NEW DIRECTIONS IN CONTACT FREE HAND RECOGNITION

Xiaoqian Jiang, Wanhong Xu, Latanya Sweeney, Yiheng Li, Ralph Gross, Daniel Yurovsky

The Laboratory for International Data Privacy
School of Computer Science
Carnegie Mellon University
Email: xiaoqian@andrew.cmu.edu

ABSTRACT

The ability to quickly compute hand geometry measurements from a freely posed hand offers advantages to biometric identification systems. While hand geometry systems are not new, typical measurements of lengths and widths of fingers and palms require rigid placement of the hand against pegs. Slight deviations in hand position, finger stretch or pressure can yield different measurements. This paper offers novel approaches to computing hand geometry measurements from frontal views of freely posed hands. These approaches offer advantages in hygiene, comfort and reliability. Our algorithms segment the hand from a known background under spot lights and locate feature points along the fingers and wrists. Given a database of 54 hand images, with three different images of the same hand of each subject, our approach uniquely identified a previously unseen hand with an overall accuracy of 92%.

Index Terms— Hand geometry, Image processing, Machine learning

1. INTRODUCTION

Because of the increased hygiene concern in biometric systems and the difficulty in recognizing fingerprints of manual laborers and elderly people, hand geometry has been currently employed in many systems for personal verification mostly as a complement to finger-print authentication.

Traditional hand geometry-based systems use low-resolution cameras or scanners to capture users' hand images with the help of pegs or by forcing them to touch a screen. Those systems measure a hand shape to extract its features, like lengths and widths of fingers, and hand contour, for recognition. Unfortunately, traditional techniques face unsolved problems: low discriminability due to low-resolution hand images and bad user acceptability because users worry about hygienic issues when they have to touch screens.

To improve discriminability and user acceptability, our new hand recognition systems must acquire high-resolution hand images without peg constraints and contacts. However, those settings rise new challenges, like hand tilting, motion and shadow, to extract hand shapes, and measure hand features.

This work was funded by the United States Department of Justice under award number 2005-IJ-CX-K046.

2. RELATED WORKS

Traditional hand-based biometric schemes measuring geometric features of fingers and palms requires rigid placement of the hand against pegs or contact with a touch screen. For example, Oden et al used geometric features and implicit polynomial invariants of fingers and his classifiers are based on Mahanlanobis distance [1]. Sanchez-Reillo classified 25 geometric features based on Gaussian Mixture Models [2]. Bulatov developed a classifier based on Chebyshev metric between feature vectors [3]. Kumar designed a correlation-like similarity measurement to distinguish different hands [4]. Erdem suggested two alternative methods: one classifier based on modified Hausdorff distance of hand contour and another classifier based on the Euclidean distance on features consist of independent components of the hand silhouette [6].

Our work stands out traditional methods as we offer novel approaches including a robust feature point extraction algorithm and a machine learning based method to classifier using hand geometry measurements from frontal views of freely posed hands in contact free settings. Another novel approach alone the same lines is the use of AAMs for both hand segmentation and localization of key feature points by Gross [8] recently. Of note that this AAMs approach based paper uses the same features and database as we used in our paper.

3. METHODS

3.1. Overview of our system

The pipeline of our system consists of four steps: 1)Image acquisition by a static video camera; 2)Hand contour segmentation based on a decision tree; 3)Measurement of local feature points extracted along the fingers and wrists; 4)Identification based on geometry measurements of a query image against a database of recorded measurements using Support Vector Machine (SVM). We propose three innovative methods

3.2. Hand Segmentation

The segmentation procedure aims to separate the skin region of the hand from the rest of an image. At a glance, this seems a trivial task to segment hand foreground from the dark background. However, in the real world data, segmentation might suffer from artifacts like rings, wristwatch, belts chains or creases around the borders. In addition, image captured under spot light usually comes with shadows. This sometimes leads to ambiguous boundary when naive classifiers are used. Furthermore, the delineation of the hand contour must be accurate since the difference between hands of different individuals can be minute. Small biases in hand contour can lead to significant errors in feature extraction and recognition.

We first model the image segmentation problem as a mixture of three Gaussian distributions: skin, background and shadow

$$p(x) = \sum_{k=1}^{3} \pi_k N(x|\mu_k, \delta_k) \tag{1}$$

A K-means algorithm is adopted to initialize the mixture coefficient π_k and distribution parameters means μ_k, covariance δ_k, $k \in \{1, 2, 3\}$. These coefficients and parameters are tuned by EM (Expectation Maximization) algorithm. Finally, we use MAP (Maximum A Posteriori) method, the converged class affiliation posterior, to conduct the segmentation. However, the maximization of the posterior does not take the region information into consideration and assign disconnected pixels-regions to the same class. The result can be further biased when the algorithm gets stucked in local maximum. This problem can be addressed if we proper initialize the centroid of each cluster. Nevertheless, the inference steps of Expectation Maximization are too expensive for a fast hand recognition system.

Then, motivated by using skin color in face detection [5], we consider performing hand segmentation by using supervised machine learning technique to hand skin color. The intuition is that the (R, G, B) color of pixels of the hand region follow some pattern. In practice, we apply decision tree on hand labeled images to learn classification rules. We obtain the following eight rules after merging each decision path.

```
((95.5<=r<119.5 && 45.5<=b<101.5) |
(119.5<=r<157.5 && 45.5<=b<118.5) |
(130.5<=r<157.5 && 118.5<=b<142.5) |
(157.5<=r<176.5 && b<160.5 && g>=99) |
(176.5<=r<191.5 && b<176) |
(176.5<=r<199.5 && b<218.5 && g<196.5) |
(b<218.5 && r>=199.5) |
(b>=218.5 && r>=227.5))
```

Using these nine rules, a boundary walking algorithm retrieves the hand boundary. Sample results are illustrated in Figure 1.

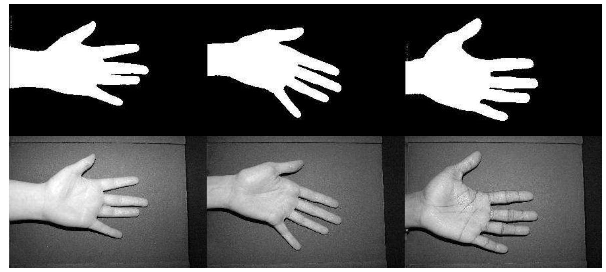


Fig. 1. Color based segmentation

3.3. Localization of Hand Extremities

Localizing hand extremities like fingertips and valleys between fingers is the first and most important step in feature extraction. The precision and robustness requirements on this procedure are high otherwise recognition will suffer from measurement errors. Based on the observation that extremities occur at high curvature locations, we conduct experiments on contour gradient gram, that is, the plot of contour gradient at various scales. Reported in [6], this naive method suffers from artifacts and ends up with unsmoothed contour. Interestingly, we observed from the previous experiment when the fingers are posed perfectly vertical, gradients of contour points symmetrically distributed near a fingertip cancel out and sum close to zero. We calculate a second version gradient gram by summing over gradients of 4 neighboring points, symmetrically distributed around any point on contour. Nine local minima clusters are collected and their centroids are sought as initial hand extremities. However, the vertical posing assumption does not always hold, (for example the thumb is almost always tiled in our captured images.)

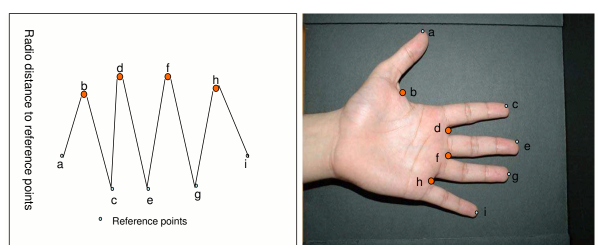


Fig. 2. Radio distance to reference points

An alternative method is known as radial distance technique [6], which yields hand extremum points correspond to local maxima and minima with respect to a reference point around the wrist region on the plot of radial distance. Given good reference point, this method is stable since the fingertips and valleys between fingers are not affected by relative small contour irregularities. The effectiveness of this method highly relies on the localization of the reference point.

Inspired by both gradient gram based algorithm and radial distance method, we proposed a new method iteratively refines hand extremities. This method makes no assumption on the fingers' pose as in the gradient gram algorithm and it outperforms the radial distance method due to the dynamic reference point localization scheme in Algorithm 1.

Algorithm 1 Fingertips and valley localization

Input: p (a set of hand contour points in counter-clockwise order)

Output: Location of fingertips and valleys between fingers

1: Plot averaged gradient gram of the contour $c_j = \sum(arctan((y_i - y_{i-s})/(x_i - x_{i-s})))$, s is the step size, centroid of clustered local minima are sought as initial fingertip locations;

2: Start from thumb's fingertip, find the radial distance maxima using two nearby fingertips as reference points. Label this maximum as valley between fingers.

3: Refine fingertips of index finger, middle finger, ring finger using their neighbor valleys found in the step 2. For thumb and little finger fingertips, use only neighbor valleys.

4: Iterate this process to the next finger until little finger fingertip is refined.

Algorithm 2 Find Wrist Extremities

Input: p (A segment of the contour containing the wrist point of interest $p_1, p_2 ... p_n$)

Output: Location of the wrist point (i, j)

Initialize stepsize $= s$, sample the contour $P = p_1, p_2 ... p_m$ into discrete points in counter-clockwise order.

Let M_P be a $size = |M|$ array, initialize each element M_i to 0;

for Each $i \in P$ **do**

  if $((p_i - p_j)(p_j - p_k)) > 0$ **then**

    $a = |p_i - p_j|, b = |p_j - p_k|, c = |p_i - p_k|$;

    $si = 1/2 * (a + b + c)$;

    $M_j = \sqrt{s(s-a)(s-b)(s-c)}$;

  end if

end for

Rank and perform clustering on top K values in M_P, to compute the centroid (i, j);

Return (i, j);

To localize upper and lower wrist points, we develop a "walking triangle" algorithm to capture subtle curvature change on flat contour regions, refer to Algorithm 2. The intuition behind this is to amplify the contour curvature using a dynamic weight scale so that large curvature corresponds to large weights and small curvature corresponds to small weights. Using hand extremities obtained so far, we extract 13 out of following 15 measurements in figure 3 for recognition excluding measurements 1 and 11.

3.4. Identification

After feature extraction, we apply Support Vector Machine (SVM) for identification against a database of recorded measurements. Support Vector Machine (SVM) classification is one of the most actively developed methodology in data mining and machine learning.

For a linear binary classification task, SVM finds the separating hyper plane ($w \cdot x = 0$) that maximizes the *margin*, denoting the distance between the hyper plane and closest data points (support vectors). To maximize the margin denoted by $\frac{1}{||w||}$ while minimizing the error, the standard SVM solution is formulated into the following primal program [7]:

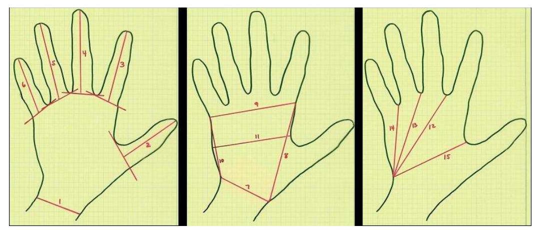


Fig. 3. 15 measurements

$$\min_{w,y} \quad \tfrac{1}{2}w'w + \nu e'y \tag{2}$$

$$s.t. \quad D(Aw - e\gamma) + y \geq e \quad and \;\; y \geq 0 \tag{3}$$

which minimizes the reciprocal of the margin ($w'w$) and the error ($e'y$). An $m \times n$ matrix A represents m data points in a n-dimensional input space. An $m \times m$ diagonal matrix D contains the corresponding labels (+1 or -1) of the data points in A. (A class label D_{ii}, or d_i for short, corresponds to the i-th data point x_i in A.) A column vector of ones of arbitrary dimension is denoted by e.The slack variable y is larger than zero when the point is on the wrong side or within the margin area. The soft margin parameter ν is tuned to balance the margin size and the error. The weight vector w and the bias γ will be computed by this optimization problem. The class of a new data x will be determined by $f(x) = w'x - \gamma$, where the class is *positive* if $f(x) > 0$, or else *negative*.

Next, we extend this general binary SVM classifier to multi-class problem using *one-against-one* scheme. This approach takes every possible class pairs to build models. In the presence of N classes, $N(N-1)/2$ SVMs must be trained so that each SVM classifier separates a pair of classes. The class label of the query is determined by majority voting of these $N(N-1)/2$ SVMs.

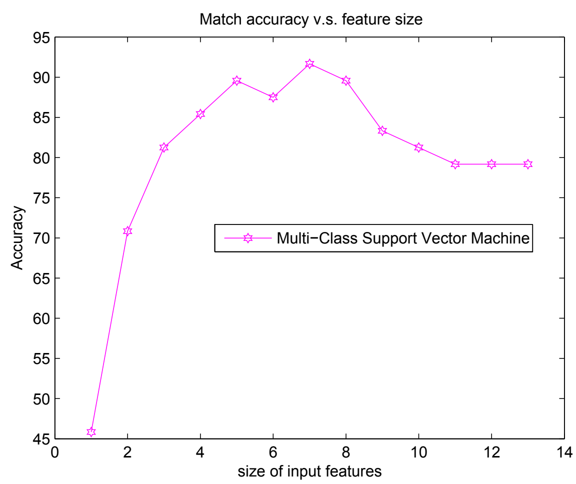
4. EXPERIMENTS

Our database consists of 54 frontal hand images. These are taken from 18 person and 3 image for each person using Professor Sweeney's "Hand Capture Device" [9]. The hand segmentation in Section 3.2 produce smooth contours like Figure 1. Feature extraction in section 3.3 collects 13 measurement features from each image.

4.1. Experiment Results

The identification accuracy is evaluated using the multi-class SVM in a 3-fold cross-validation manner using all 3 images for each object.

Fig. 4. SVM recognition performance with identified hand geometry features



We study the performance of the SVM classifier as we gradually increase input features. The input features are obtained using Relief-F algorithm, which evaluates every attribute by its ability to distinguish among instances that are near each other [10]. In specific, the features in this experiment are $[9, 7, 8, 6, 4, 3, 5, 12, 15, 14]$, ranked in descending order. We iteratively add new features and run the classifiers. Indicated by Figure 4, our best accuracy rate is 92% with input feature $[9, 7, 8, 6, 4, 3, 5]$.

Next, we compare our result to that in the AAMs paper [8], which uses the same features, image database and 3-fold cross-validation as ours. Our overall accuracy 92% is better than 90.7%, the reported accuracy of the AAMs fitted model and close to their manually established groundtruth 94.4%. In addition, our approach requires less setup and training time than AAMs model, which need intensive computation to maximize energy functions with manually labeled initials.

5. CONCLUSION

This paper introduces a novel solution for contact free hand recognition. We try to address the new challenges such as free motion, shadow and shape deformation. In this writing, we propose a robust color segmentation algorithm, a pose insensitive hand extremities localization algorithm, and a SVM based identification approach. The overall accuracy 92% seems promising and it encourages our further effort along the same lines. Upon obtaining more samples for each subject hand, we expect our algorithm to have better performance.

6. REFERENCES

[1] C. Oden, A. Ercil and B. Buke, "Combining implicit polynomials and geometric features for hand recognition," *Pattern Recognit. Lett.*,Vol.24, pp.2145-2152, 2003

[2] R. Sanchez-Reillo, C. Sanchez-Avila and A. Gonzalez-Marcos, "Biometric identification through hand geometry measurement," *IEEE Trans. Pattern Anal. Mach. Intell.*,Vol.22, pp.1168-1171, 2000

[3] Y. Bulatov, S. Jambawalikar, P. Kumar and S. Sethia, "Hand recognition using geometric classifiers," *ICBA Int. Conf. Bioinformatic and its Applications.*, 2004

[4] Y. A. Kumar, D. C. M. Wong, H. C. Shen and A. K. Jain, "Personal verification using palmprint and hand geometry biometric," *Proc. 4th Int. Conf. Audio Video-Based Biometric Person Authentication.*, pp.668-678, 2003

[5] Jure Kovac, Peter Peer and Franc Solina, "Human Skin Color Clustering for Face Detection," *International Conference on Computer as a Tool.*, 2003

[6] Yoruk, E. Konukoglu, E. Sankur and B. Darbon, J., "Shape-based hand recognition," *IEEE transactions on image processing*, Vol.15, pp.1803-1815, 2006

[7] G. Fung and O. L. Mangasarian, "Proximal support vector machine classifiers," *n Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, pp. 77-86, 2001

[8] R. Gross, Y. Li, L. Sweeney, X. Jiang, W. Xu and D. Yurovsky, "Robust Hand Geometry Measurements for Person Identification using Active Appearance Models," *Carnegie Mellon University, School of Computer Science Technical Report CMU-ISRI-06-123*, 2006

[9] L. Sweeney, "Hand Shake: a Database of Naturally Posed Hands," *Carnegie Mellon University. Data Privacy Lab Working paper 31*, 2006

[10] YX Zhang, AL Luo and YH Zhao, "An automated classification algorithm for multiwavelength data", *Proceedings of SPIE*, 2004