Histogram-based matching of GMM encoded features for online signature verification

Vivek Venugopal
On behalf of
Abhishek Sharma,Dr. Suresh Sundaram

Multimedia Analytics Laboratory, Electronics and Electrical Engineering Department
IIT Guwahati

August 8, 2018

Outline

- Introduction
- Problem Formulation
- Proposed System
- Results and Discussion

Introduction

- Signature verification system- Contrast given signature with enrolled genuine signatures of a user for authentication [1].
- Two outcomes:- Genuine , Forgery.
- Online and Offline (Static).
- Distance based [2-3] and Model based [4-5]

---

[1] A. K. Jain, F. D. Griess, and S. D. Connell, "On-line signature verification," *Pattern Recognition,* vol. 35, no. 12, pp. 2963-2972, Dec. 2002.
[2] A. Kholmatov and B. Yanikoglu, "Identity authentication using improved online signature verification method," Pattern Recognition Letters, vol. 26, no. 15, pp. 2400-2408, 2005.
[3] K. Barkoula, G. Economou, and S. Fotopoulos, "Online signature verification based on signatures turning angle representation using longest common subsequence matching," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 16, no. 3, pp. 261-272, 2013.
[4] J. Fierrez, J. Ortega-Garcia, D. Ramos, and J. Gonzalez-Rodriguez, "HMM-based on-line signature verification: Feature extraction and signature modeling, *Pattern Recognition Letters*, vol. 28, no. 16, pp. 2325-2334, 2007.
[5] E. Argones Rua and J. Alba Castro, "Online signature verification based on Generative models, *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 42, no. 4, pp. 1231-1242, Aug 2012.

Problem Statement

- A number of systems on online signature verification perform a temporal alignment between the feature vectors that are derived at each sample point of the online trace of the signatures being compared.
- Consideration of feature vector sequence in probabilistic frame work can help in capturing the user dependent characteristic of signature in better way.
- In this work, we use the parameters from a pre-learnt Gaussian Mixture Model (GMM) to encode the features.
- The histogram derived from GMM encoded feature is used for matching test signature with enrolled signatures.

Proposed System

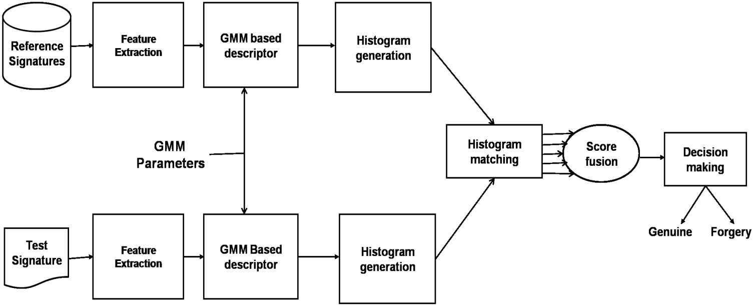

Figure: Block diagram of proposed verification scheme.

Feature Extraction

Basic attributes normalized using min-max normalization.

- **First order difference of basic features**:$\Delta x(i), \Delta y(i), \Delta p(i), \Delta \phi(i), \Delta \theta(i)$.
- **Second order difference of spatial coordinates** : $\Delta^2 x(i), \Delta^2 y(i)$.
- **Sine and cosine measures** : $\sin(\alpha(i)), \cos(\alpha(i))$.
- **Length-based features** : $l(i), \Delta l(i)$

$$\begin{aligned} l(i) &= \sqrt{(\Delta x(i))^2 + (\Delta y(i))^2} \\ \Delta l(i) &= \sqrt{(\Delta^2 x(i))^2 + (\Delta^2 y(i))^2} \end{aligned} \tag{1}$$

N genuine (reference) signatures $\{S_1, S_2, \cdots, S_p, \cdots, S_N\}$

$$\mathbf{F}_p = \{\mathbf{f}_p^1, \mathbf{f}_p^2, ..., \mathbf{f}_p^{n_p - 2}\} \qquad 1 \leq p \leq N \tag{2}$$

$$\mathbf{f}_p^j = [f_p^j(1) \;\; f_p^j(2) \;\; .... \;\; f_p^j(11)]^{\mathsf{T}} \tag{3}$$

GMM Based verification System

Log likelihood function:

$$L = \frac{1}{n_T - 2}\Big(\sum_{j=1}^{n_T-2} \ln(\sum_{i=1}^{M} w_i \mathcal{N}(\mathbf{f}_T^j \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i))\Big)$$

- Explicit contribution of each component ignored.
- Encode feature vector probabilistically with parameters learnt from GMM.
- Temporal information not adequately captured.
- Histogram generation over signature segments.

GMM based descriptor

- Each user is modelled by a specific GMM of M Gaussian components, with parameters $\{w_k, \boldsymbol{\Sigma}_k, \boldsymbol{\mu}_k\}_{k=1}^{M}$.

- Each feature vector $\mathbf{f}_p^j$ from the trace of the test signature S_p is encoded using GMM descriptor as follows:

$$g_p^j(k) = \frac{w_k \mathcal{N}(\mathbf{f}_p^j \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{c=1}^{M} w_c \mathcal{N}(\mathbf{f}_p^j \mid \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c)} \tag{4}$$

$$\mathbf{g}_p^j = [g_p^j(1) \quad g_p^j(2) \quad .... \quad g_p^j(M)]^{\mathsf{T}}$$

Histogram Generation

- Set number of bins in histogram equal to M and initialise with zero votes.
- Corresponding to each j^{th} sample point of the signature S_i from a user, the indices in histogram are voted in accordance to the elements in $\mathbf{g}_i^j$.
- Repeat accumulation across all sample points of the online trace and then normalize-Base Histogram
- To incorporate local information- first partition signature into q segments.
- Histogram comprising $q \times M$ bins is initialized with zero and voted with corresponding sample points to obtain desired histogram and then normalized.

Histogram Matching

- Histogram of test signature $\mathcal{H}_T$ is matched to $\{\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}_N\}$.
- Chi-Squared distance

$$d_i = \sum_{j=1}^{B_q} \frac{(h_T(j) - h_i(j))^2}{h_T(j) + h_i(j)} \qquad 1 \leq i \leq N$$

- $B_q = M * q$ - number of bins in histogram generated after dividing signature into q segments.
- Mean of d_is is then used for verification.

Online Signature Database

Online Signature Database: MCYT-100.

Database Name	Total Participants	Genuine Sign	Skilled Forgery	Total Signatures
MCYT-100	100	25	25	5000

- Basic attributes:x, y, pr, γ, ϕ.
- Performance measure - Equal Error Rate (EER).
- Ten repetitions.
- 3 systems implemented
  - *GMM-LIKE*:
  - *GMM-HIST1*:
  - *GMM-HIST2*:

S. G. Salicetti, N. Houmani, B. L. Van, B. Dorizzi, F. A. Fernandez, J. Fierrez, J. O. Garcia, C. Vielhauer and T. Scheidat, "Online Handwritten Signature Verification," *Guide to Biometric Reference Systems and Performance Evaluation, Chapter 6*, Nov. 2008

Experimental results

- Performance evaluation of the proposal - *GMM-HIST1* and *GMM-HIST2* systems for different number of Gaussian components M in the GMM.

# of Gaussian components M	Common Threshold		
	GMM-LIKE	*GMM-HIST1*	*GMM-HIST2*
2	20.42	14.96	13.90
4	18.69	11.65	9.16
8	16.94	8.96	6.63
16	14.94	6.77	5.11
32	12.82	5.53	4.48
64	11.61	4.97	**3.72**
128	12.94	5.62	4.49

Experimental results

- EER (%) values with different verification strategy and $M = 64$.

Scheme	Common Threshold	
	GMM-HIST1	*GMM-HIST2*
Mean	4.97	3.72
Minimum	5.26	4.39
Maximum	7.72	5.81

Comparison with prior works

Table: Survey of prior works on the MCYT database.

Method	MEER
Histogram Based Analysis [1]	4.02
Two stage normalization+DTW [2]	3.94
UBM-HMM + fuzzy cryptography [3]	5.87
User dependent features + classifiers [4]	19.4
Proposed method	**3.72**

[1] N. Sae-Bae and N. Memon, "Online signature verification on mobile devices, *Information Forensics and Security, IEEE Transactions on*, vol. 9, no. 6, pp. 933-947, June 2014.
[2] A. Fischer, M. Diaz, R. Plamondon, and M. A. Ferrer, "Robust score normalization for dtw-based on-line signature verification, in *Document Analysis and Recognition (ICDAR)*, 2015 13th International Conference on. IEEE, 2015, pp. 241-245.
[3] E. A. Rua, E. Maiorana, J. L. A. Castro, and P. Campisi, "Biometric template protection using universal background models: An application to online signature, *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 269-282, 2012.
[4] K. Manjunatha, S. Manjunath, D. Guru, and M. Somashekara, "Online signature verification based on writer dependent features and classifiers, *Pattern Recognition Letters*, vol. 80, pp. 129-136, 2016.

Thank You