

구간분할 매칭방법과 선형판별분석기법을 융합한 온라인 서명 검증

(On-line Signature Verification using Segment Matching and LDA Method)

이 대 종 † 고 현 주 ** 전 명 근 ***
(Dae-Jong Lee) (Hyoun-Joo Go) (Myung-Geun Chun)

요 약 기존의 참조서명과 입력서명을 비교하는 방법 중 분절 단위 비교 방법은 전역적 비교와 점 단위 비교 방법과 비교하여 우수한 장점을 가지고 있다. 그러나 분절 단위 비교 방법은 인식률과 직접적인 관계가 있는 분절의 불안정 문제점이 있다. 이러한 문제점을 해결하기 위하여 본 논문에서는 분절단위 방법 외에 선형판별분석에 의한 매칭방법을 고려한 서명 검증 기법을 제안한다. 최종 검증단계에서 두 개의 독립모델을 효과적으로 융합할 수 있는 확률기반의 베이지안 분류기를 적용하였다. 다양한 서명데이터를 이용하여 실험한 결과 제안된 기법은 분절단위 기반 구간분할매칭 기법에 비해 우수한 성능을 나타냈다.

키워드 : 온라인 서명검증, LDA, 구간 분할 매칭, 베이지안 식별기, 융합

Abstract Among various methods to compare reference signatures with an input signature, the segment-to-segment matching method has more advantages than global and point-to-point methods. However, the segment-to-segment matching method has the problem of having lower recognition rate according to the variation of partitioning points. To resolve this drawback, this paper proposes a signature verification method by considering linear discriminant analysis as well as segment-to-segment matching method. For the final decision step, we adopt statistical based Bayesian classifier technique to effectively combine two individual systems. Under the various experiments, the proposed method shows better performance than segment-to-segment based matching method.

Key words : On-line signature verification, LDA, Segment matching, Bayesian classifier

1. 서 론

바이오메트릭스(Biometrics) 기술에서 사용하고 있는 신체적 특징 중에서 서명은 오래 전부터 계약체결 등의 서류에 대한 증빙 목적으로 사용될 만큼 사용자에게 익

숙한 인증수단으로서, 음성, 홍채 등의 다른 생체인식기법에 비해 소형 태블릿과 같은 가격 경쟁력을 지닌 입력장비만을 필요함으로써 가격면에서도 우위를 보일 뿐만 아니라 최근에는 태블릿 마저 필요 없는 PDA의 등장과 활성화로 서명인식에 관한 연구가 주목을 받고 있다[1,2]. 그러나 서명인식의 경우 생체인식을 위한 다른 특징들과 달리 사칭자에 의해 전문 위조가 가능하고 동일인의 서명조차도 외부 환경 또는 내적인 요인에 의해 서명의 변이 차이로 시스템의 성능이 저하되는 문제점을 지니고 있다[3].

현재까지의 온라인 서명검증방식은 참조서명(Reference Signature)과 입력서명(Input Signature)의 비교 방식에 따라 크게 전역적 특징 비교, 점 단위 비교, 분절 단위 비교방법이 있다[4]. 이 중 분절 단위 비교방법은 참조서명과 입력서명을 각각 적절한 분절단위로 분할한 후, 분절간의 최적의 대응관계를 찾는 방식으로, 정적정보 또는 동적정보를 이용하여 기준 분할이 이루어진다.

† 정 회 원 : 충북대학교 BK21 충북정보기술사업단 교수
dimidori@empal.com

** 학생회원 : 충북대학교 제어계측공학과
ghjswy@chungbuk.ac.kr

*** 정 회 원 : 충북대학교 전기전자컴퓨터공학부 교수
mgchun@chungbuk.ac.kr

논문접수 : 2004년 5월 31일

심사완료 : 2007년 12월 4일

: 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 소프트웨어 및 응용 제34권 제12호(2007.12)

Copyright © 2007 한국정보과학회

여기서, 정적정보는 굴곡점, 봉우리 등의 정적특징을 이용하여 분절 단위로 기준 분할이 이루어지며[4-7], 동적정보는 최소 속력점, 압력 등의 동적특징을 이용하여 입력서명을 분절 단위로 분할한다[8,9]. 이와 같은 분절단위 비교방법은 분절간에 지역적 비교가 가능하며, 지역적으로 통계적 모델을 사용할 수 있어 분석이 용이한 장점을 지니고 있는 반면에 서명의 변이로 인해 모든 서명을 같은 분절단위로 분할하기 어려워 진서명간에도 분할점이 불안정할 수 있다는 문제점을 지니고 있다. 이러한 문제점은 진서명간의 비교하는 특징점 위치가 서로 상이하여 인식률과 직접적인 관계가 있다.

본 논문에서는 기존의 분절 단위 비교만을 이용한 방법에 대한 문제점을 보완하기 위해 선형판별분석(Linear Discriminant Analysis, LDA)기법과 융합하여 서명인식의 신뢰도를 높이는 온라인 서명 검증 기법을 제안한다. 제안한 방법에서 분절 단위 비교 방법으로는 구간 분할 매칭 방법을 사용하였고, 선형판별분석 기법은 생체인식 분야에서 주로 얼굴인식에 사용되었던 방법으로 서명에 대하여 진서명간에 일어나는 변이에 대해서는 흡수면서 다른 서명과의 차이를 높이는 특징을 추출한다. 최종 검증 단계에서는 기존에 다중 센서 시스템에서 사용되고 있는 베이저안 식별기(Baysian Classifier)에 의해 융합한다.

본 논문의 구성은 다음과 같다. 2장에서는 온라인 서명 검증 기법에 사용된 알고리즘을 설명한다. 3장에서는 제안된 온라인 서명 검증 시스템에 대한 설명하고 4장에서는 실험 결과를 바탕으로 제안한 방법의 타당성을 검토한다. 마지막으로 5장에서는 결론을 맺는다.

2. 구간분할 매칭방법과 선형판별분석 및 베이저안 식별기

2.1 구간분할 매칭방법

분절단위 비교방식에 의하여 서명 데이터를 분할할 경우 참조서명과 입력서명간에 비교되는 분할점, 즉 특징벡터를 안정적으로 추출하는 것은 검증성과 직결될 만큼 중요시 다루어지는 부분이다. 본 논문에서는 동일 서명간에 변이가 적은 1차원상의 Y축의 봉우리를 기준 분할점으로 선택한 후 분할점 간에 위치한 특징벡터인 봉우리·계곡점들을 비교함으로써 구간적으로 안정된 분할이 이루어지도록 제안된 방법을 사용하였다[9]. 그 동작특성을 살펴보면 다음과 같다.

그림 1에서 보는 바와 같이 동일한 서명이라 할지라도 서명 당시의 감정상태 또는 주변 환경 등에 의하여 Y축상의 봉우리 또는 계곡의 불일치가 발생한다. 그러나, 그림 1의 (b)에 P1, P2, P3로 나타낸 봉우리는 두 서명간의 변이에 상관없이 동일한 위치에 존재함을 알

수 있다. 즉, 서명마다 갖고 있는 고유의 성질로 인하여 서명의 변이에도 변하지 않는 P1, P2, P3와 같은 우세 봉우리가 존재한다. 이러한 우세 봉우리점을 기준 분할점으로 하여 구간별로 두 서명간의 봉우리와 계곡을 매칭시킬 경우 안정적으로 분할매칭이 가능하다. 따라서, 기존 논문에 제시되었던 P1, P2, P3 지점의 우세 봉우리를 PPP(Partitioning Peak Points)을 이용하여 입력서명의 구간별 분할매칭을 하였으며, 그 과정을 단계별로 살펴보고자 한다.

단계 1. 참조서명에 대한 PPP 및 계곡·봉우리점을 설정한다. 여기서, PPP점은 그림 1에서 보는 바와 같이 서명의 봉우리를 중에서 봉우리의 값이 큰 순서대로 3개의 봉우리를 선택하였다.

단계 2. 입력서명에 대한 Y축의 봉우리점과 계곡점을 찾은 후 봉우리점을 대상으로 입력서명에 대한 PPP점을 선택한다. 여기서, 입력서명의 PPP를 선택 시 정규화된 Y 변위 값과 시간에 대하여 참조서명의 PPP와 가장 근접한 위치에 있는 봉우리점을 찾는다.

단계 3. 참조서명의 PPP점과 입력서명의 PPP점을 기준으로 한 구간분할 매칭에 의해 입력서명의 계곡점과 봉우리점을 선택한다. 그림 1에서 보는 바와 같이 PPP점에 의해 분할된 구간 a-a', b-b', c-c', d-d'로 각각 나누어서 매칭이 이루어진다. 매칭 방법은 PPP를 지정할 때와 마찬가지로 정규화된 Y변위와 시간을 이용하여 참조서명의 봉우리점 또는 계곡점과 가장 근접한 위치에 있는 입력서명의 봉우리점과 계곡점을 찾는다.

매칭과정에서 비교서명보다 참조서명의 데이터가 많은 경우 참조서명 데이터의 Y변위와 시간을 비교서명과 비교하여 근접도가 낮은 데이터를 삭제하며, 비교서명의 데이터가 많은 경우 반대의 과정을 수행한다. 또한 매칭도를 높이기 위해 두 서명의 근접도를 비교 시 봉우리와 계곡을 분리하여 비교한다. 그림 2는 매칭된 결과 그림을 나타낸다. 그림 2에서 보는 바와 같이 구간분할매칭에 의해 참조서명의 데이터 d 부분이 삭제되었음을 확인할 수 있고, 참조서명과 입력서명 모두 안정된 분절단위로 매칭이 이루어졌음을 확인할 수 있다.

2.2 선형판별분석에 의한 서명 특징추출

선형판별분석 기법은 분류를 위해 최적의 투영을 찾는 방법 중의 하나이다. 투영된 분산의 최대화하는 투영을 찾기보다는 선형판별분석기법은 클래스내의 분산을 나타내는 행렬과 클래스간 분산을 나타내는 행렬의 비가 최대가 되는 투영 행렬을 찾는 것이다. 일반적으로 선형판별분석기법을 이용하여 분류하고자 할 경우 최적의 투영행렬을 계산하기 전에 클래스내 분산을 나타내는 행렬이 정칙이 되도록 주성분분석기법을 이용하여 입력서명을 저차원 공간으로 투영한다.

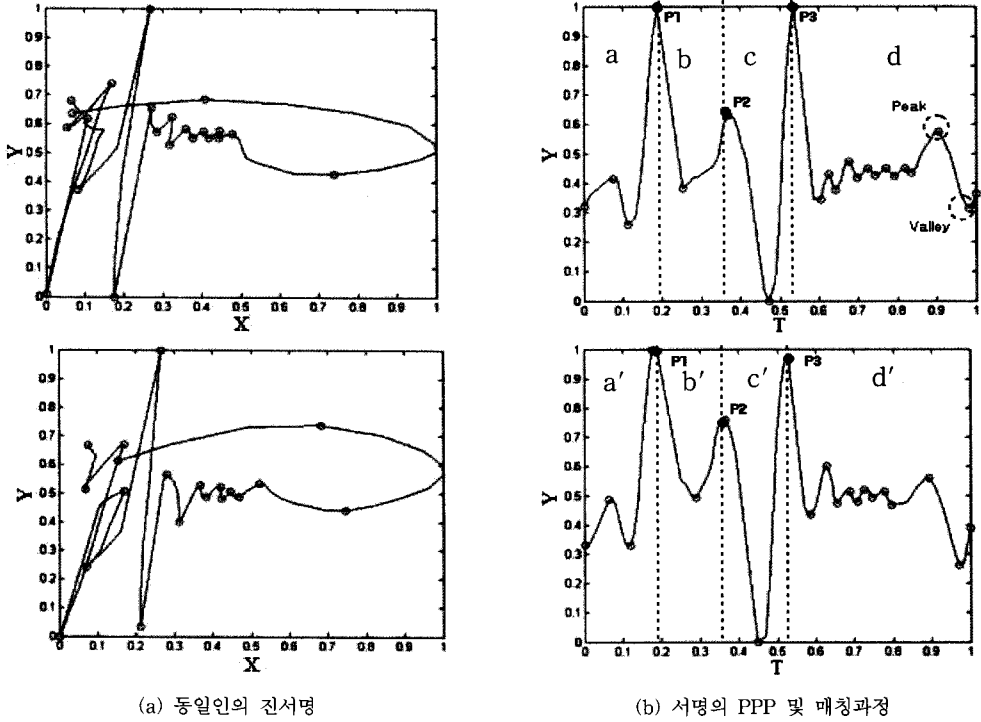


그림 1 진서명간의 봉우리와 계곡

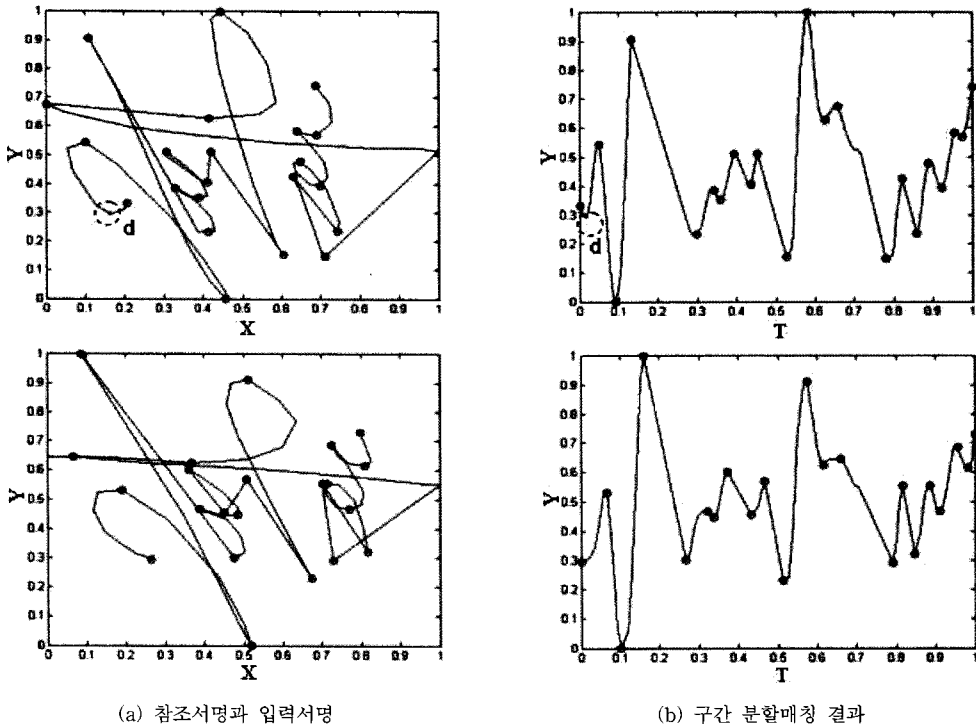


그림 2 참조서명과 입력서명의 구간분할

본 논문에서 입력 서명을 선형판별분석기법을 이용하여 특징 추출하는 과정을 설명하기 위해 입력신호의 차원축소에 사용되는 주성분분석기법과 특징추출을 위해 사용된 선형판별분석 기법에 대하여 설명하면 다음과 같다[10,11].

우선 타블렛으로부터 입력서명의 X 축 및 Y 축의 X, Y 좌표값과 좌표값에 해당하는 시간 정보를 획득한 후 다음 식과 같은 서명의 입력값을 취득한다.

$$s_i = [x_1, y_1, t_1, x_2, y_2, t_2, \dots, x_i, y_i, t_i]^T \quad (1)$$

여기서, x_p, y_p 은 시간 t_i 에서의 온라인 입력서명의 X 축과 Y 축의 좌표값을 각각 의미한다.

학습서명의 개수가 P 개일 경우 1차원 열벡터의 학습 벡터를 s_1, s_2, \dots, s_p 로 표시하면, 학습에 사용된 서명의 평균벡터는 다음과 같이 표현된다. 여기서, 입력서명의 특징벡터 s_i 가 $N \times 1$ 인 1차원 열벡터일 경우 평균벡터 m 도 ($N \times 1$)인 1차원 열벡터 크기를 갖는다.

$$m = \frac{1}{P} \sum_{i=1}^P s_i \quad (2)$$

이 식에 사용된 학습서명에 대한 1차원 벡터의 차를 구하면 다음식과 같이 표현된다.

$$\begin{aligned} \bar{S} &= [\bar{s}_1 : \bar{s}_2 : \dots : \bar{s}_p] \\ \bar{s}_i &= s_i - m \end{aligned} \quad (3)$$

식 (3)에서 $N \times P$ 행렬로 이루어진 \bar{S} 에 대한 공분산 행렬 Ω 는 식 (4)와 같으며, 이 공분산행렬에 대한 고유값과 고유벡터를 식 (5)에 의해 산출한다.

$$\Omega = \bar{S} \bar{S}^T \quad (4)$$

$$\bar{S} \bar{S}^T v_j = \lambda v_j \quad (5)$$

식 (5)에서 구해진 고유벡터 v_j 을 이용하여 학습신호 x_i 에 대한 고유벡터를 구하면 다음 식과 같다. 즉, 입력신호에 대한 특징벡터는 공분산에 대한 고유벡터를 이용하여 주성분분석기법에 의해 변환된 공간으로 다음과 같이 투영함으로써 얻을 수 있다.

$$z_i = v_i^T (s_i - m) \quad (6)$$

공분산 행렬 Ω 의 사이즈 $N \times N$ 가 클 경우 스냅샷(snap-shot) 방법을 이용하여 효율적으로 공분산 행렬의 고유벡터를 구한다[10].

한편, 선형판별분석 기법은 클래스 내의 분산을 나타내는 행렬 S_W 와 클래스 간 분산을 나타내는 행렬 S_B 의 비율이 최대가 되는 행렬을 구하여 이 행렬을 이용하여 선형 변환을 하는 기법으로, 두 행렬을 식으로 정의하면 각각 식 (7) 및 식 (8)과 같다.

$$S_B = \sum_{i=1}^c n_i (m_i - m)(m_i - m)^T \quad (7)$$

$$S_W = \sum_{i=1}^c \sum_{s \in C_i} (s - m_i)(s - m_i)^T \quad (8)$$

윗 식에서 n_i 는 i 번째 클래스 C_i 에서의 데이터 수이며, m 은 모든 데이터의 평균이다. 또한 m_i 는 클래스 C_i 에서 주성분분석기법에 의해 변환된 데이터의 평균이다.

최적의 투영행렬 W 는 식 (9)와 같이 투영된 샘플의 클래스 내의 분산을 나타내는 행렬 S_W 에 대한 클래스 간 분산을 나타내는 행렬 S_B 의 비를 최대화하는 정규 직교 행렬을 가진 행렬로서 선택되어진다.

$$W = \operatorname{argmax}_W \frac{|W^T S_B W|}{|W^T S_W W|} = [w_1, w_2, \dots, w_m] \quad (9)$$

여기서, w_i 는 식 (10)과 같이 S_B 와 S_W 에 대한 일반화된 고유벡터의 집합이다.

$$S_B w_i = \lambda_i S_W w_i, \quad i = 1, 2, \dots, m \quad (10)$$

최종적으로 입력된 서명 s_i 에 대한 특징벡터 $F = (F_1, F_2, \dots, F_N)$ 는 식 (11)에 의해 얻어진다.

$$F_i = W^T z_i = W^T v_i^T (s_i - m) \quad (11)$$

2.3 베이시안 식별기

베이스 정리(Bayes theorem)는 사전확률이 추가적 정보의 조건부확률인 사후확률로 나타나는 방식으로 실험의 결과로써 얻은 정보를 토대로 하여 어떤 사건의 알려지지 않은 확률을 구할 수 있다. 클래스 분류 문제에 있어서 베이스 정리를 이용한 식별규칙을 베이시안 식별기라하며 사전확률만을 이용한 분류기 보다 많은 정보에 근거하여 더 우수한 성능을 보인다. 본 논문에서의 융합기법은 Ben-Yacoub 연구에서 사용한 최소 비용 베이시안 식별기(minimum cost bayesian classifier)로 다중 생체 인식과 다중 감지 장치에서 적용되었다[11,12]. 베이스 식별규칙은 클래스 분류를 위한 합리적인 결정을 내리기 위해 식별에 대한 손실로서 위험부담을 적용한다. 위험부담은 특징벡터 x 에 대해 $\{w_1, \dots, w_c\}$ 의 클래스를 분류하는 경우 식 (12)과 같이 나타낼 수 있다.

$$R(x) = \sum_{j=1}^c \int_{D_j} \left(\sum_{i=1}^c l_{ij} p(x|w_j) \right) d_x \quad (12)$$

식 (12)에서 $R(x)$ 는 모든 클래스 결정에 대한 전체 위험부담을 가리키며 D_j 와 l_{ij} 는 각각 결정영역과 결정에 대한 손실을 나타낸다. 베이스 식별규칙은 $R(x)$ 를 최소화하는 결정을 따르며 이에 기반한 결정규칙을 일반화 베이스 결정규칙(Bayes decision rule)이라 한다. 그리고 베이스 결정규칙으로 찾아진 최소 위험부담을 베이스 위험부담(Bayes risk)이라고 한다. 두 클래스 문제에 있어서 오차를 최소를 하기 위해서 베이스 식별규칙은 식 (13)에 의해 결정되며 결정경계는 그림 3에서는 이중 사선이 없는 점선이 결정된다. 그림 3에서 D_1

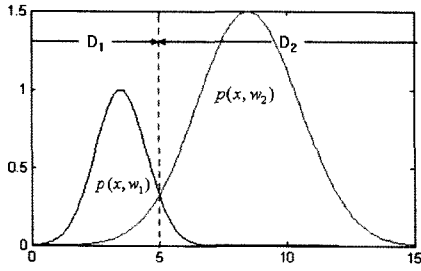


그림 3 1차원의 두 클래스 식별 문제에 대한 최적 결정 경계

과 D_2 는 결정경계에 대한 결정영역이 된다.

$$p(x|w_2)P(w_2) = p(x|w_1)P(w_1) \quad (13)$$

Ben-Yacoub 연구에서 적용된 최소 비용 베이즈안 식별기는 우도함수(likelihood function)를 베이스 정리에 적용하여 사후확률을 식 (14)과 같이 나타내었다[11].

$$p = \left\{ 1 + \left[\frac{g}{1-g} \frac{f(x|w=1)}{f(x|w=0)} \right]^{-1} \right\}^{-1} \quad (14)$$

식 (14)에서 $f(x|w)$ 는 우도함수이고 g 는 이진 사건으로 표현된 w 가 1일 때의 확률로 사전확률이다. 또한 우

도함수는 각 x 집합이 독립일 경우 $\prod_{i=1}^n f_i(x_i|w)$ 와 같이 표현할 수 있으며 식 (14)은 식 (15)와 같이 로컬 우도율(local likelihood rate) 함수로 나타낼 수 있다.

$$p = \left\{ 1 + \left[\frac{g}{1-g} \prod_{i=1}^n \frac{f_i(x_i|1)}{f_i(x_i|0)} \right]^{-1} \right\}^{-1} \quad (15)$$

식 (15)에 대한 베이스 위험 부담을 최소화하는 결정 규칙은 식 (16)을 따르며 C 는 0과 1의 값을 갖는 비용 함수이다.

$$a = \begin{cases} 1, & \text{if } \frac{f(x|w=1)}{f(x|w=0)} > \frac{1-g}{g} \frac{C_{10}-C_{00}}{C_{01}-C_{11}} \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

3. 선형판별분석기법과 구간분할 매칭방법을 융합한 온라인 서명 검증 시스템 구현

본 논문에서는 기존의 구간 분할 매칭방법의 성능 향상을 위해 선형판별분석기법과 융합한 새로운 온라인 서명 검증시스템을 제안하였다. 제안한 온라인 서명 검증 구성은 그림 4와 같이 데이터 획득, 전처리, 특징 추출 및 비교, 위조서명 추정, 융합, 검증과정으로 이루어

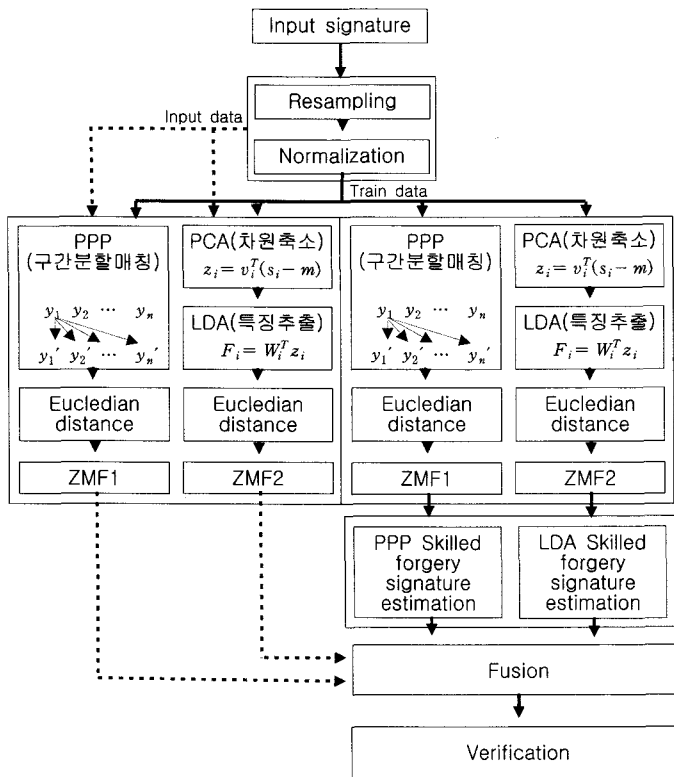


그림 4 제안된 온라인 서명 검증 시스템

진다. 각각의 처리부에 대하여 알아보면 다음과 같다.

우선, 데이터 취득부에서는 와콤(WACOM)사의 인투어스(intuos) 4×5사이즈의 태블릿을 이용하여 압력이 1 이상인 서명의 X축, Y축 및 시간(T)에 대한 온라인 서명 데이터를 획득하였다. 여기서 태블릿이나 전자펜으로부터 획득한 데이터는 동일인의 서명간에도 서명 당시의 감정 또는 환경변화에 의해 서명의 미세한 변이가 발생하고 그로 인해 취득한 서명데이터의 사이즈가 동일하지 않다. 그러나 선형판별분석기법에 의해 특징을 추출하기 위해서는 각 입력되는 데이터의 사이즈가 동일해야 하므로 본 논문은 리샘플링 과정을 거쳐 서명데이터의 사이즈를 동일화한다.

본 논문에서는 안정된 리샘플링 과정을 통하여 서명마다 사용할 서명 데이터의 사이즈를 동일하게 하였다. 즉, 이러한 리샘플링 과정은 취득한 데이터의 순서에 따라 데이터 시퀀스 번호를 부여한 후 시퀀스 번호를 균등 분할하여 데이터를 재 취득하였다. 리샘플링 과정의 첫 번째로 각 서명데이터를 동일한 조건으로 균등 분할하기 위해 데이터 시퀀스 번호의 시작점을 0, 끝점을 1로 정규화하였다. 다음으로 샘플링 데이터 개수에 따라 정규화된 시퀀스 번호를 균등 분할 후 분할영역내의 정규화된 시퀀스 번호들 중 중앙에 위치한 데이터를 선택한다. 그림 5는 서명 데이터 개수가 각각 6개와 10개인 데이터를 2개로 리샘플링하는 과정을 나타내고 있다.

그림 6은 리샘플링 전·후의 서명 데이터 결과를 나타냈다. 그림 6에서 'o' 표시는 취득 데이터의 분포를 나타낸다. 리샘플링 전의 서명데이터를 나타낸 그림 6(a)



그림 5 데이터의 시퀀스를 이용한 리샘플링 과정

의 경우 취득 데이터의 수는 123개에서 리샘플링 과정을 통하여 취득 데이터의 수를 80개로 재 취득하였으며, 그 결과를 그림 6(b)에 나타냈다. 그림에서와 같이 두 서명에 대하여 서로 유사한 위치를 갖는 데이터를 추출한 후에 각 서명간의 위치와 크기를 균일화시키기 위해 데이터를 0과 1사이의 값으로 정규화 한다.

서명의 특징추출은 2장에서 설명된 식 (11)에 의해 선형판별분석기법에 의해 특징벡터와 구간분할 매칭에 의해 선택된 봉우리 및 계곡점에 대한 시간 T의 데이터를 얻는 과정이다. 각각의 알고리즘으로부터 구한 여러 값은 베이지안 식별기에 적용시키기 위해 Z모양의 소속함수를 이용하여 낮은 에러값이 높은 소속도를 갖도록 0과 1사이값으로 변환한다. ZMF 소속함수를 설정하기 위하여 그림 7의 a 점과 b점을 설정해야 한다. 이를 위해 동일 서명내의 학습서명간의 유클리디언 거리를 측정 후 유클리디언 거리의 최소값과 최대값을 기준으로 ZMF 소속함수의 a, b 점을 설정하였다.

인식하고자 하는 입력서명에 대해서는 그림 8에서 보는 바와 같이 선형판별분석기법에 의해 입력 특징벡터를 산출한 후 미리 저장된 특징벡터와 유클리디언 거리를 산출한다. 또한 동시에 구간분할 매칭에 의한 유클리디언 거리 산출도 동일한 방법에 의하여 산출한다. 각각 산출된 유클리디언 에러값은 훈련단계에서 사용된 참조 서명의 특성을 이용하여 미리 설정된 ZMF1과 ZMF2로 표현되는 ZMF 소속함수를 통하여 입력서명의 소속도를 결정하였다.

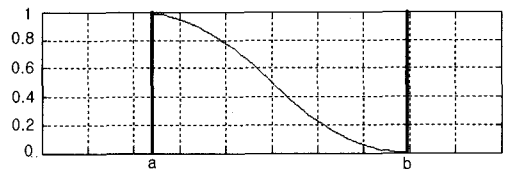
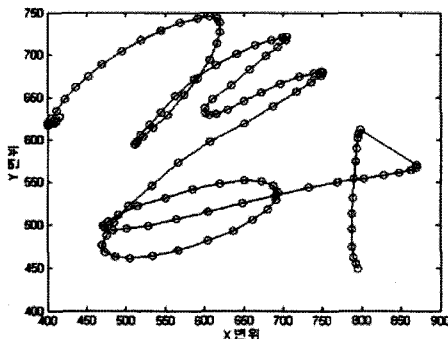
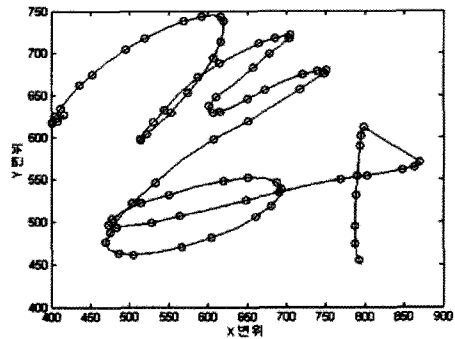


그림 7 Z-소속함수와 파라미터



(a) 리샘플링 전의 서명 데이터



(b) 리샘플링 후의 서명 데이터

그림 6 서명의 리샘플링 결과

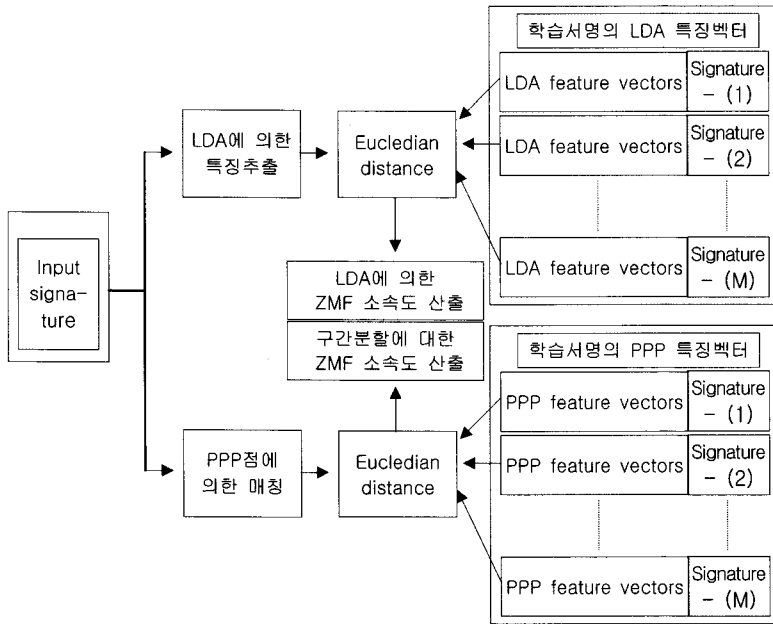


그림 8 학습서명과 입력서명간의 유클리디언 거리 측정

선형판별분석기법과 구간분할매칭에 의해 각각 산출된 소속도는 최종 인증 시 베이지안 식별기에 적용되며 식 (17)은 서명 판별을 위한 임계치로 베이지안 결정 규칙에 의해 산출하였다[13].

$$Thr = \frac{\prod_{i=1}^n f_i(x_i|w=1)}{\prod_{i=1}^n f_i(x_i|w=0)} \quad (17)$$

여기서, $f_i(x_i|w=0)$: 입력값에 대한 침입자의 가우시안 확률값

$f_i(x_i|w=1)$: 입력값에 대한 등록자의 가우시안 확률값

식 (17)로 표현되는 베이지안 식별기는 입력서명에 대한 등록자 서명의 분포와 침입자 분포를 이용하여 임계치를 결정한다. 이를 위해 등록자 서명과 침입자 서명의 분포도에 대한 확률적 모델이 선행되어야 한다. 일반적으로 다른 생체인식 보다 서명인식의 신뢰도가 낮은 원인은 서명을 획득 시마다 일어나는 서명의 변이와 전문위조가 용이함에 있다. 이 두 가지 문제점 중 전자보다 후자가 큰 비중으로 작용한다. 랜덤위조서명의 경우 참조서명으로부터 패턴을 예측하기 용이하나 전문위조서명은 위조 패턴이 다양하고 위조자에 대한 정보가 없어 예측하기 어렵다. 여기서 랜덤위조서명은 고의로 위조하지 않은 서명을 가리킨다. Richard 연구에 의하면 생체인식에 있어서, 등록자와 침입자의 매칭값의 분포는 그림 9와 같이 매칭값에 대해 일정 간격을 두고, 서로 떨어진 두 개의 가우시안 곡선을 형성한다[14-16]. 그림 9

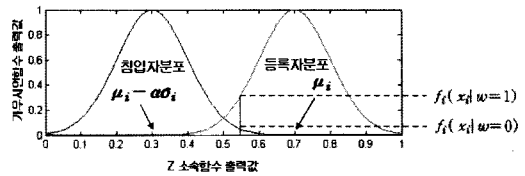


그림 9 등록자와 침입자의 가우시안 분포

에서 α 값은 일정 상수이다.

본 논문은 이 점을 고려하여 랜덤위조와 전문위조를 추정하기 위해 가우시안 함수를 사용하였고 두 위조서명을 하나의 침입자 분포로 추정하였다.

우선, 등록자에 대한 가우시안 분포는 사용된 학습서명에 대하여 Z-소속함수에 의해 산출된 소속도의 평균과 표준편차를 이용하여 가우시안 함수를 생성하였다. 위조서명에 대한 침입자 분포는 그림 9에서 보는 바와 같이 침입자 분포에 대한 평균값은 $\mu_i - \alpha\sigma_i$ (여기서, μ_i 와 σ_i 는 등록자분포의 평균값과 표준편차를 각각 나타냄)을 이용하여 산출하였다. 이와 같이 설정된 가우시안 모델에 대하여 각각 등록자 및 침입자에 대한 가우시안 확률값을 산출할 수 있다. 따라서, 입력서명 s_i 에 대한 최종 확률값 $Prob(s_i)$ 는 식 (18)에 의해 얻어진다.

$$Prob(s_i) = \frac{f_{LDA}(x_{LDA}|w=1)}{f_{LDA}(x_{LDA}|w=0)} \times \frac{f_{PPP}(x_{PPP}|w=1)}{f_{PPP}(x_{PPP}|w=0)} \quad (18)$$

여기서,

$f_{LDA}(x_{LDA}|w=1)$: 선형판별분석기법에 의한 산출된
입력 x_{LDA} 에 대한 등록자 가우시안 확률값

$f_{LDA}(x_{LDA}|w=0)$: 선형판별분석기법에 의한 산출된
입력 x_{LDA} 에 대한 침입자 가우시안 확률값

$f_{PPP}(x_{PPP}|w=1)$: 구간분할매칭기법에 의한 산출된 입력
 x_{PPP} 에 대한 등록자 가우시안 확률값

$f_{PPP}(x_{PPP}|w=0)$: 구간분할매칭기법에 의한 산출된 입력
 x_{PPP} 에 대한 침입자 가우시안 확률값

4. 실험결과

제안된 방법의 타당성을 검증하기 위하여 20명의 사
람으로부터 각각 10개씩 서명 데이터를 획득하였고 충
분한 학습과정을 통해 각 사람당 10개씩 위조서명을 취
득하였다. 동일인의 서명에 대한 10개의 진서명 중 5개
는 참조서명으로 사용하였고 나머지 5개의 진서명과 10
개의 전문위조서명을 대상으로 제안된 알고리즘의 성능
을 실험하였다. 그림 10에 사용된 본인서명과 위조서명
의 일부분을 나타냈다. 서명의 형태만을 이용할 경우 육

안으로도 진서명과 위조서명간을 분별하는데 한계가 있
음을 확인할 수 있다.

실험은 베이지안 분류기와 일반적으로 융합에 사용되
는 가중치 합에 의한 방법[16]과 비교하였다. 그림 11은
융합을 하지 않은 선형판별분석기법과 구간분할매칭방
법에 대한 검증 결과와 베이지안 기반 융합기법을 적용
한 경우의 인식률을 나타냈다. 서명인식시스템의 성능지
표로서 임계치의 변화에 대한 오거부율(False Reject
Rate, FRR)과 오인식률(False Accept Rate, FAR)을
사용하였다. 여기서, 오거부율이란 본인의 서명이 거부
된 오류율을 나타내고 오인식률은 타인의 서명이 본인
으로 인식된 오류율을 각각 의미한다. 그림 11(a)는 선
형판별기법만을 적용한 경우의 인식결과로서, 진서명에
대한 오거부율은 낮은 실선으로, 전문위조와 랜덤위조에
대한 오인식률은 각각 높은 점선과 가는 점선으로 나타
났다. 그림 11(b)는 구간 분할 매칭 방법만을 적용한 경
우, 그림 11(c)는 베이지안에 기반을 둔 융합기법을 적
용한 경우 인식결과를 나타냈다.

그림 11(a), (b)의 경우 임계치의 기준으로서 유클리
디언 에러를 사용한 반면에 융합한 그림 11(c)의 경우는

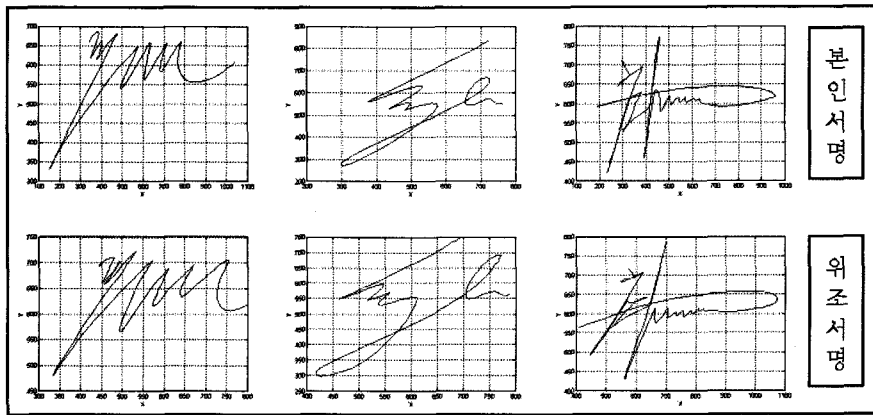
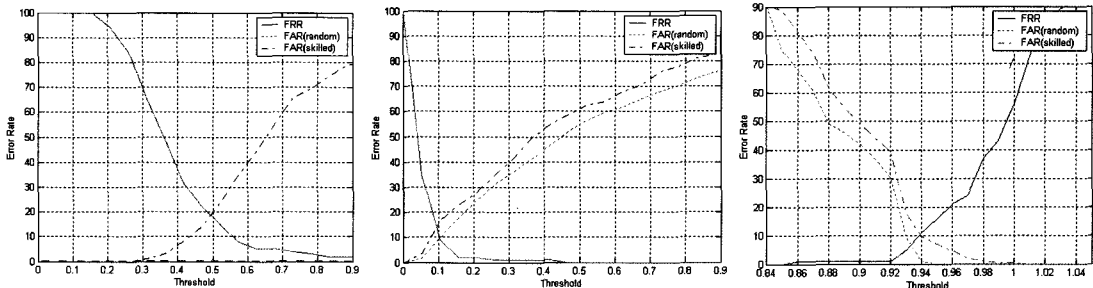


그림 10 진서명과 위조서명의 일례



(a) 선형판별분석기법을 적용한 경우 (b) 구간분할 매칭방법을 적용한 경우 (c) 베이지안 융합기법을 적용한 경우

그림 11 적용기법에 따른 오인식률과 오거부율

타인에 대한 본인의 확률값을 기준으로 설정하였다. 따라서, 그림 11(a), (b)는 어려움이 낮을수록 본인의 서명일 가능성이 높고, 그림 11(c)의 경우는 확률값이 높을수록 본인의 서명일 가능성이 높은 것을 의미한다. 그림 11의 (a), (b)로부터 선형판별분석기법에 의한 방법은 구간분할매칭방법보다 랜덤위조서명에 우수한 성능을 보인 반면에 구간분할매칭방법은 선형판별분석기법을 사용한 경우보다 전문위조서명에 우수한 성능을 보였다. 그림 11(c)의 베이지안 기반 융합기법을 적용한 경우 랜덤위조 및 전문위조에 대하여 향상된 인식결과를 나타냄을 확인할 수 있다. 그러나 그림 11의 경우 임계값을 모든 서명에 동일한 값을 적용한 실험결과로, 단일 인식 기법만을 적용한 경우보다 인식률이 향상되었지만 어려움이 다소 높게 나타나기 때문에 고도의 개인확인 시스템에 적용하는 데는 한계가 있음을 알 수 있다. 즉, 서명의 특성상 사람의 서명마다 크기 변이 및 분절의 수도 다양하기 때문에 임계값을 동일 적용하는 데는 문제점이 있다. 따라서, 개인특성에 맞는 최적의 임계값을 설정한 후 반복실험을 하였다.

표 1은 각 개인서명별로 최적의 임계값을 설정한 후 가중치합과 베이지안 분류기를 이용하여 융합한 결과를 나타냈다. 실험데이터는 참조서명과 입력서명의 데이터를 10회 랜덤하게 변경하여 반복실험을 하였다. 표 1의 결과로 제안한 알고리즘을 개별적으로 적용하였을 때 보다 융합을 하였을 때 우수한 성능을 나타냈으며 가중치합에 의한 방법이 랜덤위조에 대해 베이지안 분류기를 이용한 방법보다 높은 성능을 보였지만 오거부율과 오인식률을 전체적으로 평가해 보았을 때 보다 향상된 성능을 확인할 수 있다.

5. 결론

서명인식은 사용자에게 거부감이 없는 친숙한 생체인

식 기술 분야 중에 하나이다. 그러나, 서명인식의 경우 생체인식을 위한 다른 특징들과 달리 사칭자에 의해 전문 위조가 가능하고 동일인의 서명조차도 외부 환경 또는 내적인 요인에 의해 서명의 변이 차이로 시스템의 성능이 저하되는 문제점을 지니고 있다

따라서, 본 논문에서는 기존에 분절단위비교방법의 분절추출에 대한 불안정 문제를 선형판별기법과 융합함으로써 서명인식의 신뢰도를 향상시킬 수 있었다. 실험을 통해 제안된 방법의 타당성을 검증한 결과 선형판별분석기법에 의한 방법은 구간분할매칭방법보다 랜덤위조서명에 우수한 성능을 보인 반면에 구간분할매칭방법은 선형판별분석기법을 사용한 경우보다 전문위조서명에 우수한 성능을 보였다. 두 기법을 융합한 베이지안 기반 결정법칙을 적용한 결과 랜덤위조 및 전문위조에 대하여 향상된 인식결과를 나타냄을 확인할 수 있다. 또한 기존의 융합방법인 가중치합 결정법칙에 적용한 경우보다 제안된 베이지안 기반 결정법칙을 적용한 결과 더 나은 성능을 나타냈다.

참고 문헌

- [1] 전명근, 생체인식(Biometric) 총론, 한국정보통신교육원, pp. 8-1~8-21, 2002.
- [2] Kim, J. W, Cho, H. G., Cha, E, Y., "A study on enhanced dynamic signature verification for the embedded system," International Symposium on Brain, Vision and Artificial intelligence, pp. 19-21, 2005.
- [3] Mingming, M., Wilesona, W. S., "Automatic on-line signature verification based on multiple models," Proceedings of the Computational Intelligence for Financial Engineering, pp. 30-33, 2000.
- [4] Rhee, T. H., "On-line signature verification using model-guided segmentation and discriminative feature selection for skilled forgeries," Proceedings of

표 1 융합방법에 따른 인식률

[단위: %]

Signature Set	Weighted Sum Method			The proposed method		
	FRR	FAR		FRR	FAR	
		random	skilled		random	skilled
1	4	0	3.5	3	0.2632	2
2	6	0	4	2	0.7368	2
3	6	0	4	3	0.7368	1
4	4	0	3	3	0.0526	3
5	5	0	5	4	0.3684	2
6	4	0	1.5	4	0.3684	3.5
7	5	0	4	4	0.2632	2
8	5	0	4.5	3	0.1053	3.5
9	6	0	3.5	2	0.8947	2
10	4	0	4.5	3	1.8421	2
평균	4.9	0	3.75	3.1	0.5632	2.3

Document Analysis and Recognition, pp. 645-649, 2001.

- [5] Yue, K. W., and Wijesoma, W. S., "Improved segmentation and segment association for on-line signature verification," IEEE International Conference on Systems, Man and Cybernetics, Vol.4, pp. 2752-2756, 2000.
- [6] Brault, J. J. and Plamondon, R., "Segmenting handwritten signatures at their perceptually important points," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15, No.9, pp. 953-957, 1993.
- [7] Schmidt, C., and Kraiss, K. F., "Establishment of personalized templates for automatic signature verification," Proceedings of the Document Analysis and Recognition, Vol.1, pp. 263-267, 1997.
- [8] Kiran, G. V., Kunte, R. S. R., and Samuel, S., "On-line signature verification system using probabilistic feature modelling," Symposium on Signal Processing and its Applications, Vol.1, pp.355-358, 2001.
- [9] Ryu, S. Y., Lee D. J., Chun, M. G., "A robust on-line signature verification system," International Journal of Fuzzy Logic and Intelligent Systems, Vol.3, No.1, pp. 27-31, 2003.
- [10] Belhumeur, P. N., Hespanha, J. P., Kriegsmann, D. J., "Eigenfaces vs. Fisherfaces : recognition using class specific linear projection," IEEE Trans. on Pattern Analysis and Machine Intelligent., Vol.19, No.7, pp. 711-720, 1997.
- [11] 한학용, 패턴인식개론 : MATLAB 실습을 통한 입체적 학습, 한빛미디어, pp. 273-334, 2005.
- [12] Souheil, B. Y., Yousri, Y., Eddy, M., "Fusion of face and speech data for person identity verification," IEEE Transactions on Neural Networks, Vol.10, pp. 1065-1074, 1999.
- [13] Ross, A. A., Nandakumar, K., Jain, A. K., Handbook of Multibiometrics, pp. 99-111, Springer, 2006.
- [14] Richard, O. D., Peter, E. H., David, G. S., Pattern Classification, pp. 32-36, JOHN WILEY&SONS, INC. Second Edition, 2002.
- [15] 민준오, 류상연, 광근창, 전명근, "얼굴과 서명을 이용한 다중생체인식 시스템," BERC Biometrics Workshop, Vol.1, pp. 141-145, 2004.
- [16] Conrad, S., Kuldip K. P., "Structurally noise resistant classifier for multi-modal person verification," Pattern Recognition Letter, Vol.24, pp. 3089-3099, 2003.



이 대 중

1995년 충북대학교 전기공학과(학사). 1997년 충북대학교 전기공학과(공학석사). 2002년 충북대학교 전기공학과(공학박사). 2004년~2005년 University of Alberta Postdoc 2006년~현재 충북대학교 BK21 충북정보기술사업단 초빙 조교수. 관심분야는 Biometrics, 얼굴인식, 다중생체인식, 데이터마ining

고 현 주

정보과학회논문지 : 소프트웨어 및 응용 제 34 권 제 4 호 참조

전 명 근

정보과학회논문지 : 소프트웨어 및 응용 제 34 권 제 4 호 참조