Lip prints, Image processing,
Clustering techniques,
Data classification

Krzysztof WROBEL[1], Wojciech FROELICH[1]

RECOGNITION OF LIP PRINTS USING FUZZY C-MEANS CLUSTERING

In this paper a new method for lip print recognition is proposed. The proposed approach is based on Fuzzy c-Means clustering of the characteristics features of lip prints. First, the Hough transform is applied for the recognition of the characteristic features within lip prints, then Fuzzy c-Means clustering is performed to cluster those features. The proposed algorithm applies the results of clustering to find an unknown image withing the collected repository of lip prints. Instead of comparing all pairs of individual characteristic features, the proposed algorithm uses the representatives of clusters for the comparison of images. The advantage of using the proposed method is its increased tolerance to the noise in data and thus the increased efficiency of the recognition. The effectiveness of presented method has been verified experimentally using real-world images. The results are satisfactory and suggest the possibility of using the method in forensic identification systems.

1. Introduction

Recognition of lip prints (Cheiloscopy) is an important problem considered in criminology. The recognition of lip prints is a specific research domain relatively weakly connected with the recognition of finger prints. Contrarty to the recognition of finger prints the recognition of lip prints is weakly explored and a very small number of research papers addresses this problem.

Edmond Locard, a French criminologist, proposed in 1932, as the first, to use images of human lips for forensic identification. In 1970, Yasuo Tsuchihasi and Kazao Suzuki, Japanese scientists, carried out studies on the use of lip images for identification of persons [19], [20].

Up to 1145 individual features can be found on a single lip print [6], [13], [15], [17]. They form a unique pattern, which is different for each person. For comparison, only up to 100 individual features can be observed on a fingerprint [1], [4]. From the studies published in [14] it appears that the pattern of lip grooves is invariable and stable in people aged 21–50. In people over 50, changes in the pattern of rubor labiorum are observed, even in 5–year periods. This is caused by aging and flabbiness of facial muscles. Nevertheless, from the viewpoint of the forensic procedure, the time interval from the moment of securing a lip print at a crime scene to the moment of arresting a suspect on this basis generally does not exceed 6 months. This guarantees the invariability of rubor labiorum and allows using cheiloscopy for personal identification.

A practical classification of lip print images was presented by Jerzy Kasprzak, a Polish criminologist [13], [14]. On the basis of the studies carried out on a sample of 1500 lip

[1]Institute of Computer Science, University of Silesia, ul. Bedzinska 39, 41-200 Sosnowiec, Poland

prints, he distinguished 23 types of individual features that occur in a lip print pattern. He also proposed a division of lip print patterns into four groups: forked pattern - consisting of forks, linear pattern - consisting of straight lines, reticular pattern - consisting of intersections of lines forming a reticulum, and indefinite pattern that occurs when other patterns cannot be distinguished.

The contribution of this study is a new algorithm for the recognition of lip print images. In this algorithm we apply Hough transform [7], [12], the algorithm for the detection of segments and the Fuzzy c-Means clustering. However, our algorithm is completely new. It is presented for the first time in this paper.

The Hough transform is applied for the recognition of the characteristic features within lip prints, then Fuzzy c-Means clustering is performed to cluster those features. Our algorithm applies the results of clustering to find the unknown image within the collected repository of lip prints. Instead of comparing all pairs of individual characteristic features, the proposed algorithm uses the centers of clusters for the comparison of images. The advantage of using the proposed method is its increased tolerance to the noise in data and thus the increased efficiency of the recognition.

The reminder of the paper is organized as follows. Background knowledge on Hough transform, the algorithm for the identification of segments and basic notions of Fuzzy c-Means clustering are given in Section 2. The contribution of this study is presented in Section 3. The proposed approach has been validated and evaluated using real-world lip print images. The description of experiments is provided in Section 4. Section 5 concludes the paper.

2. Background knowledge

In this section we provide information on the background knowledge applied in our study: Hough transform, the known algorithm for the detection of segments and Fuzzy c-Means clustering.

2.1. Hough Transform

The Hough transform (HT) is a well known and documented method used in image processing [5], [9], [16]. It was introduced by Paul Hough in 1962 as a method for detecting patterns in binary images [10], [11]. Initially, the main purpose of the Hough transform was detection of straight lines in digital images. Later, it was used also for detecting other geometric objects that can be described by parametric equations, such as circles, ellipses or parabolas [7]. In 1981, a modification of the Hough transform was introduced by Dan H. Ballard [2]. This modification is called the Hough transform for irregular objects or Generalized Hough Transform.

HT is described in detail in [7], [12]. In our work we have used a standard HT where parameters of lines are detected in a normal form:

$$\zeta = xcos(\alpha) + ysin(\alpha). \tag{1}$$

As a result of using the Hough transform, a list of straight lines located in the image is obtained. These straight lines are defined by the parameters (α, ζ). Segments of different lengths can be located on a single straight line. An example of such a situation is shown in Fig. 1.

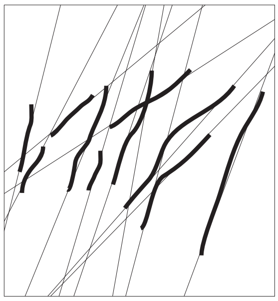


Fig. 1. Segments located on the straight lines.

2.2. THE ALGORITHM FOR FINDING SEGMENTS

On the basis of the applied Hough transform it is possible to find segments within the image. The algorithm and the idea of the algorithm for finding segments in binary image is presented below.

Input: Binary image I, a list L containing parameters of straight lines, $t \in \mathbb{N}$ - threshold.
Output: Set of segments.

For (All lines from set L) {
  For (Every point of the image I) {
    Check whether the pixel corresponding to the selected point is activated.
    If the pixel is activated then it is selected as belonging to the segment.
  }
  Count all points belonging to the selected line.
  Accept those segments with the lengths above the threshold t.
}
return (Set of segments.)

As a result of using the algorithm for finding segments, each analyzed image is described by a set of segments. Each segment is defined by the following attributes:

- the coordinates of the beginning of the segment,
- the coordinates of the ending of the segment,
- the length of the segment,
- the angle of inclination of the segment in relation to the X axis.

2.3. INTRODUCTION TO FUZZY c-MEANS CLUSTERING

The Fuzzy c-Means (FCM) algorithm was proposed by Dunn [8] and has been improved by Bezdek [3]. It assumes that a single instance of data may belong to two or more clusters.

Let us assume that $x_i \in X^d$ is the data element in d-dimensional space and let c_j be a d-dimension center of the cluster. To partition all data to clusters represented by c_j an iterative optimization of the objective function (2) is performed:

$$J_m = \sum_{i}^{N} \sum_{j}^{C} u_{ij}^m ||x_i - c_j||^2, \tag{2}$$

where m is any real number $m \geq 1$, u_{ij} is the degree of membership of x_i in the j^{th} cluster, N is the number of the considered data instances, C is the number of clusters and $||*||$ is any norm measuring the similarity between the data and the center of the cluster. In this study we apply the euclidean distance for that purpose.

The membership function is updated by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{||x_i - c_j||}{||x_i - c_k||} \right)^{\frac{2}{m-1}}}, \tag{3}$$

The cluster center is updated by:

$$c_j = \frac{\sum_i^N u_{ij}^m x_i}{\sum_i^N u_{ij}^m}. \tag{4}$$

The iteration stops when the following condition holds: $max_{ij}(|u_{ij}^{k+1} - u_{ij}^{k}|) < \epsilon$, where $\epsilon \in [0, 1]$ is the parameter and k is the iteration step.

As a result, each data element x_i is associated with each cluster by the membership level. This membership level indicates the strength of the association between x_i and the cluster center c_j.

3. Fuzzy c-Means clustering for lip print recognition

The contribution of this paper is an algorithm for the recognition of the unknown lip print image within the collected repository of lip print images. The proposed algorithm assumes the transformation of all images already assigned to persons (historical data) to the corresponding sets of characteristic features recognized within them. In this Section we formalize the problem and on that basis, we describe the proposed algorithm.

Let U be the set of lip print images that are already assigned to the known persons and let u_x be the image that is still not classified (the test image). First, for u_x and for every $u_j \in U$ a preprocessing is performed resulting in corresponding to every image set of segments (characteristic features). In this way we obtain F_x and F_j sets of features corresponding to the images u_x and u_j respectively.

The proposed approach is given as Algorithm 1.

Algorithm 1: Recognition of lip print image.

Input: The sequence of images: $U = \{u_1, u_2, ..., u_{card(U)}\}$,
where $u_x \in U$ is assumed as unclassified.
Output: The recognized image $u_s \in U$.

```
Preprocessing u_j ∈ U;
j = 1;
While (j ≤ card(U), j ≠ x) {
   d_j = 0;
   K_j = FuzzycMeans(F_j, c);
   While (k ≤ card(F_x)){
      d_min(f_k) = K_j(f_k ∈ F_x);
      d_j = d_j + d_min(f_k);
      k = k + 1;
   }
```

$j = j + 1;$
}
$s = \mathrm{argmin}_j(d_j);$
return u_s;

All sets F_j are clustered using Fuzzy c-Means function $FuzzycMeans(F_j, c)$, where c is the parameter determining the number of clusters. The result is a lazy classifier denoted as K_j.

Then, every feature $f_k \in F_x$ is classified using classifiers K_j. As an additional outcome the distances $d(f_k)$ of every f_k to the prototypes from are calculated. We denote $d_{min}(f_k) = K_j(f_k \in F_x)$ as the shortest of those distances, indicating the cluster (prototype) that is most strongly associated with the feature f_k.

The procedure is repeated for every feature $f_k \in F_x$ and image $u_j \in U$. Then for every image $u_j \in U$ we sum all distances $d_{min}(f_k)$ as $d_j = \sum_k^{card(F_x)} d_{min}(f_k)$. Afterward we select from U the image $u_s \in U$ with the lowest summarized distance. The index of that image in U is $s = \mathrm{argmin}_j(d_j)$.

4. Experiments

The proposed algorithm was validated and evaluated using real-world lip print images. The aquisition of images was the first step of the experimental setup.

4.1. Experimental setup

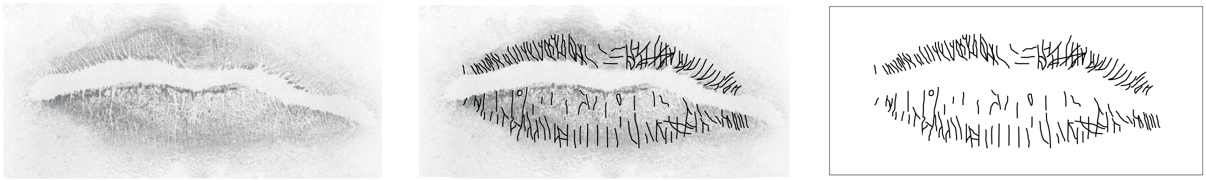
The acquisition of lip print images was performed using a traditional method with the use of a special tools: fingerprinting roller (Fig. 2a), magnetic powder (Fig. 2b), magnetic powder applicator (Fig. 2c) and cosmetic cream to moisturize lips during acquisition process [18], [21], [23]. Acquisition by use of the mentioned tools is described in detail in [14] and [22].

(a) (b) (c)

Fig. 2. Tools used during acqusition process.

An example of a lip print image obtained by means of the mentioned acquisition method is shown in Fig. 3a. On the images prepared in this way, a forensic expert marked a linear pattern consisting of segments representing the real lines occurring on rubor labiorum. Fig. 3b presents an example of a lip print image with segments marked on it by an expert. The image containing only segments marked by the expert was saved as a bitmap and provided as an input data for the research method. Fig. 3c presents an example of an input image.

Next, the lip print images were subjected to a feature extraction. Extracted features were segments located in the lip images. In order to find these segments, the Hough transform has been used to find straight lines and then an algorithm for searching for segments lying on straight lines has been applied.



(a) (b) (c)

Fig. 3. (a) Lip print image. (b) A lip print image with segments marked by an expert. (c) Image of segments characteristic of lip prints.

Finally, each test lip image was described by a set of segments represented by the previously provided attributes (Section 2.2).

4.2. Results

A number of 45 lip print images were used in the experiments. For the validation of the proposed Algorithm 1 we applied the Leave-one-out cross-validation (LOOCV) with respect to lip prints. It means that every selected image u_x from the repository of 45 lip prints has been assumed as unknown, i.e., assigned to the unknown person. Then the image u_x has been tested using Algorithm 1. If the image u_s returned by the Algorithm 1 has been assigned to the same person as the tested image u_x, we assumed that the correct classification occurred.

All considered images were obtained from 15 persons, 3 lip prints from each of them. Each image has been adequately prepared using the previously described methods and subjected Algorithm 1. The number of clusters was varied in the range from 10 to 100 every 10. Efficiency of classification was calculated for each measurement. The results of the experiment are presented in Fig. 4.

When analysing the results it can be seen that the best results were obtained for a number of cluster equal 80 or more. In this case the efficiency of classification was 82.2%. In Figure 4, the effectiveness of classification depending on the number of clusters is presented.

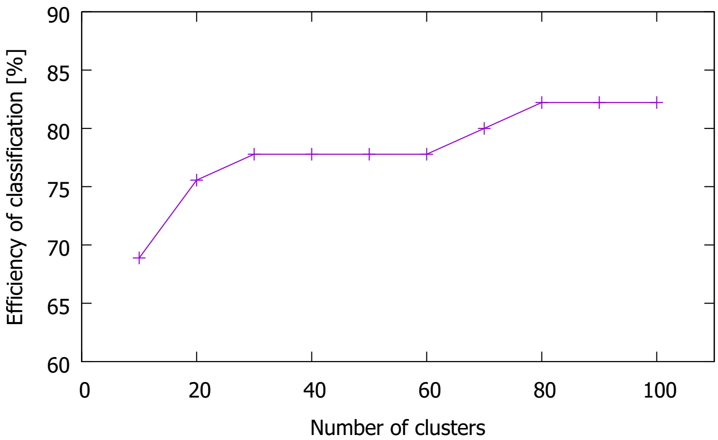


Fig. 4. The effectiveness of classification depending on the number of clusters.

As can be noted in Fig. 4 the accuracy of forecasting grows while increasing the number of clusters. First it raises rapidly in the range [10,30] and then stabilizes. Another increase of the accuracy can be noted in range [60,80]. Increasing the number of clusters up to 80 stabilized the accuracy. On the basis of the obtained results a conclusion can be drawn that the optimal number of cluster is about 80.

5. Final remarks

In this paper we addressed the problem of lip prints recognition. We proposed a novel approach based on Fuzzy c-Means that enables to increase the accuracy of forecasting. The approach presented in this paper, proved to be useful in the recognition of lip prints. The results are satisfactory and suggest the possibility of using the method in forensic identification systems. The results of the studies encourage to carry out further experiments, so the described method can also be the basis for further studies.

Bibliography

[1] Agarwal G., Ratha N., Bolle R. Biometric verification: Looking beyond raw similarity scores. Workshop on Multibiometrics (CVPR), 2006. pp. 31–36.

[2] Ballard D. Generalizing the hough transform to detect arbitrary shapes. Pattern Recognition, 1981, Vol. 13. pp. 111–122.

[3] Bezdek J. C. Pattern recognition with fuzzy objective function algoritms. 1981. Plenum Press.

[4] Bhatnagar J., Kumar A. On some performance indices for biometric identification system. Lecture Notes in Computer Science, 2007, Vol. 4642. Springer, pp. 1043–1056.

[5] Castleman K. Digital image processing. 1996. Printice–Hall, New Jersey.

[6] Choras M. The lip as a biometric. Pattern Analysis And Applications, 2010, Vol. 13. Springer, pp. 105–112.

[7] Duda R., Hart P. Use of the hough transformation to detect lines and curves in pictures. Comm. ACM, 1972, Vol. 15.

[8] Dunn J. C. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. Journal of Cybernetics, 1973, Vol. 3. pp. 32–57.

[9] Gonzalez R., Woods R. Digital image processing. 1992. Addison-Wesley.

[10] Hough P. Machine analysis of bubble chamber pictures. Proc. Int. Conf. on High Energy Accelerators and Instrumentation, 1959.

[11] Hough P. Method and means for recognizing complex pattern. U.S. Patent No. 3069654, 1962.

[12] Illingworth J., Kittler J. A survey of the hough transform. Computer Vision, Graphics and Image Processing, 1988, Vol. 44(1). pp. 87–116.

[13] Kasprzak J. Possibilities of cheiloscopy. Forensic Science International, 1990, Vol. 46. pp. 145–151.

[14] Kasprzak J., Leczynska B. Cheiloscopy. human identification on the basis of lip trace (in polish). 2001.

[15] Newton M. The encyclopedia of crime scene investigation. 2008.

[16] Parker J. Algorithms for image processing and computer vision. 1987. John Wiley & Sons.

[17] Petherick W., Turvey B., Ferguson C. Forensic criminology. 2010. Elsevier Academic Press.

[18] Porwik P., Orczyk T. Dtw and voting–based lip print recognition system. Lecture Notes in Computer Science, 2012, Vol. 7564. Springer, pp. 191–202.

[19] Suzuki K., Tsuchihashi Y. Personal identification by means of lip prints. Journal of Forensic Medicine, 1970, Vol. 17. pp. 52–57.

[20] Tsuchihashi Y. Studies on personal identification by means of lip prints. Forensic Science, 1974. pp. 127–231.

[21] Wrobel K., Doroz R. Method for identification of fragments of lip prints images on the basis of the generalized hough transform. Journal of Medical Informatics & Technologies, 2013, Vol. 22. pp. 89–93.

[22] Wrobel K., Doroz R., Palys M. A method of lip print recognition based on sections comparison. IEEE Int. Conference on Biometrics and Kansei Engineering (ICBAKE), 2013. pp. 47–52.

[23] Wrobel K., Doroz R., Palys M. Lip print recognition method using bifurcations analysis. Lecture Notes in Computer Science, 2015, Vol. 9012. Springer, pp. 72–81.