

Automatic Tongue Recognition Based on Color and Textural Features

Ryszard S. Choraś^(✉)

Institute of Telecommunications and Computer Sciences,
UTP University of Science and Technology, S. Kaliskiego 7,
85-796 Bydgoszcz, Poland
choras@utp.edu.pl

Abstract. This paper proposes a method of tongue recognition. Tongue images have many advantage for personal identification and verification. In this paper a tongue features are extracted based on color and texture features. These features can be used in forensic applications and with other robust biometrics features can be combined in multi modal biometric system.

1 Introduction

Personal identification is crucially significant in a variety of applications. Conventional person's identification systems used:

- Something you have:
 - Token: key, card, or badge
- Something you know:
 - Password
 - PIN numbers
- Something you are:
 - Biometric
 - Physiological
 - Behavioral

A password for personal identification has the risk that the user forgets it or the other persons use it. User's have problems in terms of theft, loss, and reliance on the own memory.

Images play an important role in the identification process of people [2,4]. Biometric identification systems are systems that use pattern recognition ways to identify a specific person by establishing the authenticity of a specific physiological or behavioral characteristic of that person [8].

Biometric systems have four main components: sensor, feature extraction, biometric database, matching-score and decision-making modules. The input subsystem consists of a special sensor needed to acquire the biometric signal. Invariant features are extracted from the signal for representation purposes in the feature extraction subsystem. During the enrollment process, a representation

(called template) of the biometrics in terms of these features is stored in the system. The matching subsystem accepts query and reference templates and returns the degree of match or mismatch as a score, i.e., a similarity measure. A final decision step compares the score to a decision threshold to the comparison a match or non-match.

A.K. Jain et al. [8] also defines the following requirements that a given measure must satisfy to be a biometric: The ideal biometric characteristics have five qualities:

1. Robust: Unchanging on an individual over time. “Robustness” is measured by the probability that a submitted sample will not match the enrollment image.
2. Distinctive: Showing great variation over the population. “Distinctiveness” is measured by the probability that a submitted sample will match the enrollment image of another user.
3. Available: The entire population should ideally have this measure in multiples. “Availability” is measured by the probability that a user will not be able to supply a readable measure to the system upon enrollment.
4. Accessible: Easy to image using electronic sensors. “Accessibility” can be quantified by the number of individuals that can be processed in a unit time, such as a minute or an hour.
5. Acceptable: People do not object to having this measurement taken on them. “Acceptability” is measured by polling the device users.

Tongue recognition is attracting a great deal of attention because of its usefulness in many applications [3]. Traditional, tongue recognition are often classified into two groups:

- Tongue recognition and analysis for the patient disease diagnosis. Tongue recognition for diagnosis has played an important role in traditional Chinese medicine (*TCM*) and in this area most investigation has been focused on extraction of chromatic features [10, 18], shape and textural features [5, 6, 11].
- Tongue recognition for biometric personal identification.

Our work concerns the biometric applications of the tongue recognition and efficient feature extraction.

Tongue image analysis have received much attention in image analysis and computer vision. Tongue texture has many advantages for human identification and verification [9, 20]. The identification of people can be based on the texture features. As a biometric identifier, tongue image has the following properties:

- Tongue images are unique to every person. Texture features of the tongue are distinctive to each person,
- Texture features of an individual tongue are stable and unchangeable during the life of a person,
- The human tongue is well protected in mouth and is difficult to forge.

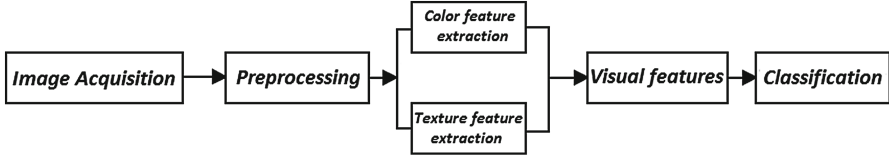


Fig. 1. Tongue recognition system



Fig. 2. Tongue images

Tongue recognition system is presented in Fig. 1 and it involves five major modules: tongue image acquisition, preprocessing, tongue feature extraction, visual features and classification.

Images which are considered in this paper are displayed in Fig. 2.

2 Preprocessing

Before performing feature extraction, the original tongue images are subjected to some image processing operations, such as:

1. Color conversion. The extraction of color features can be performed in different color spaces [17]. Each image is represented using three components of the color space. A color transformation that reduces the psychovisual redundancy and correlation of the image is highly desired. The YC_rC_b is an encoded non-linear RGB signal for image processing work. Color is represented by luminance, computed from nonlinear RGB [14], constructed as a weighted sum of the RGB values, and two color difference values C_r and C_b that are formed by subtracting luminance from RGB red and blue components. The two color

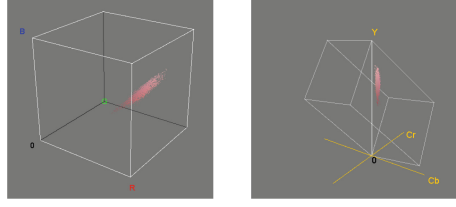


Fig. 3. RGB and YC_bC_r color spaces

spaces i.e. RGB and YC_bC_r were used for extraction color features (Fig. 3).

$$\begin{aligned} Y &= 0,299R + 0,587G + 0,114B \\ C_r &= 0,713(R - Y) \\ C_b &= 0,564(B - Y) \end{aligned} \quad (1)$$

or

$$\begin{bmatrix} Y \\ C_r \\ C_b \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ 0,500 & -0,419 & -0,081 \\ -0,169 & -0,331 & 0,500 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2)$$

An image histogram refers to the probability mass function of the image intensities. This is extended for color images to capture the joint probabilities of the intensities of the three color channels. More formally, the color histogram is defined by,

$$h_{A,B,C} = N \cdot Prob(A = a, B = b, C = c) \quad (3)$$

where A , B and C represent the three color channels (R, G, B or YC_rC_b) and N is the number of pixels in the image. Computationally, the color histogram is formed by discretizing the colors within an image and counting the number of pixels of each color.

2. Extraction of region of interest (ROI) from original tongue images. The tongue images are normalized with respect to position, orientation, scale, reflection, as follows.

The new invariant coordinates (x, y) of image pixels and the old coordinates (x', y') are related by

$$\begin{aligned} [x, y, 1] &= [x', y', 1] \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -i_0 & -j_0 & 1 \end{bmatrix} \\ &\times \begin{bmatrix} \frac{1}{\delta_x} & 0 & 0 \\ 0 & \frac{1}{\delta_y} & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \cos \beta & \sin \beta & 0 \\ -\sin \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (4)$$

where x_0, y_0 is the centroid of image; δ_x and δ_y represent standard deviation relative to variable x, y ; and β is an angle between the major axis of an object and the vertical line

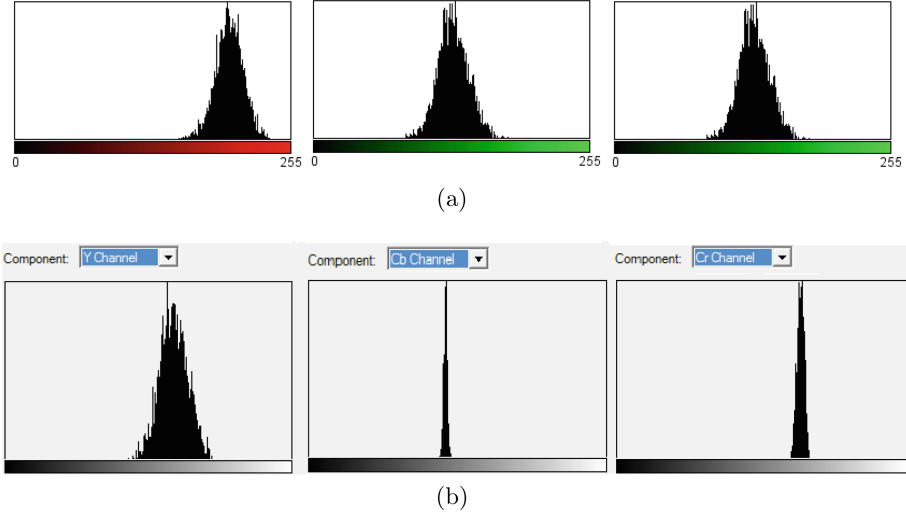


Fig. 4. Histograms: (a) RGB channels and (b) YC_bC_r channels

$$\tan 2\beta = \frac{2 \sum_x \sum_y (x - x_0)(y - y_0)}{\sum_x (x - x_0^2) - \sum_y (y - y_0^2)} \quad (5)$$

Next, the ROI 's tongue blocks are automatically selected on the centroid of tongue normalized images. The size of whole ROI is $w_x \times w_y$ where $w_x = (x_0 + \frac{K}{2}) - (x_0 - \frac{K}{2})$, $w_y = (y_0 + \frac{K}{2}) - (y_0 - \frac{K}{2})$ where $K = 128$ pixels (Fig. 5). Next, the ROI image is divided into the four sub-blocks. The size of sub-block is $\frac{K}{2} \times \frac{K}{2}$ pixels (Fig. 6).

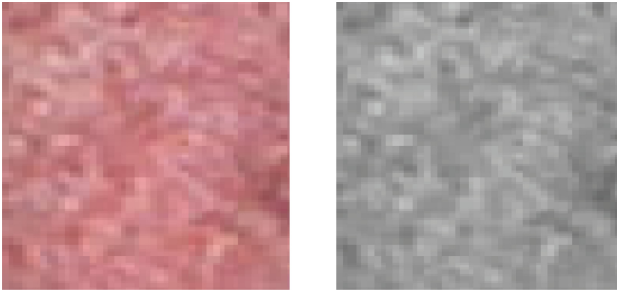


Fig. 5. Tongue ROI

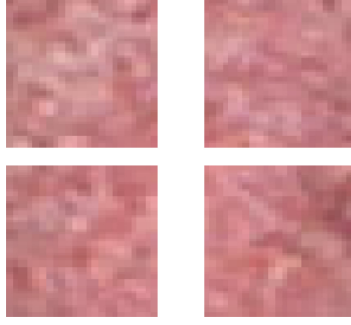


Fig. 6. Four *ROI*'s sub-blocks

3 Feature Extraction

3.1 Feature Extraction Based on Color Moments

The distribution of color was represented by color histograms (Fig. 4), and formed the image's feature vectors. The mathematical foundation of this approach is that any probability distribution is uniquely characterized by its moments.

Color moments have been successfully used in many image processing/biometrics systems. The first order (mean), the second (variance), the third order (skewness) and the fourth order (kurtosis) color moments have been proved to be efficient and effective in representing color distributions of images [16]. Mathematically, the first three moments are defined

$$\mu_c = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N f_c(x, y) \quad (6)$$

$$\sigma_c = \left(\frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N (f_c(x, y) - \mu_c)^2 \right)^{\frac{1}{2}} \quad (7)$$

$$s_c = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N \left[\frac{f_c(x, y) - \mu_c}{\sigma_c} \right]^3 \quad (8)$$

$$k_c = \left\{ \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N \left[\frac{f_c(x, y) - \mu_c}{\sigma_c} \right]^4 \right\} - 3 \quad (9)$$

where $f_c(x, y)$ is the value of the c -th color component of the image pixel (x, y) , and MN is the number of pixels in the image.

The color features are computed in RGB and YC_bC_r color spaces.

Since only 24 (four moments for each of the three color components in two color spaces) numbers are used to represent the color content of each image, color moments are a very compact representation compared to other color features (Table 1).

Table 1. Moments of color components

Image	Color components	Mean μ_c	Variance σ_c	Skewness s_c	Kurtosis k_c
Tongue ROI Fig. 5	R	198,72	13,07	0,201	0.267
	G	129,14	14,39	0,889	0,158
	B	132,95	14,44	-0,020	-0,013
Tongue ROI Fig. 5	Y	149,81	13,63	-0,030	0,184
	C_b	117,71	1,92	-0,074	0,389
	C_r	161,99	3,454	-0,074	-0,3

3.2 Gabor Filters for Feature Extraction

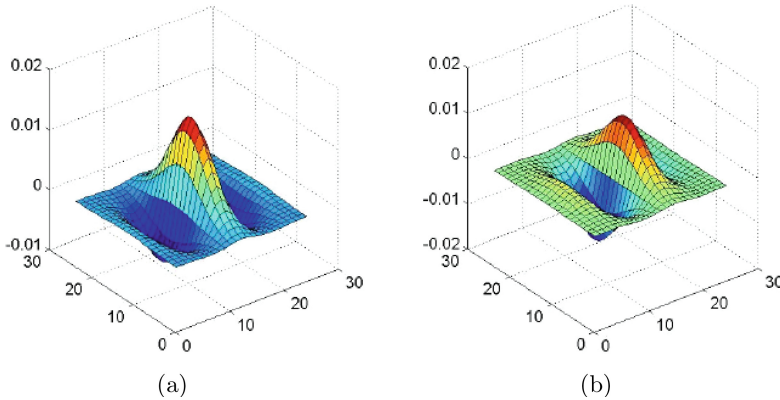
Gabor filters are a powerful tool to extract texture features and in the spatial domain is a complex exponential modulated by a Gaussian function. In the most general the Gabor filters are defined as follows [1, 13, 15].

The two-dimensional Gabor filter is defined as

$$Gab(x, y, W, \theta, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{\left[-\frac{1}{2}\left(\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right) + jW(x \cos \theta + y \sin \theta)\right]} \quad (10)$$

where $j = \sqrt{-1}$ and σ_x and σ_y are the scaling parameters of the filter, W is the radial frequency of the sinusoid and $\theta \in [0, \pi]$ specifies the orientation of the Gabor filters [7].

Figure 7 presents the real and imaginary parts of Gabor filters.

**Fig. 7.** The real and imaginary parts of Gabor filters

In our work we use a bank of filters built from the real part of Gabor expression called as even symmetric Gabor filter. Gabor filtered output of the image

is obtained by the convolution of the image with Gabor even function for each of the orientation/spatial frequency (scale) orientation (Fig. 8).

Given an image $F(x, y)$, we filter this image with $Gab(x, y, W, \theta, \sigma_x, \sigma_y)$

$$FGab(x, y, W, \theta, \sigma_x, \sigma_y) = \sum_k \sum_l F(x - k, y - l) * Gab(x, y, W, \theta, \sigma_x, \sigma_y) \quad (11)$$

The magnitudes of the Gabor filters responses are represented by three moments

$$\mu(W, \theta, \sigma_x, \sigma_y) = \frac{1}{XY} \sum_{x=1}^X \sum_{y=1}^Y FGab(x, y, W, \theta, \sigma_x, \sigma_y) \quad (12)$$

$$std(W, \theta, \sigma_x, \sigma_y) = \sqrt{\sum_{x=1}^X \sum_{y=1}^Y |FGab(x, y, W, \theta, \sigma_x, \sigma_y) - \mu(W, \theta, \sigma_x, \sigma_y)|^2} \quad (13)$$

$$Energy = \sum_{x=1}^X \sum_{y=1}^Y [FGab(x, y, W, \theta, \sigma_x, \sigma_y)]^2 \quad (14)$$

By selecting different center frequencies and orientations, we can obtain a family of Gabor kernels, which can then be used to extract features from an image. The feature vector is constructed using *mean* - $\mu(W, \theta, \sigma_x, \sigma_y)$, *standard deviation* - $std(W, \theta, \sigma_x, \sigma_y)$ and *energy* as feature components (Table 2).

We defined the vectors of features as follows:

$$FV = (Feature_{Color\ moments}, Feature_{Gabor}) \quad (15)$$

The first part of the FV contains the 24 color moments. The features in second part of FV are listed as follows $Feature_{Gabor} = ((\mu_1(x, y), std_1(x, y), Skew_1) \dots (\mu_t(x, y), std_t(x, y), Skew_t))$.

To reduce dimension of feature vector [12, 19], we use the Principle Component Analysis (PCA) algorithm to keep the most useful Gabor features.

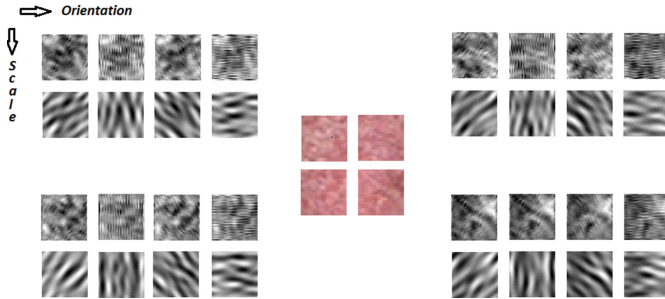


Fig. 8. Gabor images of tongue ROI 's

Table 2. Features of tongue *ROI*'s.

ROI 1				
Scale	Orientation	Energy	Mean	Std
2	45°	1968755.0	480.65308	35.064735
2	90°	1999251.1	488.09842	40.68314
2	135°	1968758.9	480.65402	34.23929
2	180°	1999252.2	488.0987	42.19372
8	45°	3.181323E7	7766.903	669.66876
8	90°	3.1827126E7	7770.2944	657.5264
8	135°	3.181321E7	7766.897	762.5107
8	180°	3.1827142E7	7770.2983	728.3237
ROI 2				
Scale	Orientation	Energy	Mean	Std
2	45°	1968755.0	480.65308	35.064735
2	90°	1999251.1	488.09842	40.68314
2	135°	1968758.9	480.65402	34.23929
2	180°	1999252.2	488.0987	42.19372
8	45°	3.181323E7	7766.903	669.66876
8	90°	3.1827126E7	7770.2944	657.5264
8	135°	3.181321E7	7766.897	762.5107
8	180°	3.1827142E7	7770.2983	728.3237
ROI 3				
Scale	Orientation	Energy	Mean	Std
2	45°	1867638.8	455.9665	36.981506
2	90°	1896568.8	463.02948	44.25503
2	135°	1867638.6	455.96646	36.304142
2	180°	1896567.2	463.0291	41.741203
8	45°	3.0179278E7	7367.988	707.88116
8	90°	3.0192456E7	7371.205	717.7458
8	135°	3.0179264E7	7367.984	774.104
8	180°	3.0192524E7	7371.221	614.612
ROI 4				
Scale	Orientation	Energy	Mean	Std
2	45°	1906920.6	465.5568	27.945246
2	90°	1936458.5	472.7682	35.911766
2	135°	1906919.8	465.55658	28.149256
2	180°	1936459.8	472.7685	35.161488
8	45°	3.0813984E7	7522.9453	786.99506
8	90°	3.08275E7	7526.245	728.3674
8	135°	3.0814002E7	7522.9497	562.7714
8	180°	3.082751E7	7526.2476	521.4142

Let $X = [x_1, x_2, \dots, x_n]$ denote an n -dimensional feature vector. The mean of the vector X and the total scatter covariance matrix of the vector X are defined as: $\bar{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$ and $S_X = \sum_{i=1}^n (x_i - \bar{\mu}) \cdot (x_i - \bar{\mu})^t$.

The *PCA* projection matrix S can be obtained by eigen-analysis of the covariance matrix S_X . We compute the eigenvalues of $S_X : \lambda_1 > \lambda_2 > \dots > \lambda_n$ and the eigenvectors of $S_X : s_1, s_2, \dots, s_n$. Thus $S_X s_i = \lambda_i s_i$, $i = 1, 2, \dots, m$. s_i is the i th largest eigenvector of S_X , $m \ll n$ and $S = [s_1, s_2, \dots, s_m]$.

Any vector x can be written as a linear combination of the eigenvectors (S is symmetric, s_1, s_2, \dots, s_n form a basis), i.e. $x = \sum_{i=1}^n b_i u_i$. For dimensionality reduction we choose only m largest eigen values, i.e. $x = \sum_{i=1}^m b_i u_i$. m is choose as follows: $\frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^n \lambda_i} > t$ where t is threshold.

By removing the principal components that contribute little to the variance, we project the entire feature vector to a lower dimensional space, but retain most of the information.

4 Conclusion

In the paper, are presented some approaches for tongue recognition from images. To evaluate the performance of tongue recognition methods we use own tongue database that consists 30 images. We proposed a method which combines the recognition results of Gabor filters and color moments features to tongue recognition. The proposed system will be evaluated on other tongue databases in the future study.

Acknowledgment. The research was supported by the UTP University of Sciences and Technology by the Grant BS01/2014.

References

1. Choras, R.S.: Iris-based person identification using Gabor wavelets and moments. In: Proceedings of 2009 International Conference on Biometrics and Kansei Engineering ICBAKE, pp. 55–59. CPS IEEE Computer Society (2009)
2. Choras, R.S.: Thermal Face Recognition, Image Processing & Communications Challenges 7. AISC, vol. 389, pp. 37–46. Springer, Berlin (2016)
3. Choras, R.S.: Tongue recognition from images. In: Proceedings of the First International Conference on Advances in Signal, Image and Video Processing SIGNAL 2016, Lisbon, Portugal, pp. 6–11 (2016)
4. Choras, R.S.: Vascular Biometry, Image Processing & Communications Challenges 6. AISC, vol. 313, pp. 21–28. Springer, Berlin (2015)
5. Huang, W., Yan, Z., Xu, J., Zhang, L.: Analysis of the tongue fur and tongue features by naive bayesian classifier. In: 2010 International Conference on Computer Application and System Modeling (ICCA SM 2010), vol. 4, pp. V4–304, IEEE, October 2010
6. Huang, B., Zhang, D., Li, Y., Zhang, H., Li, N.: Tongue coating image retrieval. In: Proceedings of the 3rd International Conference on Advanced Computer Control, Harbin, pp. 292–296 (2011)

7. Jain, A.K.: Fundamentals of a Digital Signal Processing. Prentice-Hall, Englewood Cliffs (1989)
8. Jain, A.K., Bolle, R., Pankanti, S.: Biometrics: Personal Identification in Networked Society. Kulwer Academic, London (1999)
9. Lahmiri, S.: Recognition of tongueprint textures for personal authentication: a wavelet approach. *J. Adv. Inf. Technol.* **3**(3), 168–175 (2012)
10. Li, C., Yuen, P.: Tongue image matching using color content. *Pattern Recogn.* **35**(2), 407–419 (2002)
11. Li, W., Hu, S., Yao, J., Song, H.: The separation framework of tongue coating and proper in traditional Chinese medicine. In: *Proceedings of the 7th International Conference Information, Communications and Signal Processing*, Macau, December, pp. 1–4 (2009)
12. Liu, C.J., Wechsler, H.: Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Trans. Image Process.* **11**(4), 467–476 (2002)
13. Liu, D.H., Lam, K.M., Shen, L.S.: Optimal sampling of gabor features for face recognition. *Pattern Recogn. Lett.* **25**(2), 267–276 (2004)
14. Poynton, C.A.: Frequently asked questions about colour (1995). <ftp://www.inforamp.net/pub/users/poynton/doc/colour/>
15. Shen, L., Bai, L.: Face recognition based on gabor features using kernel methods. In: *Proceedings 6th IEEE Conference on Face and Gesture Recognition*, Korea, pp. 170–175 (2004)
16. Stricker, M., Orengo, M.: Similarity of color images. In: *SPIE Storage and Retrieval for Image and Video Databases III*, vol. 2185, pp. 381–392 (1995)
17. Swain, M.J., Ballard, D.H.: Color indexing. *Int. J. Comput. Vis.* **7**(1), 11–32 (1991)
18. Wang, Y., Yang, J., Zhou, Y.: Region partition and feature matching based color recognition of tongue image. *Pattern Recogn. Lett.* **28**(1), 11–19 (2007)
19. Zhang, B.C., Shan, S.G., Chen, X.L., Gao, W.: Histogram of Gabor phase patterns (HGPP): a novel object representation approach for face recognition. *IEEE Trans. Image Process.* **16**, 57–68 (2007)
20. Zhang, D., Liu, Z., Yan, J., Shi, P.: Tongue-print: a novel biometrics pattern. In: Lee, S.-W., Li, S.Z. (eds.) *ICB 2007. LNCS*, vol. 4642, pp. 1174–1183. Springer, Heidelberg (2007). doi:[10.1007/978-3-540-74549-5_122](https://doi.org/10.1007/978-3-540-74549-5_122)

Image Processing and Communications Challenges 8
8th International Conference, IP&C 2016 Bydgoszcz,
Poland, September 2016 Proceedings
Choraś, R.S. (Ed.)
2017, XI, 280 p. 121 illus., Softcover
ISBN: 978-3-319-47273-7