhou Traditional Chinese Medicine Hospital. In a subsequent paper, authors in [20] computed the entropy and energy functions to represent the texture features, and employed a k-Means algorithm to select the clusters and finally used 3-D visualization to classify 11 normal tongue images and 8 tongue images from patients with gastro cancer. The author concluded that color and texture features are sensitive to abnormal tongues. The authors in [21] designed a Computerized Tongue Examination System (CTES) to automate the diagnostic of tongue images based on chromatic and spatial textural properties. In particular, colors of substance and coating, thickness of coating and the detection of grimy coating have been measured. Indeed, textural features including the angular second moment (ASM), contrast, correlation, variance and entropy were used to determine the grimy coating of the tongue. The k-NN algorithm successfully classified 86% of the tongue images. The authors concluded that there is a real potential for computerized tongue diagnosis. Authors in [22] proposed a Tongue-Computing Model (TCoM) for the diagnosis of appendicitis. Chromatic and textural metrics are the basic features that are jointly used to classify tongue images. In addition, they have proposed a new measurement called Grade of Differentiation (GOD) to evaluate the classification performance of different metrics. Then, the nearest distance rule for the classification of each metric is used to classify images. Finally, a survivor metric is employed to obtain a final decision. The experimental results from 912 tongue images show that the ratio of the correct classification is 92.98% and that of false classification is 8.52%.

As mentioned above, a novel biometric system for person identification based on the tongueprint was proposed and tested in [12]. According to the authors, the major advantage of their proposed system is its noninvasiveness. For instance, to recognize persons, [12] employed a fusion approach that makes use of both static (geometric, crack, and texture features) and dynamic features (squirm). The obtained recognition rate was 95%. This fusion approach performed better than using separately geometric features, crack features, textures features, or squirm features. The tongueprint is a promising candidate for biometric identification and worthy of further research. The authentication system proposed by [12] uses both static and dynamic features. They are contained in two modules. The first module is called enrolment module where static and dynamic features are extracted. The second module is called recognition module that operates following two steps. The first step is the liveness detection where information regarding the squirm of the tongue is used to detect whether the subject is alive. In the second step, both static physiological features and dynamic squirm features are extracted. The extraction of static physiological features step is aimed to extract geometric features, crack features and textural features. The geometric features are measures of the width of the tongue, its thickness, and the curvature of its contour. Combining these three measures allows forming the geometric vector. The crack features correspond to the lines found on the centre of the tongue surface. Therefore, a region of interest (ROI) is defined and along the centre of the tongue surface and two-dimensional Gabor filter is applied to extract crack features from ROI. Finally, frequency domain images are represented using the polar coordination system and energy is computed to form the textural feature. In the step of the extraction of dynamic features, the tongue squirm is captured in a sequence of continuous images. Then, the authors used the orthogonal neighbourhood preserving projections (ONPP) to reduce the tongue squirm into a lower dimensional feature space. According to the authors [12], the ONPP technique was used since it is a linear dimensionality reduction that allows preserving both local and global geometry of high dimensional data samples. A dynamic descriptor is calculated using the mean and the variance obtained from the analysis of the low dimensional manifold obtained using the ONPP technique. The dynamic descriptor is employed to classify the input image sequence as valid or invalid. Then, the squirms of all subjects are clustered. Finally, static features and dynamic features are grouped to form one vector of features for each subject. Then, the mean and the variance of each vector are calculated. The classification of subject is made by minimum distance method using the computed mean and variance.

The proposed fusion system of [12] provides higher classification rate (95%). However, this system comes with several drawbacks as mentioned in introduction. The tongue contour extraction is difficult since the surface color of the tongue is highly similar to that of the ambient biological tissue [12]. In addition, this system is not easy to implement and data processing is time consuming. Finally, spectral analysis was chosen to perform images and extract features; which is not an adequate approach.
Gabor filter banks are not mutually orthogonal, they need for an optimal tuning of parameters and leads to high computational costs.

The purpose of this study is to compare the performance of the $k$-NN classifier given the type of the approach followed to extract texture features from tongueprints: wavelet analysis, Gabor filter, and spectral analysis approach. The main hypothesis is that relying only on texture analysis using wavelets transform may lead to higher classification performance compared to Gabor filter banks and spectral analysis employed in [12]. In addition, relying on texture only and ignoring geometric and crack features allow implementing a simple and fast identification system. In addition, texture analysis has been proven to helpful for the extraction of features [3][4][5][8]. Moreover, a system where textural properties of a tongue are used can be implemented in clinical medicine for diagnosis purposes [13].

In sum, the contribution of our paper is to experimentally shed light on the benefit of the DWT-based features in tongueprint textures recognition. Indeed, in previous studies textural features were directly extracted from tongueprint images [13] [20][21][22] or from spectral and Gabor processed images [12][19]. However, these approaches come with serious drawbacks. In one hand, features extracted directly from tongueprint images do not contain high frequency information that characterizes biological tissue. On the other hand, as mentioned in introduction spectral analysis suffer from artefacts and limited resolution, and Gabor transform requires optimal tuning of its parameters and may lead to correlated features. As a result, using DWT to extract features would allow obtaining better high frequency features than spectral and Gabor transform. Indeed, the disadvantages of Gabor filter can be avoided if wavelet transform is adopted since it provides a precise analysis of a signal at different scales [14][15].

III. METHODOLOGY

The purpose is to compare the recognition ability of wavelet features in comparison with Gabor processed images and spectral features. The research methodology consists of four parts as follows:
1) Extraction of region of interest (ROI) from the original images.
2) Processing ROI using wavelet, Gabor, and spectral analysis.
3) Computing statistics of the processed ROI.
4) Computing statistics of the non-processed ROI.
5) Classification of features using $k$-NN.
6) Comparison of the results.

A. Spectral Analysis

Based on the Fourier transform, spectral analysis allows detecting high-energy bursts in the spectrum of texture. The spectrum is expressed in polar coordinates to obtain a function $S(r, \theta)$ where $S$ is the spectrum function and $r$ and $\theta$ are the variables in the polar coordinate system. For each direction $\theta$, the function $S(r, \theta)$ is considered as a one dimensional function: $S_\theta (r)$. Meanwhile, for each frequency $r$, the function $S(r, \theta)$ is considered as a one dimensional function: $S_r (\theta)$. Then, the spectral measures to describe texture are obtained by summing the functions $S_\theta (r)$ and $S_r (\theta)$ as follows:

$$S(r) = \sum_{\theta=0}^{\pi} S_\theta (r)$$

$$S(\theta) = \sum_{r=1}^{\pi} S_r (\theta)$$

where $r_0$ is the radius of a circle centered at the origin. For each tongueprint image, the features $S_\theta (r)$ and $S_r (\theta)$ constitutes the spectral energy that describes its entire texture as in [12]. Then, they are fed to the classifier. The biometric system based on texture spectra for tongueprints verification is shown in Figure 1.

B. Gabor Filter

The one dimension (1-D) Gabor filter was first defined by Gabor [22] and was later extended to 2-D by Daugman [23]. The Gabor filter is extensively used in texture analysis since it decomposes an image into components corresponding to different scales and orientations [24-26].

Therefore, the two-dimensional (2D) Gabor filter is able to capture visual properties such as spatial localization, orientation selectivity, and spatial frequency. Thus, Gabor filter is well-adapted for image processing applications; especially texture analysis. The 2D Gabor filter is the product of a 2D Gaussian and a complex exponential function. For instance, the Gabor function is a complex sinusoid centered at a given frequency and modulated by a Gaussian envelope. The Gabor filter comprises both real and imaginary parts. The general form of the real part of a 2-D Gabor function is defined as follows:

$$G(x, y, \sigma_x, \sigma_y, f, \theta) = \exp \left[ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] \cos(2\pi f r')$$

$$x' = x \cos(\theta) - y \sin(\theta)$$

$$y' = x \sin(\theta) + y \cos(\theta)$$
where $\sigma_x$ and $\sigma_y$ are respectively the standard deviations of the Gaussian envelope along the $x$ and $y$ axes respectively. The parameters $f$ and $\theta$ are respectively the central frequency and the rotation of the Gabor filter. Then, to obtain the Gabor-filtered image $f(x,y)$ of a given input image $h(x,y)$ the 2-D convolution operation ($\ast$) is performed as follows:

$$f(x,y) = G(x,y,\sigma_x,\sigma_y,f,\theta) \ast h(x,y)$$  

(6)

The selection of the parameters ($\sigma_x$, $\sigma_y$, $f$, $\theta$) is crucial and difficult. Bianconi and Fernández [25] investigated the effects of Gabor filter parameters on texture classification using a large set of experiments on different textures. They conclude that the number of orientations did not show significant effects on the percentage of correct classification. In other words, increasing the number of orientations would produce a considerable waste of computational time, without tangible increase in the correct classification rate [25]. Thus, the value of $\theta$ is set to 0 in our study. On the other hand, they found that best classification rate is obtained when the parameters $\sigma_x$, $\sigma_y$, and $f$ are set to their lowest level (0.5). In other words, they concluded that low selectivity in the frequency positively affects correct texture classification. Therefore, the standard deviations of the Gaussian envelope along the $x$ and $y$ axes and the central frequency of the Gabor filter are all set to one half in our study. The biometric system based on Gabor filter is shown in Figure 2.

![Diagram](image)

**Figure 2.** Biometric system with Gabor filter.

### C. Wavelet Transform

The discrete wavelet transform (DWT) decomposes an image into several sub-bands according to a recursive process (Figure 3). These include LH1, HL1 and HH1 which represent detail images and LL1 which corresponds to the approximation image. The approximation and detail images are then decomposed into second-level approximation and detail images, and the process is repeated to achieve the desired level of the multi-resolution analysis. The obtained coefficients values for the approximation and detail sub-band images are useful features for texture categorization [18][19]. To obtain the set of features that characterize a given texture image, the 2D-DWT wavelet transform is used to find its spectral components. This allows transforming each texture image into a local spatial/frequency representation by convolving the image with a bank of filters. Then, the image features are extracted from the obtained 2D-DWT representation. A 1D-DWT is defined as follows:

$$f(x) = \sum_{i,j} c_{i,j} \psi_{i,j}(x)$$  

(7)

Where $\psi_{i,j}(x)$ are the wavelet functions and and $c_{i,j}$ are the DWT coefficients of $f(x)$. They are defined by:

$$c_{i,j} = \int_{-\infty}^{\infty} f(x) \psi_{i,j}(x)$$  

(8)

A mother wavelet $\psi(x)$ is used to generate the wavelet basis functions by using translation and dilation operations:

$$\psi_{i,j}(x) = 2^{j/2} \psi(2^{-j} x - i)$$  

(9)

where $j$ and $i$ are respectively the translation and dilation parameters. The one-dimensional wavelet decomposition can be extended to two-dimensional objects by separating the row and column decompositions [27][28]. A second level decomposition is considered. For instance, features are extracted from HH2. The Daubechies-4 wavelet [29] is chosen as the mother wavelet since it has the advantage of better resolution for smoothly changing signals [30]. The biometric system based on wavelet analysis and used for tongueprints verification is shown in Figure 4.

![Diagram](image)

**Figure 3.** Two level 2-D DWT decomposition of an image.

---

1 Features were also extracted from Gabor filter banks with three orientations; for instance $\theta = 0$, $\pi/4$, $\pi/2$ and following the design rules as in [14]. However, the obtained classification results were similar to the approach we follow and which is proposed in [25].
The classifiers used to describe the processed images are the mean, standard deviation, smoothness, third moment, uniformity, and entropy. They are chosen since they are widely used in pattern recognition [31]. The statistics are expressed as follows:

\[
\text{Mean} = m = \frac{1}{L} \sum_{i=0}^{L-1} z_i p(z_i) \tag{10}
\]

\[
\text{St.Dev} = \delta = \sqrt{\mu_2} = \sqrt{\sigma^2} \tag{11}
\]

\[
\text{Smoothness} = R = 1 - \frac{1}{L + \delta^2} \tag{12}
\]

\[
3\text{th.Moment} = \mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i) \tag{13}
\]

\[
\text{Uniformity} = U = \frac{1}{L} \sum_{i=0}^{L-1} p^2(z_i) \tag{14}
\]

\[
\text{Entropy} = e = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i) \tag{15}
\]

where z is a random variable indicating intensity, p is the probability density of the ith pixel in the histogram, and L is the total number of intensity levels. A vector is constructed with these statistics as feature components to be fed to the k-NN classifiers.

E. The Classifier

The k-nearest neighbor algorithm (k-NN) was first introduced by [32] and is a nonparametric method that assigns query data to the class that the majority of its k-nearest neighbors belong to. For instance, the k-NN algorithm uses the data directly for classification without the need of an explicit model. The performance of k-NN depends on the number of the nearest neighbor k. In general, there is no solution to find the optimal k. However, trial and error approach is usually used to find its optimal value. Therefore, our objective is to find the value of k that maximizes the classification accuracy. The main advantage of k-NN algorithm is the ability to explain the classification results. On the other hand, its major drawback is the need to find the optimal k and to define the appropriate metric to measure the distance between the query instance and the training samples. In this paper, the distance metric chosen is the Euclidean distance. The standard algorithm of k-NN is given as follows:

1. Calculate Euclidean distances between an unknown object (o) and all the objects in the training set;
2. Select k objects from the training set most similar to object (o), according to the calculated distances;
3. Classify object (o) with the group to which a majority of the K objects belongs.

IV. EXPERIMENTAL RESULTS

The images are provided by the Biometric Research Centre of The Hong Kong Polytechnic University [33] (PolyU/HIT Tongue database). The PolyU/HIT Tongue Database contains 12 color images in BMP image format, which are used in [13][17][30]. According to [33] these images are captured using a black box with special lighting. Thus the capturing environment is stable. The number of images in the database is twelve. An example of the original images and ROI are given in Figures 5 and 6. The original tongue database in [12] contains 174 images. We were provided only with twelve images. Nine images are used for training and three images for testing. The training database is composed of six true images and three impostor images. The images that form training and testing databases were randomly selected. Finally, the correct recognition rate is computed for each experiment. The Matlab software is used to perform wavelet analysis, Gabor filter, and spectral analysis and to train and test the k-NN classifier. The classifier have been trained and tested with k varying from one to five. The best recognition rates were obtained for k=1 and k=2 which give similar performances.
Then, increasing \( k \) above 2 leads to increasing the error rate of the classifier for all types of features. Therefore, \( k=1 \) is chosen as the optimal value. Finally, three performance measures are used. They are the correct classification rate (CCR), sensitivity, and specificity. The performance measures are given as follows:

\[
\text{CCR} = \frac{\text{CCS}}{\text{CS}} \quad (16)
\]

\[
\text{Sensitivity} = \frac{\text{CCPS}}{\text{TPS}} \quad (17)
\]

\[
\text{Specificity} = \frac{\text{CCNS}}{\text{TNS}} \quad (18)
\]

where CCS is correctly classified samples, CS is classified samples, CCPS is correctly classified positive samples, TPS is true positive samples, CCNS is correctly classified negative samples, and TNS is true negative samples. The experimental results for each experiment given the type of features are shown in Figures 7 and 8. On the other hand, Figures 9, 10, and 11 show the scatter plots of spectral, Gabor and DWT approach respectively. By analyzing the results, the \( k \)-NN correctly classifies the testing data with 92\% accuracy. On the other hand, the correct detection rate obtained with Gabor filter features is 83\%. Finally, \( k \)-NN achieves the lowest performance when features from spectral analysis are considered as inputs (75\%). Besides, the experimental results suggest that all three approaches successfully detected the true persons with 100\% sensitivity. However, they all fail to correctly detect impostors. For instance, the specificity statistic obtained by DWT, Gabor and spectral analysis is 83\%, 67\%, and 50\% respectively.

The results show evidence of two findings. First, features obtained with wavelet transform perform much better that features obtained with Gabor filter and spectral analysis. Then, our choice for wavelet transform to extract features from tongue texture is justified. Second, features obtained with Gabor filter perform much better than features obtained with spectral analysis. This result is consistent with finding in [12].
V. CONCLUSION

The purpose of this study is to compare the performance of the \( k \)-NN classifier given the approach used to extract features from tongueprint texture for person verification. Unlike previous literature [12] that uses Gabor filter and spectral analysis for feature extraction from tongueprint for person authentication, the wavelet transform is considered in this study. The results show strong evidence that features obtained with wavelet transform help improving the correct detection rate. In addition, features obtained with spectral analysis perform the worst. For instance, features extracted with Gabor filters perform much better than spectral analysis features. This finding is consistent with the results found in [12]. Finally, all features extraction approaches provide a 100% sensitivity statistic. Therefore, they are all capable to detect true persons. On the other hand, the wavelet approach gives the highest detection rate of the impostors. The wavelet-based system for tongueprint authentication has proven its effectiveness and could be a very promising approach to be implemented in clinical medicine for diagnosis purposes mainly the Traditional Chinese Medicine.

The spectral analysis which is based on the Fourier transform characterizes the spatial-frequency distribution, but ignores the information in the spatial domain. On the other hand, Gabor filter jointly considers specific frequency and orientation characteristics to analyze textured images. Then, the performance of \( k \)-NN with Gabor features is higher. However, the design of optimal Gabor filter is a very hard since it depends on many parameters. Thus, human intervention is required to select the appropriate parameters for texture analysis. Finally, the wavelet transform for feature extraction is useful since it provides uncorrelated data and improves recognition accuracy with information extracted from HH subbands. This indicates that high frequency channels contain more information regarding the texture of the biological tissue. Furthermore, wavelet transform does not require a lot of parameters like Gabor filter. Indeed, only the type of mother wavelet and level of resolution should be selected.

In sum, it is suggested to use wavelet transform in tongueprint texture identification and verification. For future work, a larger database will be considered. In addition, it is recommended to explore different types of mother wavelets at different resolution levels to examine their effects on the recognition performance of a biometric system based on tongueprints.

REFERENCES


Salim Lahmiri (M.Eng, Ph.D, Canada) is interested in biomedical signal and image processing, computer vision, and pattern recognition.