An Extraction and Recognition of Tongue-Print Images for Biometrics Authentication System

Manoj Diwakar
Department of Computer Science & Technology, Dehradun Institute of Technology (DIT) Dehradun, Uttrakhand, India

Manish Maharshi
Department of Electronics & Communication Engineering, FET, Mody institute of technology and science Laxmangarh, Rajasthan, India

ABSTRACT

Biometrics authentication system promises to deliver a level of uniqueness to identification applications that uses Tongue print, fingerprint or facial scans and iris or voice recognition to identify users. The tongue is a unique organ in that it can be stuck out of mouth for inspection, in this act offering a proof of life, and yet it is otherwise well protected in the mouth and is difficult to forge. The tongue also presents both geometric shape information and physiological texture information which are potentially useful in identity verification applications. Tongue biometric can function as an extremely reliable means for personal identification and act as a general biometric in all applications. This paper demonstrates two phases. First, find out spots on the tongue with the help of histogram and second phase to extract image tongue and recognize from the tongue-image database.

Keywords: tongue-print; biometric; Histogram; pattern recognition system; Tongue code; Biometrics authentication system, SIFT Feature Extraction.

1. Introduction

A biometric scanning device takes a user's biometric data, such as an iris pattern or tongue-print scan, and converts it into digital information a computer can interpret and verify. Since it is more difficult and it is unlikely that a user will misplace or misuse his or her biometric data, this form of technology a greater level of assurance than other methods of identification. Biometrics can be used for both physical access to corporate buildings and internal access to enterprise computers and systems.

A biometric system is a real-time identification system which identifies a person by measuring a particular physical or behavioural characteristic and later comparing it to a library of characteristics belonging to many people [1]. Fingerprint and other biometric devices consist of a reader or scanning device, software that converts the scanned information into digital form, and wherever the data is to be analyzed, a database that stores the biometric data for comparison with previous records. When converting the biometric input, the software identifies specific points of data as match points. The match points are processed using an algorithm into a value that can be compared with biometric data scanned when a user tries to gain access.

Thus biometric devices can be explained with a 3-step procedure. They are -

(i) A sensor takes an observation. The type of sensor and its observation depend on the type of biometrics device used. This observation gives us a Biometric Signature of the individual.

(ii) A computer algorithm normalizes the biometric signature so that it is in the same format (size, resolution, etc.) as the signatures on the system's database. The normalization of the biometric signature gives us a Normalized Signature of the individual.

(iii) A matcher compares the normalized signature with the set (or sub-set) of normalized signatures on the system's database and provides a Similarity score that compares the individual's normalized signature with each signature in the database set (or sub-set). What is then done with the similarity scores depends on the biometric system's application.

Thus, a biometric system is essentially a *pattern recognition system,* which makes a personal identification by determining the authenticity of a specific physiological or behavioral characteristic possessed by the user.

2. Tongue Biometric

To begin with, the tongue is unique to each person in its shape (see Figure 1) and in its surface textures (see Figure 2). Second, the tongue is the only internal organ that can quite normally and easily be exposed for inspection.

System design includes the basic architecture, algorithms to be implemented and the flow diagram of the complete system being developed [2]. Matching feature points between images is one of the most fundamental issues in computer vision tasks. Tongue recognition is followed by two Feature of Tongue that is-

- Shape parameter is calculated by using control points which gives the prominent outlines of shape feature of tongue.

Fig.1 Different shape of Tongue

- Texture Feature is calculate by SIFT Algorithm which is Pre-processed by Histogram Equalization

Fig.2 Different texture of Tongue

3. Shape feature extraction Algorithm (control Points) for Tongue Images

The shape of the tongue (to be represented as a shape vector) is measured using a set of control points. The control points are P1, P2, …, P11, Ptip and Pm (shown in Fig. 3).These control points demarcate important areas of the ROI (here, the part below the segment P1,P2 L).

The following describes how five measures, length, bend, thickness, width of the tongue, and the curvature of its tip, are formed as our measurement vectors:
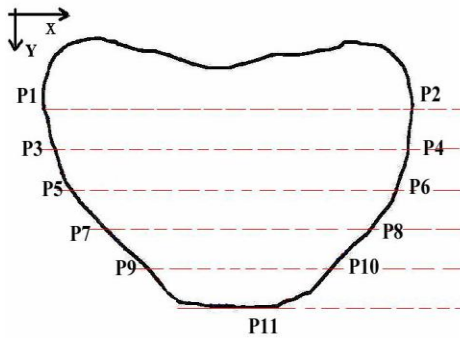

Fig.3 The shape feature model for the frontal and profile view images (Control Points)

Width: We define four segments ($L_{p3,p4}$, $L_{p5,p6}$, L , $L_{p7,p8}$,$L_{p9,p10}$) that are parallel to the segment $L_{p1,p2}$ in the regions of interest mentioned above. In addition, these segments follow the rule formularized by Eq. (1):

$$d(L_{p1,p2},L_{p3,p4})=d(L_{p3,p4} \quad , \quad L_{p5,p6},)=d(L_{p5,p6}, \quad L \quad , L_{p7,p8})=d(L_{p7,p8},L_{p9,p10}) \qquad (1)$$

Where d(i)represents the distance between two parallel segments. We then use the length of these five segments to construct the width vector W.

Length: The length of the tongue in the profile view is defined by the distance between Ptip and Pm (Ptip denotes the tip of the tongue, and Pm denotes the corner of the mouth, as shown in Fig. 3 as follows Length=P_{tip}−P_m (2)

4. Extraction of Tongue Image from Face

4.1 System Level Design:

The computerised image classification uses a tongue image acquiring device and tongue diagnostic software.
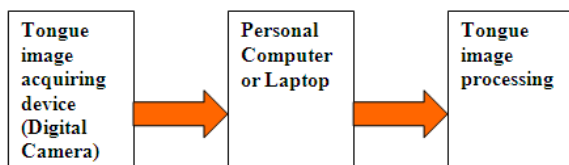

Fig. 4 Simplified computerised tongue-image classification

For the tongue image acquisition, a digital camera is used. It has an extreme sensor super CCD EXR. The new Super CCD EXR provides superior picture quality, enabling a "3-in-1" sensor combination of "Fine Capture Technology" (High Resolution) that reproduces the smallest detail of the subject in bright

beautiful colours, "Pixel Fusion Technology" (High Sensitivity & Low Noise) which Reveals subtle shades and details of the subject even in high contrast lighting, and "Dual Capture Technology" (Wide Dynamic Range) capturing the subject with smooth texture and natural brightness free from noise [2]. With an innovative colour filter array and image processing technology, the EXR ensures an advanced reproduction in imaging with exceptionally balanced quality. With the sensor and every settings automatically optimized for the scene, the image can be captured just as human eyes sees them.

The height and distance of the digital camera is taken into considerations. This is to ensure good clarity and accuracy of the tongue image. The height of the camera is supported by a tripod stand of 14 cm. The person's head is supported by a head rest of height 15 cm. The distance between the head rest and camera is 9cm.

The environment is also taken into consideration. The tongue images should be taken in an enclose area, to ensure that the lighting of the tongue image in control. This is also to ensure that the colour accuracy of the tongue image is not affected in the diagnosis. If lighting of the place is not enough for the colour accuracy of the tongue image. Flash will be provided.

The tongue image extraction module is explained in detail in the next part for algorithm design and other subsequent sections.
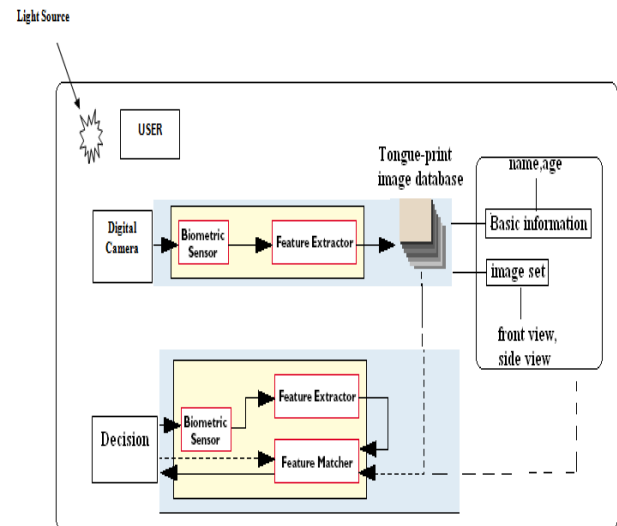

Fig 5: Biometric tongue-print system setup

4.2 Tongue-print image management:

To manage the data of tongue image, use proper data-structure. Set the data with their basic information. The quality of the tongue-print image database is evaluated through the validation experiments. The validation study addresses the question of whether the samples of the database can be used to for verification and whether they can represent the general feature extraction [3]. To do so, we conducted an analysis and test against our tongue-print image database.

4.3 Algorithm Level Design:

The tongue image extraction module consist of capture tongue images, images of tongue in database, spots on the tongue, display scale setting, read pixel, zoom in and zoom out.

Steps:

1. Create a set-up to take the tongue-images.

2. Capture the tongue images with the help of high sensor digital camera.

3. Store that image into a temporary location

4. Process on that image to find the spots on tongue.

5. Whether tongue image has a spot or no spot, process to find the information of image with the help of histogram.

6. With the help of histogram matching technique and shape matching, match with the stored tongue-image database.

7. Because of spot on tongue, 90% matching considered, tongue image is matched.

8. If it is not matched, that image is stored in tongue image database.

5. Extraction of Spots on tongue-print

Histogram technique:

Histogram equalization is a technique to enhance the images for achieving more image information and image data and getting better results by image processing [3].

Histogram equalization technique based on distribution of the intensity values of the pixels -which consist of three separate colour values, i.e. RGB- onto the desired intensity interval.

In this step we applied histogram equalization technique on the frames of the specified gestures in our database in order to make more accurate matching.
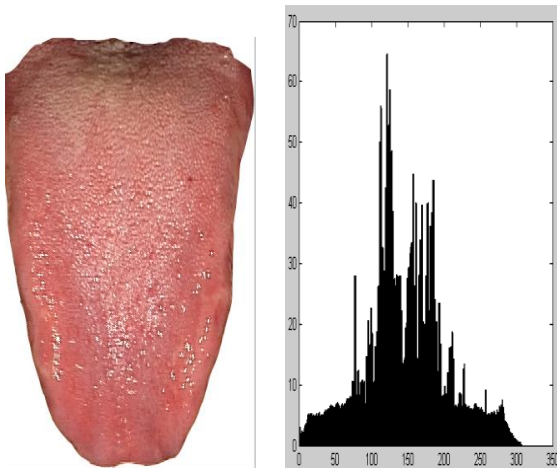

Fig 6: Input Image and its Histogram

The histograms of the frames with the intensity level in the range of 0 to 255 are evaluated by this function:

$$h\left(r_k\right)=n_k$$

k^{th} intensity value — Number of pixels in the image with intensity r_k

r_k is the intensity value between 0 and 255, n_k is number of the pixels which has the value of r_k. This evaluated function is normalized by the total number of pixels in the frame.

$$p\left(r_k\right)=\frac{n_k}{MN}, \quad k=0,1,...L-1$$

$p\left(r_k\right)$ is the probability of occurrence of the kth intensity value in the image that has the width of M and the height of N as pixel.
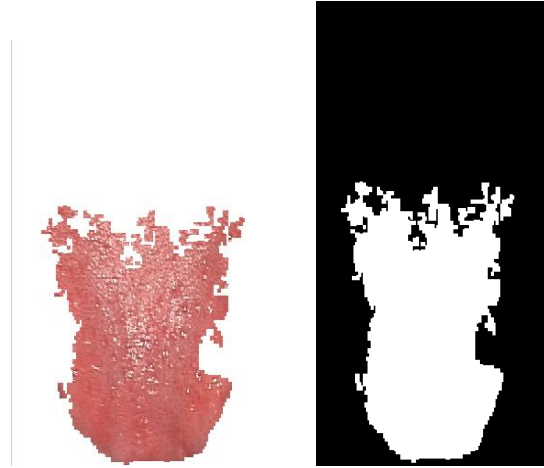

Fig 7: Processed Image

After calculating the probability of occurrence of each intensity value, the new value of that intensity is evaluated by this histogram equalization transform function:

$$s_k=T(r_k)=(L-1)\sum_{j=0}^{k}p_r(r_j)=\frac{(L-1)}{MN}\sum_{j=0}^{k}n_j \quad k=0,1,2,...,L-1$$

Hence the value of r_k is mapped to a new intensity value of s_k by summing all the probability and multiplying the result with the image size till the specified value (k). The produced values should be rounded to an integer value as a matter of digital image processing.
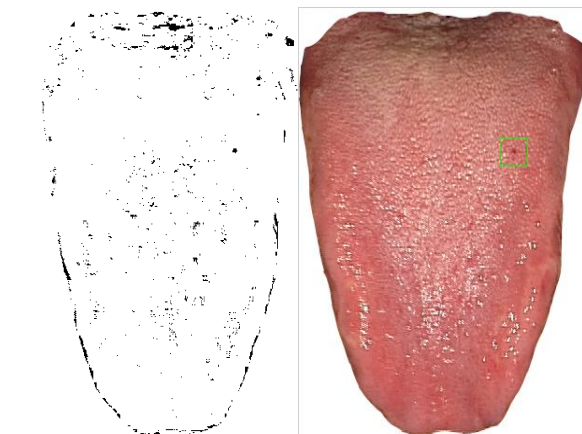

Fig 8: Out-put Image

6. Design and Implementation:

A tongue-print in its narrow sense is an impression left by the friction ridges of a human face. In a wider use of the term, tongue-prints are the traces of an impression from the friction ridges of any part of a human face. Tongue-prints are one of the many forms of biometrics that can be used to identify a person

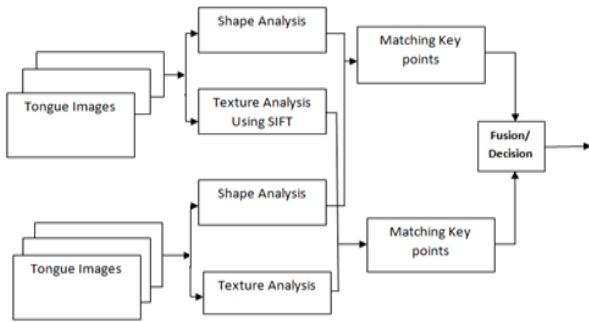and an automated process can be developed that can verify and find a match between two tongue-prints.

Fig 9: Architecture of Tongue Recognition Approach

One of the methods towards verifying a tongue-print is to find out the minutiae in the image of a tongue-print. The two features of a minutia are ridge endings and bifurcation points. A ridge ending is the point where a ridge ends, and a bifurcation point is where two ridges meet and continue as a single ridge (or conversely a single ridge splits into two ridges).

6.1 SIFT Feature Extraction

SIFT was proposed to detect local image features invariant to image scaling, translation, and rotation. This is achieved by selecting key locations at local maxima and minima of a difference of Gaussian function applied in scale space, which is constructed by successively down sampling the input image [4]. Maxima and minima of this scale space function are determined by comparing each pixel to its neighbours.

i. Scale-space extreme detection: The first stage of computation searches over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation.

ii. Key point localization: At each candidate location, a detailed model is fit to determine location and scale. Key points are selected based on measures of their stability.

iii. Orientation assignment: One or more orientations are assigned to each key point location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations.

iv. Key point descriptor: The local image gradients are measured at the selected scale in the region around each key point. These are transformed into a representation that allows for significant levels of local shape distortion and changes in illumination.

6.2 Matching Process:

Point wise matching is used for matching the objects in pattern recognition based on SIFTS features. Generally, Euclidean distance is used as distance metric in point wise Matching. In order to improve the efficiency, it is cheaper to compute dot product between unit vectors rather than Euclidean distance. Matching is performed by comparing each local extreme a based on the associated descriptors [5]. Suppose we want to match two images I1 and I2. Given a feature point p11 in I1, its closest point p21, second closest point p22, and their distance ratio is calculated. A match is accepted only if its distance is less than the distance ratio (empirically chosen distance ratio is 0.6) times the distance to the second closest match. The matching score between two images can be decided based on the number of matching points and their geometric configuration.

6.3 Algorithm:

Assuming that the target object is not deformed, all features of the target image can be considered as being affected with constant scaling and rotational factors. This can be used to optimize the SIFT-feature matching phase where the outliers' rejection stage of the original SIFT-method is integrated into the SIFT-feature matching stage.

6.3.1 Scaling factor calculation:

Using the SIFT-operator, the two object images (model and test) are transformed into two SIFT-image feature sets. These two feature sets are divided into subsets according to the octaves in which the feature arise [7]. Hence, there is a separate subset for each image octave as shown in figure 10. To carry out the proposed new strategy Improved SIFT-Features matching for Object Recognition
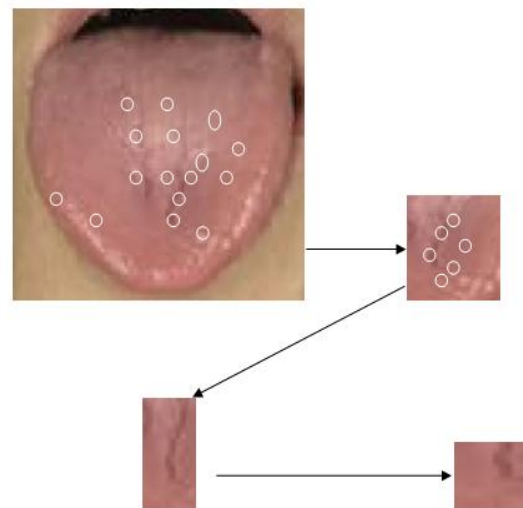

Fig. 10: Steps of the procedure for scale factor calculation.

The features subsets obtained are arranged so that a subset of the model image feature set is aligned with an appropriate subset of the test image feature set [6]. The process of alignment of the model image subsets with the test image subsets is indicated with arrows in Fig 11. The alignment process is performed through the (n+m-1) steps, where n and m are the total number of octaves (subsets) corresponding to the model and test image respectively. For each step all pairs of aligned subsets must have the same ratio ν defined as ν: $2^{0_1}/2^{0_2}$ where 0_1 and 0_2 are the octaves of the model image subset and the test image subset respectively. At every step, the total number of positive matches is determined for each aligned subsets pair. In order to realize the procedure mathematically, a quality-integer function F(x) is defined as:

$$F(x) = \sum Z \ (M_1^{n-1-x+j,} \ M_2^j)\ldots.if(x<n) \qquad (3)$$

$$\sum (M_1^{n-1-x+j,} \ M_2^j)\ldots if \ (n<x<n)$$

m>=n and x€[0,m+n-2)   (4)

$$\sum (M_1^{n-1-x+j}, M_2^{x-m+1+j})\dots \text{if}(x>=m) \quad (5)$$

Where limit of $\sum$ ranging from j=0 to j=x for Eq.3, from j=0 to j=n-1 for Eq.4 and from   j=0 to j=m+n-2-x for Eq.5
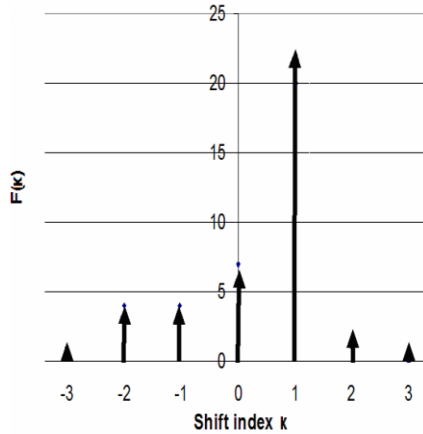

Fig.11 Shift index by which Sift Descript key points

6.3.2   Retrieval of the correct matches:

Among all found matches it can happen that a lot of correct matches exceed Lowe's   threshold τ. In order to retrieve these correct matches, the ratio between the Euclidean distance to the nearest and the second nearest feature neighbour must be reduced [8]. This can be done either by reducing the smallest distance D1 (F1i , F2j0) or by increasing the next smallest distance D 2 (F1i , F2j1) . In practice, the first alternative is impossible while the enlargement of next smallest distance can be achieved by limiting the search area for both the nearest and next nearest feature to the feature F1i within a specified domain. For a better explanation of this idea, suppose that a feature F1i from the model image feature set is correctly assigned to the feature F2j0 from the test image feature set. Also, suppose that F2j1 is the second nearest feature to the F1i while F2j2 is the second nearest feature to it when the search is limited only to the octave in which the F2j0 is found.   Since $D_2 (F_1^i, F_2^{j1}) \leq D_3 (F_1^i, F_2^{j2})$ always holds the following:

$D_1 (F_1^i, F_2^{j0}) / D_2 (F_1^i, F_2^{j1}) \geq D_1 (F_1^i, F_2^{j0}) / D_3 (F_1^i, F_2^{j2})$ is obtained. Thus, by reducing the search area it is possible to decrease the ratio related to the feature F1i and make it less than threshold τ . In this way the number of correct matches is increased.

Minutiae Extraction involves the following three steps

   i.   Pre-Processing of Image

   ii.   Minutiae Extraction

   iii.   Post-Processing

## 6.4 Implementation on MATLAB:

6.4.1 Pre-Processing of Image

The pre-processing of the image involves taking the image and applying various processes on the image so that it can easily be processed to find out the ridge endings and bifurcation points [9]. The two major steps in the pre-processing are:
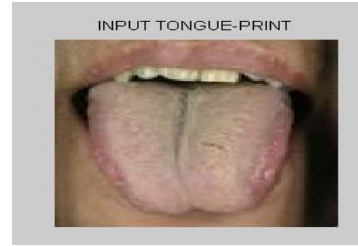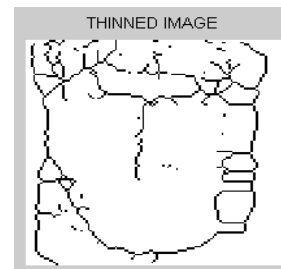

6.4.2 Binarizing:

In this step the colours of the image are binarized so that the output image consists of only two colours, black and white.
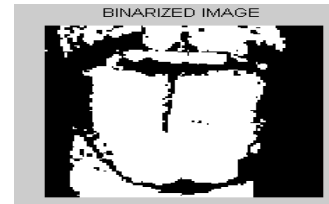
6.4.3 Thinning:

The width of the ridges is reduced to the size of one pixel as this makes the minutiae extraction process easily applicable [10].

Other steps involve the removal of unwanted pixels.

6.4.4 Minutiae Extraction

The minutiae are extracted by processing the whole image pixel by pixel using a 3x3 matrix with the centre element being the focus of each process.


6.4.5 Post-Processing

As you can see that in the above diagram, too many minutiae have been detected on the edge of the image. This is because those points passed the processing test of being ridge ends, but actually they are not ridge ends but only the points beyond which the image was not scanned [11]. So now these false minutiae have to be filtered out and removed.
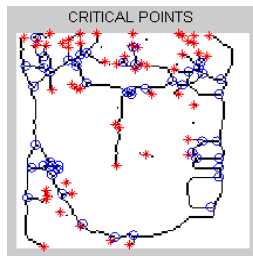
Other minutia points that have to be removed are ones that are too close to each other. They also do not qualify as valid minutiae. Any pair of ridge or bifurcation points is removed if it is found that the distance between the two is smaller than a certain number of pixels (taken as 6 in the algorithm) [12].
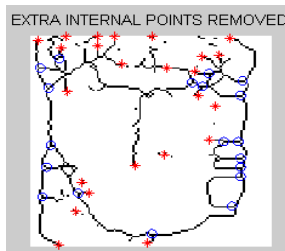
Next we need to take care of the minutiae present on the edges and for this we find ROI or Region of Interest. This region is simply a mask that is applied on the above image that removes the focus on the edge minutiae.
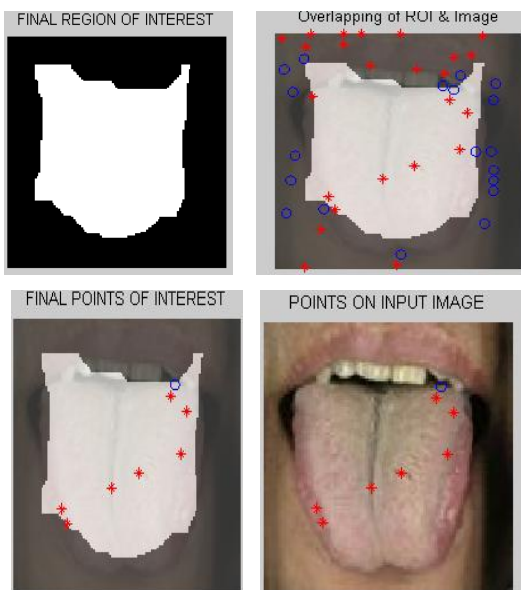

The coordinates of the final minutiae are saved in the data base in a new text file.

6.5 Matching the tongue-prints:

The algorithms that we have applied to match two tongue-prints involves calculating Euclidian distance between each ridge end and all other ridge ends and similarly between all bifurcations and then taking the average for both and at last add the results for achieving high accuracy and precision level [13]. This process is applied on both the input images and the results of the two images are compared to give the percentage match between the two images.

Match Percentage: 100.000000 percent
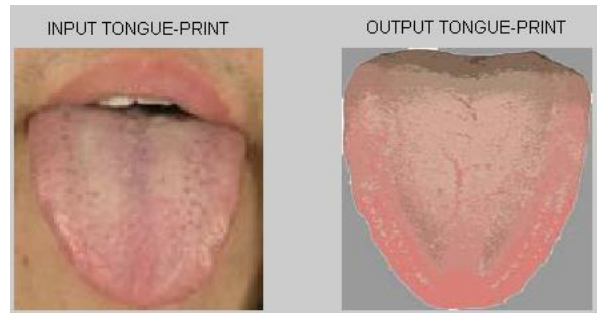
Match Percentage: 48.3576 percent

6.6 Conclusion & Future Work

Although so many biometrics have been used and developed, to our best knowledge, there is not so work has been done in tongue print recognition system and its use in any application yet. With increasing identity fraud and emphasis on security, there is a growing and urgent need to efficiently identify humans both locally and remotely on a routine basis. The human tongue promises to deliver a level of uniqueness to identification applications that other biometrics cannot match in context of that it is well protected in mouth and is difficult to forge. As the only internal organ that can be protruded from the body, the human tongue is well protected and is immune to forgery. The explicit features of the tongue cannot be reverse engineered, meaning that tongue verification protects the privacy of users better than other biometrics. The paper [14] shows that tongue-prints of the human tongue qualify as a feasible new member of the biometrics family. Matching a given image with one or many others is a key task in many computer vision applications such as object recognition, images stitching and 3D stereo reconstruction. for shape feature extraction of tongue algorithm of collecting points give efficient template for tongue image while for texture analysis normalized histogram with SIFT (Scale Invariant Feature Transform) is successfully used .combining both feature extraction techniques template Tongue verification matching is done.

To make tongue recognition more secure, Make tongue biometric Template with three views of tongue image that is Left lateral view, Right lateral view and Profile view (because tongue is a non rigid organ and it's very difficult for a person to keep it straight).

For make it more simple Tongue should be extracted from face in a pre-processing stage with more efficient and simple algorithms like segmentation, it will work better if it could be done online in verification process. Since human tongue is non-rigid when it moves, we expect to collect the video of tongue to extract some rules of its movements.

7. References

[1] Zhi Liu, Jing-Qi Yan, David Zhang, Qun-Lin Tang; "A Tongue-Print Image Database For Recognition", in proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, August 2007,pp.19-22.

[2] J.Sheeba Rani, D.Devaraj, R. Sukanesh; "A Novel Feature Extraction Technique for Face Recognition", proceedings of International Conference on Computational Intelligence and Multimedia Applications 2007, pp.431-435.

[3] P.Ekman, W. Friesen, "Constants across cultures in the faceand emotion,*" Journal of Personality and Social Psychology*,1971. **17**(2): pp. 124-129.

[4] Biometrics, Personal Identification in Networked Society, A.Jain, R. Bolle, and S. Pankanti, eds., Kluwer AcademicPublishers, Boston, July 1999.

[5] Chandrashaker, Siv Aramakrishna and Gordon Lee; "Facial information Retrieval Using Component-based Classification and Scale Invariance", World Automation Congress (WAC), 2010,pp.1-6 .

[6] Chao Li, Armando Barreto; "an Integrated 3D Face Expression Recognition Approach." In proceeding of ICASSP 2006 vol 3 pp.1132-1135

[7] Giovanni Garibotto ElsagDatamat spa; "Video Surveillance and Biometric Technology Applications" in proceedings of sixth IEEE International conference Advanced Video and Signal Based Surveillance 2009,pp.288.

[8] Michael Negin ,Thomas A.Chmielewski, Jr.Marcos Salganicoff ,Theodore A.Camus,Ulf M. Cahnvon Seelen, Péter L.Venetianer, Guanghua G.Zhang. "An Iris Biometric System for Public and Personal Use, of The Software Engineering Coordinating Committee Leading the next generation of software Feb. 2000".

[9] Wangmeng Zuo, Kuanquan Wang, Zhang, D., Hongshi Zhang, Combination of polar edge detection and active contour model for automated tongue segmentation, Proceedings of Third International Conference of Image and Graphics, Dec. 2004,pp.270 -273.

[10] Liu Zhi, Jing-qi Yan, Tao Zhou, Qun-lin Tang, Tongue Shape Detection Based on B-Spline ICMLC2006, Aug. 2006, Vol. 6, pp. 3829-3832.

[11] Zheng Jun, Zhu Jing, Image matching based on adaptive genetic algorithm, Chinese Journal of Zhejiang University:Engineering Science,2003.

[12] J.D. Daugman, "High-Confidence Visual Recognition of Persons by a Test of Statistical Independence," IEEE Trans. Pattern Matching and Machine Intelligence,Nov. 1993, pp. 1,148-1,160.

[13] Stephen Milborrow, Fred Nicolls, "Locating Facial Features with an Extended Active Shape Model," Lecture notes in computer science,2008.

[14] Jin-Woong Park, Sun-Kyung Kang, Sung- Tae Jung in" Tongue Diagnosis System Based on PCA and SVM" published in " World Academy of Science, Engineering and Technology 60 2011".