구조적 서명 검증에서의 참조 서명의 데이터 크기 고정화 기법

이이섭* · 김성훈**

A Technique for Fixing Size of Reference Signature Data in Structural Signature
Verificaiton

Lee-Sub Lee* · Seong-Hoon Kim**

이 논문은 2008년도 경북대학교 신임교수정착연구비에 의하여 연구되었음

요 약

구조적 접근방법에 의한 서명검증은 서명을 부분획 단위의 집합 구조로 표현하여 서명 데이터를 분석하기 때문에 서명의 국부적인 변화의 특성을 효율적으로 흡수하여 우수한 검증 효과를 나타낸다. 그러나 이 방법은 실용화할 때 데이터의 크기가 고정되지 않는 문제가 발생한다. 이 논문에서는 이 문제를 해결하기 위하여, 상대적으로 중요한 부분을 선택하여 데이터를 고정크기로 만듦으로서 데이터의 크기를 줄이는 방법을 제안한다. 실험 결과, 고정된 크기로 선택된 서브패턴에 대해서 국부적인 변화도와 복잡도에 의한 가중치를 적용한 결과, 더 좋은 검증율을 보였으며, 이 때 고정된 크기를 만들기 위한 최소 크기의 참조 모델의 서브패턴의 개수는 일정 갯수의 범위에서 적절히 고정될 수 있음을 보였다.

ABSTRACT

The structural approach in the signature verification, representing a signature as a structural form of local primitives, shows an excellent performance since it counts in the local characteristics such as local variation, stroke complexity, and etc. However, this method has a problem of template data sizing which can not fix the number of subpatterns comprising a signature. In this paper, we proposed a new algorithm to reduce the signature data into a fixed size by selecting a fixed number of subpatterns which is considered as important parts.

As a result, it shows more excellent performance when the fixed sized sub-patterns is applied with local weights extracted from variational characteristics and complexities in local part. And the number of subpatterns representing a signature reference model can be fixed under a certain number of segments determined appropriately.

키워드

고정크기 참조서명, 필기 분할, 온라인 서명검증

Key words

Fixed-sized Reference Signature, Handwriting Segmentation, On-line Signature Verification

* 금오공과대학교 컴퓨터공학부

** 경북대학교 컴퓨터정보학부 (교신저자, shkim1454@knu.ac.kr)

접수일자 : 2010. 04. 13

심사완료일자 : 2010. 04. 29

I. 서 론

정보화와 네트워크 시대가 도래함에 따라 수많은 정보와 개인정보가 디지털화되어 네트워크 환경에 노출되어 있다. 이에 따라 각종 정보기기 및 데이터베이스 접근에 대한 보안의 필요성이 증대한 문제점으로 대두되고 있다. 비밀번호나 ID카드, 열쇠와 같은 기존의 보안 방법은 유출, 분실, 훼손, 도난 등의 문제 발생 소지가 있고, 이를 보완하고자 지문, 홍채, 얼굴, 서명, 등의 생체인식 시스템이 도입되고 있다. 생체인식은 안전성과 편리함을 모두 갖춘 미래의 사용자 인증기술로써, 미국의 911 테러 사태이후 세계적으로 관심이 더욱 집중되고 있다.

이중, 서명은 다른 생체인식 방법에 비해 사람들에게 친숙하며 상대적으로 저렴한 장비를 사용하여 단시간에 본인 여부의 확인이 가능하기 때문에 큰 인기를 얻고 있다. 요즘은 태블릿PC와 스마트폰의 폭발적인 보급으로 펜과 터치 인터페이스가 모바일 단말장치에 일반적인 인터페이스로 보급됨에 따라 생체인식의 방법으로 서명인식이 새롭게 주목받고 있다. 또한, 국내 전자여권이 도입되어 생체 정보를 본인확인 방법으로 활용될 수 있는 여건이 마련되었다. 하지만, 지문 정보 획득의 거부감으로 인해 얼굴 사진 만에 의한 본인

확인의 한계를 아직도 내포하고 있어서, 이를 보완할 추가적인 생체정보로 서명이 하나의 대안으로 생각될 수 있다[1][2].

금융카드나 교통카드, 그리고 전자여권 등에 내장된 스마트카드 칩은 제한된 컴퓨팅 환경을 갖는다. 특히, 기록 가능한 기억장치인 EEPROM은 기껏해야 수십 KB (64KB~80KB)에 지나지 않는다. 이러한 환경에서는 생체 데이터는 고정된 크기의 용량으로 제약될 수밖에 없다.

서명검증의 문제는 특정인의 ID로 입력된 서명을 분석하여 서명의 진위 여부를 판별하는 것으로 정의되며, 서명 검증 시스템이란 이러한 작업을 수행하는 시스템을 말한다. 서명 검증 시스템은 서명의 입력 방식에 따라 온라인과 오프라인으로 구분된다.[3] 오프라인 검증은 지면에 남겨진 서명을 스캐너나 카메라로 입력 받아서 영상 데이터를 획득하여 이를 검증하는 방법이며, 온라인 검증은 서명자가 서명하는 순간에 태블릿과 전자펜을 이용하여 데이터를 직접 입력받아 사용하는 방법이다. 그리고, 온라인 서명 검증의 방법으로는 함수적 접근 방법과 매개변수적 접근방법, 그리고 구조적 접근방법으로 나뉘어진다. 이중 매개변수적 접근방법은 고정된 크기의 특정 데이터를 참조서명의 데이터로 저장하지만, 일반적으로 검증률이 좋지 못하다. 이에 반해, 함수적 접근방법은 검증률이 가장 좋은 것으로 평가되지만,

그림 1. 일반적인 서명 검증 과정
Fig. 1. General Process of On-line Signature Verification

시간에 대한 함수로 입력되는 서명데이터를 모두 저장해야 하므로, 서명에 따라서 저장해야될 데이터를 고정된 크기가 될 수 없다.

본 논문에서는 구조적 접근 방법을 도입하여 저장 데이터의 크기가 사람마다 유동적인 문제점을 해결하는 즉, 고정된 크기의 서명 데이터를 얻고자 하는 방법을 모색하고자 한다. 제안하는 방법은 서명을 구조적으로 분석하여 중요한 부분과 그렇지 않은 부분을 구분하여, 고정된 데이터 크기의 참조모델을 구성하는 것이다.

II. 구조적인 서명 검증 기법의 개요

서명 검증을 위한 일반적인 처리 과정은 그림 1과 같이 등록 절차와 검증 절차로 나누어진다[3]. 등록 절차는 서명자의 고유 번호와 함께 서명 샘플을 입력받아 전처리 과정과 특징 추출 과정을 거쳐 데이터베이스에 저장하는 단계이고, 검증 절차는 입력 서명에 대해 전처리 과정과 특징 추출과정을 거쳐 이미 데이터베이스에 저장된 참조 서명의 특징 값과 비교하여 진위를 판단하는 단계이다. 두 단계에서 필요로 하는 세부 과정은 전처리 과정, 특징 추출 과정, 데이터베이스 등록과정, 비교과정, 진위 판단과정 등이 있다.

전처리 과정은 특징 추출 방법에 따라 필요한 과정이 달라지는데, 일반적으로 태블릿 표면의 미끄러움이나 손의 떨림 등을 제거시키는 서명 획의 평활화(smoothing), 서명의 회전(rotation), 크기(scale) 보정을 행하게 된다.

구조적 서명 검증 기법은 서명의 분할(segmentation)로부터 얻어지는 프리미티브 패턴을 서명의 최소 기본단위로 하여 서명 전체를 구조적인 계층 관계로 구성하고 이를 바탕으로 서명 검증 과정을 행하는 방법이다[4].

2.1 프리미티브 패턴의 정의

본 논문에서는 서명의 분할로부터 얻어지는 서명의 최소단위인 프리미티브는 최소 속력점을 분할점으로 하여 얻어지는 부분획으로 정의하였다.

2.2 서브패턴의 형성 및 구분

서브패턴의 모양은 프리미티브 패턴의 연결 형태에 따라 결정된다. 본 논문에서는 서브패턴의 유형을 두 개의 프리미티브 패턴을 결합시켜 단순 회전(rotate), 첨형(cusp), 종형(bell)의 3가지 유형으로 정의하였다(그림 3).

그림 3. 서브패턴의 유형
(a) 단순회전 (b) 첨형 (c) 종형
Fig. 3. Types of Subpatterns
(a) Rotation (b) Cusp (c) Bell

그림 2. 서명의 분할
Fig. 2. Segmentation of Signature

III. 데이터 크기가 고정된 구조적 서명 검증

3.1 고정 크기의 서명 데이터 형성 과정

새로운 서명을 등록하는 과정은 서명자의 식별 번호와 함께 M 개의 진서명 샘플을 입력받아 진서명에 대한 참조 데이터를 만드는, 즉 고정크기를 형성하는 과정이다. 일반적으로 진서명에 대한 참조 데이터는 서명의 특징과 판단 임계치가 된다.

이 논문에서는 서명을 서브패턴 단위의 세그먼트로 구성하므로 서명에 대한 특징은 각 세그먼트에 대한 특징 벡터들로 구성된다. 샘플들로부터 추출된 특징 벡터들로 표현된 세그먼트들의 나열(sequence)을 예비 참조 모델이라 부른다. 예비 참조 모델에 포함되는 각 세그먼트들은 샘플들로부터 지정된 개수만큼 모양이 가장 안정되게 나타난 부분들만을 추출하여 이들로부터 특징 벡터를 얻도록 한다. 이때 안정되게 얻어진 부분의 특징 벡터를 참조 모델이라고 부른다. 각 세그먼트들에 대한 중요도를 반영하기 위해 국부적 가중치를 산출하였다. 이 과정에서 국부적인 변화도와 복잡도를 사용한다. 그러므로 서명에 대한 참조 모델은 안정된 세그먼트로 선택된 부분의 특징벡터와 국부적인 가중치가 된다.

M 개의 진서명 샘플들을 입력받아 참조 모델을 형성하는 과정은 그림 4에서와 같이 5단계의 과정을 거쳐서, 서명 샘플들에 대해 가장 안정된 세그먼트들의 특징 벡터를 구한다. 분할 및 특징 추출 과정은 입력된 N 개의 진서명 샘플을 분할하여 각 세그먼트별로 특징을 추출하는 과정이다. 동일 모양 서브패턴 매칭과정에서는 서명 샘플들간에 동일 모양을 갖는 세그먼트들 간의 매핑 관계를 얻기 위해 서브패턴 단위의 DP 정합을 행한다. 그리고 다음 과정인 고정 세그먼트 그룹핑 과정을 거쳐서 각 세그먼트들에 대해 동일 부분으로 매핑된 샘플 세그먼트들을 그룹핑한다. 각각의 세그먼트들에 대해 샘플 세그먼트들의 특징값의 평균과 표준 편차를 산출하고 변별력 있는 안정된 서브패턴을 추출하여 최종적으로 참조 모델 E 를 구하게 된다.

1. 분할 및 특징 추출

M 개의 서명 샘플들을 속력의 국부극소점에 의해 서명을 분할하고 서브패턴 단위로 특징을 추출한다.

그림 4. 참조 모델 형성 과정
Fig. 4. Generation of Reference Model

서명 샘플들의 집합을 S 라 하자.

$$S = \{S_1, S_2, S_3, \dots, S_m, \dots, S_M\} \quad (1)$$

서명 샘플 S_m 이 분할 과정을 거쳐 서브패턴 단위의 세그먼트들로 구성된 경우, 다음과 같이 표현한다.

$$S_m = [s_1^m, s_2^m, s_3^m, \dots, s_n^m, \dots, s_{N_m}^m] \quad (2)$$

여기서, N_m 은 서명 샘플 S_m 의 세그먼트의 개수

s_n^m 은 m 번째 샘플의 n 번째 세그먼트를 나타내는 특징 벡터이다.

$$s_n^m = [\text{특징 1}, \text{특징 2}, \dots, \text{특징 k}, \dots, \text{특징 K}] \quad (3) \\ = [s_{n,1}^m, s_{n,2}^m, \dots, s_{n,k}^m, \dots, s_{n,K}^m]$$

여기서, K 는 특징의 개수이다.

2. 동일 모양의 서브패턴 매칭

김성훈[4]이 제안한 서브패턴 단위의 DP 매칭을 이용하여 기준 서명을 선택한다.

$$\min_{S_i \in S} \left| \sum_{S_j \in S} D(S_i, S_j) \right| \quad (4)$$

여기서, $D(S_i, S_j)$ 는 서브패턴 단위의 DP정합의 차이가 계산값이다.

표 1. 세그먼트의 특징표
Table 1. Features of Segments

특징들	설 명
1	세그먼트의 폭
2	길이
3	첫번째 피크의 폭
4	두번째 피크의 폭
5	두 원시 패턴의 연결 각도
6	세그먼트의 필기 시간
7	최대 속력
8	평균 속력
9	최대속력 - 평균 속력
10	최대 압력
11	최소 압력
12	평균 압력
13	최대 압력 - 최소 압력
14	최대 압력 -평균 압력

3. 고정세그먼트의 그룹핑

선택된 기준 샘플을 각 세그먼트들에 대해 매칭 위평 함수로부터 다른 샘플들에서의 세그먼트를 그룹핑하는 과정(그림 5)이다.

그림 5. 고정 세그먼트 그룹핑 과정의 예
Fig 5. Example of Grouping Fixed Segments

4. 예비 참조 모델 형성

예비 참조 모델의 각 세그먼트들의 특징 벡터는 그룹핑된 고정 세그먼트들로부터 특징 공간상의 중심값과 표준편차를 계산하여 구한다.

$$r_n = \langle m_n, \sigma_n \rangle, \quad 1 \leq n \leq N_r \quad (5)$$

$$m_n = [m_1^n, m_2^n, \dots, m_k^n, \dots, m_K^n]$$

$$\sigma_n = [\sigma_1^n, \sigma_2^n, \dots, \sigma_k^n, \dots, \sigma_K^n]$$

여기서 n 번째 세그먼트에 대하여, r_n 은 특징 벡터의 평균과 표준편차의 쌍이며, m_n 은 특징벡터의 평균값이고, σ_n 은 특징벡터의 표준편차이다.

5. 참조 모델 형성

변별력 있게 선택된 서브패턴 E 은 그룹핑된 고정세그먼트 집합에서 n 번째 세그먼트의 예비 참조 모델과 학습 샘플들의 세그먼트간의 평균차이 값을 구하여 결정한다. 이 평균값을 참조 모델의 국부적인 변화도 V 라고 정의하였다. 국부적인 변화도 V 의 차이 값이 작을수록 안정하고 변별력 있는 서브패턴이라고 정의하였다. 고정 크기를 만들기 위해 차이 값이 작은 순서로 고정된 수의 서브패턴을 선택하여 저장하였다.

$$V = [v_1, v_2, \dots, v_n, \dots, v_{N_r}] \quad (6)$$

$$v_n = \frac{1}{M} \sum_{m=1}^M d(r_n, s_n^m) \quad (7)$$

$$d(r_n, s_n^m) = \sqrt{\sum_{k=1}^K \left(\frac{s_{n,k}^m - m_k^m}{\sigma_k^n} \right)^2} \quad (8)$$

여기서, 식 (8)은 n 번째 세그먼트와 m 번째 샘플의 n 번째 세그먼트와의 거리(Euclidean distance)이다.

$$E = [e_1, e_2, \dots, e_n, \dots, e_{N_r}] \quad (9)$$

여기서, e_n 은 예비 참조 모델 집합 R 에서 m 번째 변별력 있고 안정된 서명을 새롭게 n 번째로 정의한 것이다. 이때 서브패턴의 집합 E 를 참조 모델(그림 6)이라고 정의하였다.

그림 6. 선택된 참조 모델 집합의 예
Fig 6. Example of Selected Reference Models Set

3.2 국부적 가중치

세그먼트에 대한 가중치는 국부적인 변화도와 복잡도를 이용하여 산출한다.[4] 변화도는 모든 세그먼트들의 국부적인 변화도의 평균치로 산정한다. 복잡도는 Plamondon[5]의 정의에서 서브패턴 유형의 반복에 의한 감쇄효과를 추가하여 산정한다.[4]

변화도와 복잡도를 동시에 고려한 국부적 가중치 w_n 는 다음과 같이 n 번째 세그먼트의 변화도 w_n^v 와 복잡도 w_n^c 두 요소의 기하 평균으로 산정한다.

$$W = [w_1, w_2, \dots, w_n, \dots, w_{N_r}] \quad (10)$$

$$\text{where } w_n = \sqrt{w_n^v * w_n^c}$$

3.3 서명 검증

판단 임계치의 설정은 기존의 변화도에 의한 설정 방법, $\Theta = C \cdot \mu$ 을 이용하였다[6].

입력된 서명 T 와 참조 모델 E 와의 동일한 형태를 갖는 세그먼트들끼리의 매핑관계를 얻기 위해 세그먼트 단위의 DP 정합을 하여 전체 상이도는 다음식과 같이 정의된다.

$$Diss(E, T) = \frac{\sum_{n=1}^{N_e} d(r_n, t_n) \cdot w_n}{\sum_{n=1}^{N_e} w_n} \quad (10)$$

IV. 실험 결과

실험을 위한 데이터는 총 10명의 서명으로 구성되어 있다. 각각의 서명은 진서명 70, 모조 서명 각 50개로 구성되어 있다. 모조 서명은 10개의 진서명에 대해 각각 서로 다른 모조자를 정하여 1개월간의 모조 훈련을 거친 후, 1인당 50개의 샘플을 획득하였다.

(a)

(b)

그림 7. 구조적 방법의 검증률
(a) 가중치를 적용하지 않은 검증 오류율
(c = 1.1일 때, 최소 오류율 7.1%)
(b) 가중치 산출시의 검증 오류율
(c = 1.1일 때, 최소 오류율 6.6%)

Fig. 7 Performance of Structural Approach
(a) Error-Rate Curve of Applying No Weight-factor
(when c = 1.1, Minimum Error Rate is 7.1%)
(b) Error-Rate Curve of Applying Weight-factor
(when c = 1.1, Minimum Error Rate is 6.6%)

그림 8 고정 크기에 따른 Error Trade-off 곡선
Fig. 8 Error Trade-off Curves in Fixed-number Segments

실험은 진서명 샘플들에서 20개의 샘플들을 임의로 추출하여 등록과정에 사용하였고, 유형 I 오류 실험 샘플로 진서명의 등록과정에 포함되지 않은 50개의 샘플들을, 유형 II 오류 실험 샘플로 50개의 모조서명 샘플들을 사용하였다. 그리고 10사람의 서명을 분할하여 형성된 서브패턴의 평균 개수는 11~56개이다.

그림 7은 구조적 방법에 의한 서명 검증에서 서브패턴에 대한 국부적 가중치를 적용하였을 때의 성능의 향상을 보이기 위함이다. 그림 7-(a)에서 균등한 가중치로 모든 서브패턴을 사용하였을 때 평균 에러율 7.1%이고, 그림 7-(b)에서 국부적 가중치를 적용한 경우 평균 에러율 6.6%를 얻었다.

최적의 고정된 서명 데이터의 크기를 얻기 위한 실험은 서브패턴의 갯수를 줄여가면서 검증률을 구하는 실험을 행하였다. 그림 8의 실험 결과에서 보면, 참조 모델의 개수가 30개일 때 메모리와 검증률 측면에서 최적의 성능을 나타내었고, 이 때의 평균 에러율은 8.9%이며, 참조 모델의 저장용량은 약 1.8KB가 요구된다.

V. 결 론

본 논문에서는 서명 검증을 위해 중요한 부분을 선택하여 데이터의 크기를 고정된 크기로 사용하는 구조적 접근 방법을 제안하였다. 서명의 중요 부분들을 선택하기 위해서 변화도 정보와 복잡도 정보를 이용하여 참조 모델을 형성하였다. 이러한 국부적 변화도와 복잡도를 바탕으로 한 국부적 가중치 정보를 이용하여 서브패턴의 선별적 적용이 가능하므로, 고정된 개수의 서브패턴을 선택적으로 사용할 수 있게 되었다.

고정된 개수로 선택된 서브패턴에 대해서 국부적인 변화도와 복잡도에 의한 가중치를 적용한 결과, 가중치를 사용하지 않은 경우보다 더 좋은 검증율을 보였으며, 이 때 고정된 크기를 만들기 위한 최소 크기의 참조 모델의 서브패턴의 개수는 일정 개수의 범위 내에서 적절함을 알 수 있었다.

참고문헌

- [1] A. Leyenberger, "Using a PDA," PC Laptop Computers Magazine, pp.44-71, Feb. 1995.
- [2] S. Lee and J. C. Pan, "Offline Tracing and Representation of Signatures," IEEE Trans. on Systems Man and Cybernetics, Vol. 22, No. 4, pp.755-771, July/Aug. 1992.
- [3] R. Plamondon and G. Lorette, "Automatic Signature and Writer Identification - The State of Art," Pattern Recognition, Vol. 22, No. 2, pp. 107-131, 1989.
- [4] 김성훈, "구조적 접근방식의 온라인 자동 서명 검증 기법", 한국컴퓨터정보학회논문지, 제10권, 제4호, pp.385-396, 2005년 9월.
- [5] J. J. Brault and R. Plamondon, "A Complexity Measure of Handwritten Curves : Modeling of Dynamic Signature Forgery," IEEE Trans. on Systems Man and Cybernetics, Vol. 23, No. 2, pp.400-413, Mar./Apr. 1993.
- [6] J. Kim, J. R. Yu, S. H. Kim, "Learning of Prototypes and Decision Boundaries for a Verification Problem having only Positive Samples," Pattern Recognition Letters, Vol. 17, No. 1, pp.691-697, 1996.

김성훈(Seong-Hoon Kim)

한국해양정보통신학회 논문지
제14권 제5호 참조

저자소개

이이섭(Lee-Sub Lee)

- 2004년 9월 : 고려대학교 대학원
컴퓨터학과(이학박사)
 - 1990년~1992년 :
삼성종합기술원 주임연구원
 - 1993~2004년 : 삼성SDS 정보기술연구소
수석연구원
 - 2004년 9월~현재: 금오공과대학교 컴퓨터공학부
조교수
- ※ 관심분야: 소프트웨어공학, 모바일 및 임베디드
컴퓨팅, 지식관리