

매개변수적 서명 검증에서 개인화된 특징 집합의 가중치 유클리드 거리 산출 기법

김성훈*

A Technique of Calculating a Weighted Euclidean Distance with a Personalized Feature Set in Parametric Signature Verification

Seong Hoon Kim

Abstract

In parametric approach to a signature verification, it generally uses so many redundant features unsuitable for each individual signature that it causes harm, instead. This paper proposes a method of determining personalized weights of a feature set in signature verification with parametric approach by identifying the characteristics of each feature. For an individual signature, we define a degree of how difficult it is for any other person to forge the one's (called 'DFD' as the Degree of Forgery Difficulty). According to the statistical characteristics and the intuitional characteristics of each feature, the standard features are classified into four types. Four types of DFD functions are defined and applied into the distance calculation as a personalized weight factor. Using this method, the error rate of signature verification is reduced and the variation of the performance is less sensitive to the changes of decision threshold.

Key Words: weighted euclidean distance, parametric signature verification, personalized feature set, forgery difficulty degree

1. 서론

온라인 서명검증은 전자펜을 이용하여 입력된 좌표와 입력의 떨기 정보를 얻어 처리하고, 함수적 방법과 매개변수적 방법이 있다[1]. 함수적 방법은 입력된 서명 신호를 시간에 대한 특징 함수로 나타내고 대응되는 특징값의 차이를 누적시켜 비교하는 방법이다. 매개변수적 방법은 서명의 특징이 될 매개변수를 추출하고, 이것을 매개변수 공간상에 대응시켜 등록된 서명을 대표하는 중심값과 검증하려는 서명사이의 매개변수 공간상의 거리로서 참, 거짓을 판별하게 된다[1][2]. 매개변수적 방법은 각 서명에 대하여 저장해야 할 정보가 사용하는 특징들의 벡터이므로 저장할 정보량이 적은 장점이 있다. 서명 검증 시스템을 구현하는 실제적인 문제에 있어 입력받는 서명의 개수가 제한되므로 특징간의 상호관계를 정확히 구하기 어려우므로, 일반적으로 검증하려는 서명과 대표 벡터와의 거리를 구하는데, 각 특징들간의 상관관계가 없는 것으로 가정하는 유클리드 거리(Euclidean Distance)를 사용한다. 다만, 각 특징들간의 상호 반영하는 크기가 서로 다르므로, 각 특징축을 표준화하여 사용하게 된다[2].

이론적으로는 매개변수는 모든 참, 거짓을 구별하는 능력에 따라 선택되어야 한다. 그러나 현실적으로는 교묘하게 모조된 서명의 샘플을 사전에 충분히 확보할 수 없다는 근본적인 한계로 인해, 부분적으로 얻어진 소수의 모조 데이터로부터 구분력을 판단하여 특징을 추출하게 된다. 또한, 각 서명의 모양과 내부 변화에 따라 그 서명에 적합한 특징들을 추출하여 사용해야 하지만 서명 검증 시스템에서는 각 사람마다 모든 가능한 모조 서명을 구할 수 없기 때문에 서명 검증 시스템을 만드는 과정에서 부분적으로 얻게 되는 모조 서명의 샘플을 가지고 일반적으로 모든 서명에 사

용될 특징을 선택하는 방법을 적용하여 왔다[2]. 그러나, 이와 같이 모든 사람에게 동일한 특징집합을 사용하는 경우, 검증하려는 서명에 적합한 특징만을 모은 것이 아니기 때문에 사용하는 특징들중 어떤 것은 오히려 나쁜 결과를 초래할 수 있다.

본 논문에서는 검증하려는 서명에 대해 미리 획득한 진서명만을 이용하여, 그 서명의 검증에 적합한 특징 집합을 선택하는 문제를 다루고자 한다. 특징 집합의 선택에서는 각 특징들의 사용여부를 직접 판단하는 것이 아니라, 특징별로 검증하려는 서명에 적합한 정도를 판단하여 가중치를 적용한다. 즉, 유용하다고 판단되는 특징에는 큰 가중치를 설정하는데, 가중치는 그 서명에 있어 특징의 모조하기 어려운 정도를 구하여 사용한다. 본 논문에서는 특징의 모조하기 어려운 정도를 모조난이도(Degree of Forgery Difficulty; DFD)라고 하였고, 모조난이도를 판단하기 위해 다음의 두 분포를 비교한다. 첫 번째로 검증하려는 서명의 각 특징별 분포, 두 번째로 사람의 필기행동을 반영하기 위해 많은 서명을 이용해 얻은 각 특징별 일반적인 사람들의 분포를 사용한다. 즉, 일반 사람들의 필기 행동에 비추어 검증하려는 서명의 어떤 요소가 모조하기 어려운지를 분석하는 과정이다. 본 논문에서는 분포를 이용하여 특징들을 몇가지 유형으로 분류하고, 특징별로 모조난이도를 나타내는 함수를 도입한다. 이는 모조난이도를 통해 각 사람별로 검증에 유용한 특징을 알아내기 위함이며, 거리 계산에서 이 특징들을 강조하는 효과를 주기 위한 것이다.

본 논문에서는 기존의 방법과 제안한 방법을 비교 실험하였다. 2장에서는 특징 선택의 중요성과 문제점을 다루며, 3장에서는 가중치에 대한 설명과 서명별로 가중치를 설정하는 방법을 서술하였다. 그리고 4장에서는 비교실험을 다루고서 5장에 결론을 내렸다.

2. 특징 선택(Feature Selection)

매개변수는 진서명과 모조서명을 잘 구별하는 분별력에 따라 추출하여야 한다. 그러나 특징들의 물리적인 의미만으로 구분력을 판단하기에는 많은 어려움이 따른다. 그러므로 각 특징 벡터들에 대한 선험적인 지식이 없다면, 유용하다고 생각되는 특징들을 미리 선정하고 이 특징들중에서 시스템에 적합한 특징 집합을 선택하는 것이 바람직하다[2].

특징선택이란 이미 정의된 특징집합 $S_m =$

$\{f_1, f_2, \dots, f_m\}$ 이 있을때, 주어진 특징 성능 평가를 만족하며 가장 좋은 성능을 내는 특징 집합 S_n 을 찾는 문제이다[3].

여기서, $(S_n \subseteq S_m)$ (1)

특징 선택에는 '특징들의 부분집합(feature subset)의 성능을 비교할 성능 평가 함수로 무엇을 사용할 것인가'하는 문제와 'm개의 특징 벡터들이 존재할 때, 탐색해야 할 모든 경우의 수가 $2^m - 1$ 개 이므로 어떻게 탐색공간(search space)를 줄일 것인가'하는 두가지 문제가 있다[4].

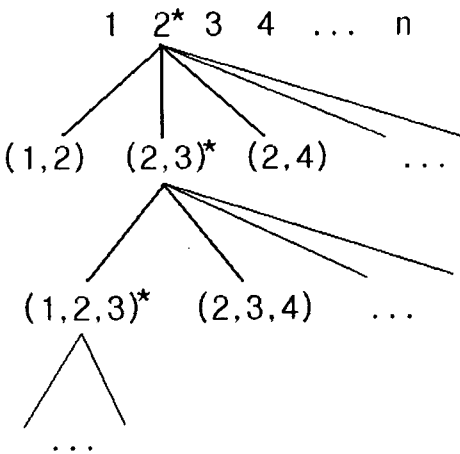
<표 1> 최초 선택한 특징 집합

특징번호			특징 설명	
X-좌표치	Y-좌표치	압력치		
0	1	2	속도의 평균	
3	4	5	서명의 크기/펜이 눌린 시간	
6	7	8	속도의 절대값의 평균	
9	10	11	양의 속도의 평균	
12	13	14	음의 속도의 평균	
15	16	17	평균 가속도	
18	19	20	가속도의 절대값의 평균	
21	22	23	속도의 분산	
24	25	26	속도의 절대값의 분산	
27	28	29	가속도의 분산	
30	31	32	최대 속도	
33	34	35	최대 가속도	
36	37	38	최소 속도	
39	40	41	최대 속도와 평균 속도의 차	
42	43	44	양의 최대 속도와 양의 평균 속도의 차	
45	46	47	양의 속도 빈도수	
48	49	50	음의 속도 빈도수	
51	52	53	속도의 영교차 빈도수	
54	55	56	양의 가속도 평균	
57	58	59	최대 속도와 최소 속도의 차	
60	누적된 각		68	획의 개수
61	누적된 각/총서명시간		69	펜의 들린 시간
62	속력의 평균		70	펜의 눌린 시간
63	서명의 X-Y 크기의 비		71	누적된 각/펜의 눌린 시간
64	입력 서명 모양의 고유치의 비		72	속력의 분산
65	서명의 무게중심에서의 영교차 비		73	최대 속력
66	서명의 무게중심에서의 좌우 크기의 비		74	평균속력에서의 영교차 수
67	서명의 무게중심에서의 상하 크기의 비		75	총 서명 시간

실제 문제에서는 많은 수의 특징들과 충분치 않은 데이터로 인해 각 특징들간의 확률 분포를 규명하기 어려우므로 대개 직접 조작 방법(direct manipulation)에 의해 성능을 평가하는 방법이 널리 사용되고 있다[2][3][5][6].

특징 선택의 두 번째 문제로 탐색 공간 축소이다. 이 문제는 개별적으로 최상의 성능을 내는 특징들만으로 문제에 적합한 특징집합(feature set)이 구성되지 않는기 때문에 생긴다[7]. 즉, 특징을 많이 사용할수록 특징집합의 크기에 비례하여 특징집합의 성능이 향상되지는 않는다. 그러므로 본 논문에서 최초 선택한 특징들 중 우수한 특징집합을 선택해야 하고, 최초 선정한 특징들은 표1에 제시된 바와 같이 기존의 매개변수를 이용한 서명 검증 시스템[2]에서 사용된 내부적 속성을 반영하는 특징들에 서명의 모양을 반영하는 특징들을 추가한 것이다.

문제에 가장 적합한 특징집합의 선정은 모든 경로의 탐색(exhaustive search)만을 통해 가능하나[8], 실제로 응용되기 어렵다. 탐색공간을 줄이는 방법으로 전향, 후향 탐색(forward, backward search)[5]과 분기와 경계선(branch and bound)[9], 등의 방법이 있다. 본 논문에서는 전향탐색법을 이용하여 특징집합을 선택하였다.



<그림 1> 전향 탐색

전향탐색이란, <그림 1>과 같이 n개의 특징들 중 최상의 구분력을 보이는 특징 하나를 선택한 후 그 특징을 포함한 다른 특징 집합들 중 가장 구분력이 좋은 특징집합을 선정하고, 이 특징집합을 포함하는 또 다른 특징집합을 선정하는, 이러한 과정을 반복하여 최상의 분별력을 갖는 집합을 선정하는 방법이다[5].

본 논문에서는 위에서 언급되는 실험에 의해서 전향 탐색을 이용하여 최초 76개의 특징들 중에서 23개 특징이 선택되었다. 이것은 각각의 서명에 가장 분별력이 우수한 특징 집합이 아니라 공통적으로 우수한 특징집합이며, 이를 '표준 특징 집합'이라고 명한다.

3. 개인별 가중치 설정 방법

사람마다 서명의 모양이 다르고 서명별로 진서명과 모조서명을 가장 잘 구별하는 특징 집합이 다르게 존재한다. 2장에서 선택된 표준 특징 집합은 실험에서 사용한 서명에서 공통적으로 좋은 결과를 낳는 특징 집합이었을 뿐이며, 각 서명별로 최상의 분별력을 갖는 특징집합은 아니다. 실제로 각 서명별로 모조서명을 만들어 모조서명과 진서명을 구분하는 분별력에 따라 특징집합을 선택할 경우, 유형 I 오류(진서명을 거부하는 오류)와 유형 II 오류(모조서명을 인정하는 오류)의 평균이 1% 이내의 결과를 보였다. 그러나, 서명 검증 시스템에 새로운 서명을 등록할 때마다 모조서명을 만들 수는 없으므로, 각각의 서명별로 모조서명을 이용하여 적합한 특징집합을 선정하는 것은 불가능하다. 그러므로, 진서명만을 이용하여 서명별로 적합한 특징집합을 선정하는 방법이 필요하므로, 본 논문에서는 검증하는 서명별로 특징마다 유용한 정도를 판단하는 가중치를 적용하고자 한다.

사용하는 특징이 N개이고, 검증하려는 특징 벡터를 \vec{x} 그 서명의 대표 벡터를 \vec{m} 이라 할 때, 가중치(p)를 적용한 거리는 다음과 같다.

$$Dist(\vec{x}, \vec{m}) = \sqrt{\frac{1}{P} \sum_{i=1}^N p_i \left(\frac{x_i - m_i}{\sigma_i} \right)^2}$$

여기서, $P = \sum_{i=1}^N p_i$ (2)

가중치를 설정하는데 검증하려는 서명의 각 특징별 분포와 사람의 필기행동을 반영하는 각 특징별 일반적인 사람들의 분포를 사용한다. 이 분포들을 이용하여 모조난이도를 구하는 예로서 속도특징에 대하여 알아보면, 사람이 모조를 하려고 할 때 모조하려는 서명이 빠른 속도로 쓰여진 것이면 천천히 쓰여진 서명보다 모조하기 어려울 것이다. 그러므로 검증하는 서명이 일반적인 다른 서명들과 비교하여 어느 정도 빠른 속도로 쓰여진 것인지를 판단한다면, 그 서명에 대한 속도 특징의 유용한 정도를 판단할 수 있다. 속도 특징의 경우 일반적인 분포의 중심 즉, 평균값보다 검증하려는 서명 분포의 중심이 더 큰 경우 그 서명에 있어 모조하기 힘든 특징이라고 분류할 수 있다.

각 사람의 특징 i 에 대한 모조난이도 DFD_i 는 다음과 같이 <표 2>의 4가지 요소의 함수로 표현된다.

<표 2> 모조난이도를 구성하는 요소들

기 호	의 미
m_i	많은 사람들의 서명을 통해 구한 I번째 특징에 대한 평균
m'_i	검증하려는 서명의 I번째 특징에 대한 평균
σ_i	많은 사람들의 서명을 통해 구한 I번째 특징에 대한 표준편차
σ'_i	검증하려는 서명의 I번째 특징에 대한 표준편차

$$DFD_i = f(m_i, m'_i, \sigma_i, \sigma'_i) \quad (3)$$

즉, 두가지 분포를 비교하는데 각 분포별로 평균과 표준편차를 이용한다.

본 논문에서는 DFD함수의 설정과 응용을 구분하여 설명한다.

3.1 모조난이도 함수의 설정

각 특징별로 함수를 설정하려면, 위의 두 분포가 어떻게 위치할 때 모조하기 어려운지 파악해야 한다. 그러나, 각 특징별로 이를 정확하게 구분한다는 것은 불가능하므로, 본 논문에서는 사람의 반응하는 형태가 시그모이드 함수와 유사하다고 생각하여 DFD함수는 다음의 형태를 따른다고 설정하였다.

$$DFD_i = \frac{1}{1 + \exp(-k * g_i)} \quad (4)$$

k 는 가중치 함수의 기울기와 관련된 계수로 값이 클수록 기울기가 크게 된다. 실험을 통하여 값을 변화시켜 가며 성능에 따라 결정한다. g_i 는 특징별로 어느 경우 모조하기 어려운지를 반영하는 함수로서, 본 논문에서는 표 3의 4가지 유형으로 분류한다. 4가지 중에는 분포의 중심(평균)을 이용하여 비교한 경우 세가지와, 분포의 폭(표준편차)을 이용하여 비교한 경우가 있다. 각 특징별로, 검증하려는 분포의 중심이 일반적인 분포의 중심보다 작을 때 모조가 힘든 특징인 경우(유형 1)와 클 때 모조가 힘든 특징인 경우(유형 2), 그리고 검증하려는 분포의 중심이 일반적인 분포의 중심으로 부터 멀리 위치할 때 모조가 힘든 특징인 경우(유형 3)와, 일반적인 분포와 비교할 때 검증하려는 분포의 폭이 좁을 때 모조하기 힘든 경우로 분류된다. 표 3의 λ_1 과 λ_2 는 4가지 유형의 함수 f_i 의 값을 표준화하기 위한 것으로 λ_1 을 1로 할 때 상대적인 값으로 λ_2 를 설정할 수 있다. 즉, 평균을 비교한 유형 1, 2, 3과 관련된 특징들 값의 분포를 비교하여 구한다.

유형 1, 2, 3은 일반적인 특징의 분포 중심으로 부터 검증하려는 서명 특징의 분포 중심이 떨어진 정도에 관련되어 모조난이도가 결

정되므로 일반적인 특징분포의 중심에서 시작하는 모조난이도 함수와 검증하려는 서명의 중심과 만나는 지점의 값이 모조난이도라 할 수 있다. <표 3>의 유형 2는 앞에서 언급한 속도특징이라 할 수 있고, 검증서명의 중심이 모든 서명의 중심과 일치할 때 모조난이도 함수값은 1/2이 된다. 또한 검증서명의 중심과 일치할 때 모조난이도 함수값은 1/2이 된다. 또한 검증서명의 중심이 양의 방향으로 위치할수록 큰 모조난이도 함수값을 갖는다. 모조난이도 함수와 검증하려는 서명의 중심과 만나는 지점의 값이 모조난이도이다. 마찬가지로 유형 4의 한 예로서 서명의 X-Y 크기비 특징을 고려할 수 있다. 검증하려는 서명에 유용한 특징이라 할 수 있으므로, 일반적으로 어느 정도의 표준편차가 이 값보다 작게 되면 유용한 특징이라 할 수 있다.

각 특징이 앞에서 설정한 어느 유형에 속하는지를 구분하는 과정은 다음 규칙들을 바탕

으로 이루어진다.

- 규칙 1) 속도와 가속도 특징은 특정한 서명의 평균이 클수록 모조가 힘들다. <유형 2>
- 규칙 2) 모조서명자는 모양을 묘사하는데 치중하므로 압력과 같은 보이지 않는 요소는 일반적인 평균과 차이가 클수록 모조가 힘들다. <유형 3>
- 규칙 3) 모조서명자의 경우 한 획이 끝난 후 다음획 시작까지의 이동 시간이 길기 때문에 펜이 들린 시간이 작을수록 모조가 힘들다. <유형 1>
- 규칙 4) 서명의 X-Y비나 음으로 움직이는 횟수 같은 경우 서명의 모양을 반영하므로, 특징의 평균보다는 얼마만큼 일정한 모양을 그리는지를 나타내는 표준편차의 의미가 더 크다. 표준편차가 작을수록 모조가 힘들다. <유형 4>
- 규칙 5) 총 서명시간과 같은 시간관련 특징은

<표 3> 모조난이도의 4가지 유형

유형	모조가 난해한 상황	g_i	DFD_i 의 곡선
1	$m_i' < m_i$	$\lambda_1 \frac{m_i - m_i'}{\sigma_i}$	
2	$m_i' < m_i$	$\lambda_1 \frac{m_i' - m_i}{\sigma_i}$	
3	$m_i' < m_i$ $m_i' < m_i$	$\lambda_1 \frac{ m_i - m_i' }{\sigma_i}$	
4	$\sigma_i' < \sigma_i$	$\lambda_2 \frac{\sigma_i - \sigma_i'}{\sigma_i}$	

모양과 관련 특징으로 고려할 수 있다.

<유형 4>

규칙 6) 특정 특징값의 표준편차를 특징으로 사용한 경우 평균이 작을수록 변화가 작은 것이므로 모조가 난해하다. <유형 1>

규칙 7) 속력의 변화 (영교차)가 많을수록 모조가 난해하다. <유형 2>

규칙 8) X,Y축의 속도가 가속도의 변화(영교차)는 좌우 이동을 반영하므로 모양관련 특징으로 분류할 수 있다. <유형 4>

규칙 9) 압력은 z축으로 볼 수 있으므로 모양관련 특징으로 고려하여, 음의 값으로 나온 횡수와 같이 모양을 반영하는 경우 압력도 모양 관련 특징으로 동일하게 취급할 수 있다. <유형 4>

규칙 10) 평균과 최대값, 최소값의 비는 클수록 모조가 난해하다. <유형 2>

3.2 DFD 함수의 응용

특징별로 식4의 형태를 갖는 DFD함수를 설정할 수 있다. DFD함수를 이용하여 임의의 서명을 검증하려면, 식2의 거리식에서 특징 i 에 대한 가중치 p_i 에 대응하는 DFD_i 를 적용하여 구할 수 있다.

$$Dis(\vec{x}, \vec{m}) = \sqrt{\frac{1}{DFD} \sum_{i=1}^N DFD_i \left(\frac{x_i - m_i}{\sigma_i} \right)^2} \quad (5)$$

$$\text{여기서, } DFD = \sum_{i=1}^N DFD_i$$

즉, 검증하는 서명에 대한 각 특징의 유용한 정도인 DFD_i 를 구하여 가중치로 이용된다.

4. 실험 및 비교

본 논문에서는 기존의 매개변수를 이용한 결과와 제안한 가중치를 적용한 결과의 비교 실험을 행하였다. 실험환경으로, 한 사람의 서

명자로부터 얻은 125개의 진서명과 한 사람의 모조실험자로부터 얻은 125개의 모조서명에서 20개의 진서명 학습표본, 100개의 진서명 실험표본, 100개의 모조서명 실험표본을 임의적으로 추출하였다. 이것으로부터 얻어진 결과를 한 프레임이라고 했을 때, 통계적인 신뢰성을 가하기 위해 한사람의 진서명과 그에 대한 모조서명에 대해 12개의 프레임을 얻었다. 총 9사람의 자료를 얻었으므로 각 특징 집합에 대해 108개의 프레임이 설정된다. 그러므로 전체특징의 결과를 얻기 위해 540개의 프레임의 결과를 구했다.

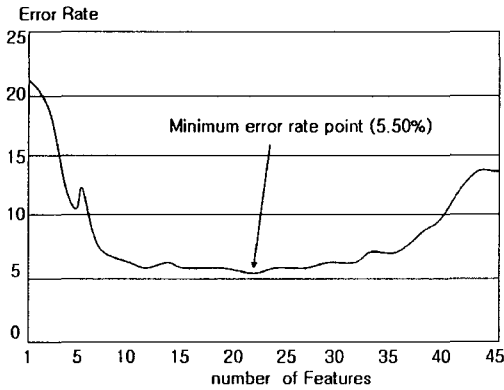
본 논문에서는 임계치의 결정을 자신의 변화도에 따라 다르게 주는 방법을 적용하였다. 즉 학습용 서명에서 각 서명끼리의 변화가 큰 경우에는 큰 임계치를 주는 방법이다. 변화도를 μ 로 표시하면 임계치 θ 의 결정은 $\theta = c * \mu$ 가 되고 이때 계수 c 는 실험을 통해 구하여 모든 서명에 공통으로 적용하였다[10].

실험은 다음의 세단계로 나누어 실시하였다.

1. 전향탐색을 이용하여 표1의 76개 특징들 중 서명검증에 유용한 특징집합을 선정하고 이때의 오류율을 구한다.
2. 단계 1에서 구한 특징집합을 대상으로 제안한 DFD를 적용하여 오류율을 구한다.
3. 단계 1과 단계 2의 오류율을 비교한다.

첫번째 실험으로 기존의 방법을 적용한 경우는, 2장에서 언급한 전향탐색을 이용하여 표 1의 최초 선택한 특징 76개 중에서 서명검증에 적합한 특징집합을 선정한다. <그림 2>는 전향탐색을 이용하여 특징의 수를 늘려가며 오류율을 고려한 경우로 23개의 특징을 사용할 때 최소 오류율을 나타내고 있고, 특징의 수가 증가해도 오류율이 더 이상 감소하지 않는 것을 알 수 있다.

이때의 23개의 특징은 <표 4>와 같고 이를 표준 특징 집합이라 한다.



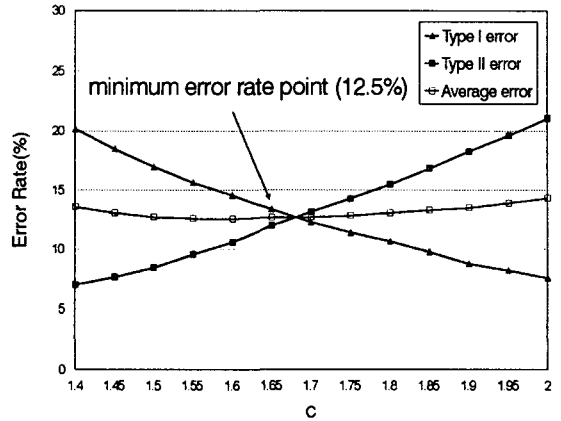
<그림 2> 전향탐색에 의한 결과

<표 4> 4개의 유형으로 분류한 표준 특징 집합

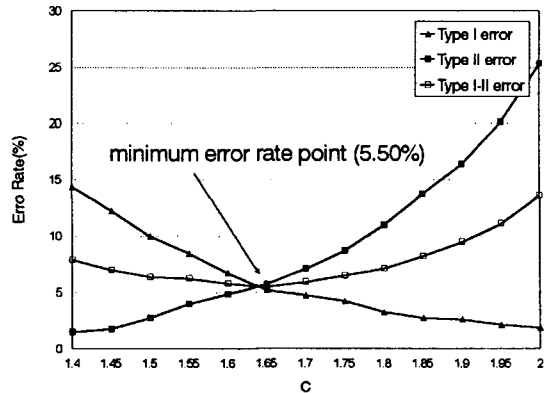
특징 번호	특징	적용한 유형
0	X축의 평균 속도	유형4
1	Y축의 평균 속도	유형4
2	압력의 평균	유형3
6	X축 속도의 절대값의 평균	유형2
25	Y축 속도의 절대값의 표준편차	유형1
31	Y축 최대 속도	유형2
34	Y축 최대 가속도	유형4
38	압력 최소 속도	유형3
40	Y축 최대속도와 평균속도의 차	유형2
43	Y축 최대속도와 양의 평균속도의 차	유형2
47	압력 양의 속도 빈도수	유형4
48	X축의 음의 속도 빈도수	유형4
49	Y축의 음의 속도 빈도수	유형4
50	압력 음의 속도 빈도수	유형4
52	Y축 속도의 영교차 빈도수	유형2
56	압력 양의 가속도 평균	유형1
60	누적된 각	유형4
61	누적된 각/총 서명시간	유형4
65	서명의 무게중심에서의 영교차	유형4
66	서명의 무게중심에서의 좌우 크기의 비	유형4
67	서명의 무게중심에서의 상하 크기의 비	유형2
70	펜의 눌린 시간	유형4
74	평균속력에서의 영교차 수	유형4

<그림 3>은 최초의 특징집합인 76개의 특징을 모두 사용한 경우로서, 임계치를 늘려가며 오류율을 구한 결과이다. 임계치 계수 c 가 1.70 일때 최소 오류율 12.5%를 보이고 있다. <그림 4>는 <그림 2>의 최소의 오류율을 내는 23개

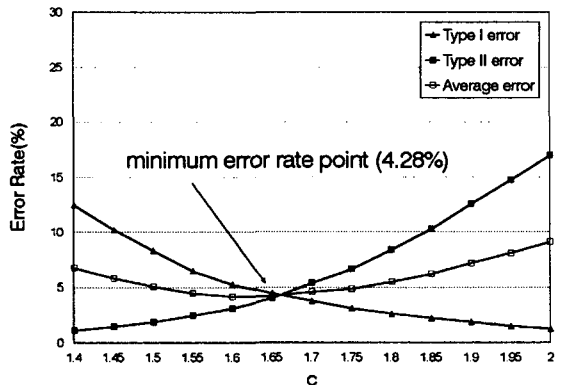
의 특징 집합을 사용할 경우이다. 계수 c 가 1.65일 때 최소 오류율 5.50%를 보이고 있다.



<그림 3> 최초의 특징집합(76개)을 모두 사용한 경우의 검증 오류율



<그림 4> 최적의 23개의 특징을 사용한 경우



<그림 5> 개인별 가중치를 사용하는 경우

마지막 실험으로, 동일한 표준 특징 집합을 사용하며, <표 3>에 기술한 바와 같이 제안한 가중치를 적용한 경우의 결과는 <그림 5>와 같다. 제안한 방법을 적용한 경우, 임계치를 나타내는 계수 c 가 1.59일 때 최소 오류율 4.28%를 보이고 있다.

<표 5> 실험 결과 비교

사용된 특징	76개의 최초 특징	23개의 최적화 특징	개인별 가중치 적용한 특징
최소 오류율	12.5%	5.50%	4.28%

<표 5>에서 보이듯이, 76개의 모든 특징을 사용하는 경우 최소 오류율 12.5%의 결과를 보였으며, 서명검증에 유리한 특징으로 23개의 특징을 걸러낸 경우 최소 오류율 5.50%의 더 좋은 결과를 내었고, 개인별 가중치를 적용한 경우에 최소 오류율 4.28%의 결과로 더욱 향상된 검증률을 얻을 수 있었다.

5. 결론

본 논문에서는 각 서명자 별로 다른 특징을 적용하는 방법을 제안하였다. 검증에 사용하는 각 특징에 대해 진서명의 분포와 대부분의 서명이 분포하는 형태를 비교하여 네 가지 경우로 구분하고, 이러한 구분을 통해 각 특징 별로 검증하려는 서명의 중심과 모든 서명의 중심과의 거리로서 그리고 검증하려는 서명의 표준 편차와 모든 서명의 표준 편차의 비로 모조난이도를 설정하여 특징간의 가중치를 사용한다. 서명의 진위 여부는 특징들의 대표값과의 거리를 비교해 행하고 모조의 난이도를 통해 각 서명별로 유용한 특징이 거리의 중요한 부분을 차지하도록 한다. 이러한 방법을 통해 기존의 방법[2]의 오류율 (12.5%)보다 낮은 오류율(4.28%)을 얻었다.

일반적인 서명을 나타내는 평균과 표준편차

를 실험에 사용한 서명으로부터 얻었지만 더 많은 서명의 획득을 통해 이를 결정한다면 더욱 향상된 결과를 얻을 수 있을 것이다.

참고문헌

- [1] Rejean Plsmondon and Guy Lorette, "Automatic Signature Verification and Writer Identification The State of The Art", Patten Recognition, vol. 22, no. 2, pp. 107-131, 1989.
- [2] Hewitt D. Crane and John S. Ostrem, "Automatic Signature Verification Using a Three Axis Force-Sensitive Pen", IEEE Trans. on Syst., Man, Cyber., vol. 13, no 3, pp. 329-337, May/june, 1983.
- [3] Lian Li, "A Hypothesis Testing Approach to word Recognition Using Dynamic Feature Selection "Proc.of 11th.Conf.on Pattern Recognition, pp. 586-589, 1992.
- [4] Manadu Ichino and Jack Sklansky. "Optimum Feature Seletion by Zero-One Integer Programing", IEEE Trans. on Syst., Man, Cyber., vol. 14 no. 5, pp. 737-746, September /October. 1984.
- [5] Laveen .Kanal, "Patterns in Pattern Recognition : 1968-1974", IEEE trans. on Information Theory. vol. IT-20, no. 6, pp. 697-722, November, 1974.
- [6] Dobald H. Foley, "Consideration of Sample and Feature Size", IEEE Trans. on Information Theory, vol. IT-18, no. 5, pp. 618-626, September, 1972.
- [7] T. M Cover, "The Best Two Independent Mesurements are not the Two best", IEEE Trans. on Syst., Man, Cyber., pp. 116-117, January, 1974.
- [8] T. M Cover, "Computational Complexity Aspects of Dimensionality Reduction and Classification ". IEEE Trans. on

- Syst., Man, Cyber., pp. 402-771, October, 1971.
- [9] Keinosuke Fukunaga, Introduction to Statistical Pattern Recognition. 2nd edition, Academic Press, 1994.
- [10] J. Kim, J.R. Yu and S. H. Kim, "Learning of Prototypes and Decision Boundaries for a Verification Problem have only Positive Samples," Pattern Recognition Letters, Vol.17, pp.691-697, 1996.

주 작 성 자 : 김 성 훈

논 문 투 고 일 : 2005. 08. 30

논 문 심 사 일 : 2005. 08. 31(1차), 2005. 09. 08(2차)
2005. 09. 09(3차)

심 사 판 정 일 : 2005. 09. 09

● 저자소개 ●



김성훈

1988 서강대학교 이공대학 전자공학과 공학사
1990 연세대학교 공과대학 전자공학과 공학석사
1996 연세대학교 공과대학 전자공학과 공학박사
1996 ~ 현재 영동대학교 컴퓨터공학과 교수
관심분야: 패턴인식, 서명검증, 생체인증, 지능시스템