

임베디드 생체인식 시스템에서 특징 추출1)

김병주⁰ 김일곤

영산대학교 정보통신학과 경북대학교 컴퓨터학과

bjkim@ysu.ac.kr ikkim@knu.ac.kr

Feature Extraction on Embedded Biometric Authentication System

Byung-Joo Kim⁰, Il-Kon Kim

Yongsan University Dept. of Information & Communication

Kyungpook National University Dept. of Computer Science

요 약

정보화 사회의 진행과 더불어 최근 스마트카드(smart card) 시스템을 비롯한 임베디드(embedded) 시스템의 사용이 활발해짐에 따라 위/변조나 도용에 강건한 인증 시스템의 필요성이 그 어느 때 보다도 높아지고 있다. 그러나 카드 내부의 메모리 크기 및 프로세스의 처리 능력은 매우 제한적이어서 일반 컴퓨터 환경에서의 인증 알고리즘이 수행되지 않을 수 있다. 따라서 적은 메모리와 제한적 처리 능력 하에서 동작 가능한 생체인증 알고리즘의 개발이 필요하다. 본 논문에서는 임베디드 생체인식 시스템을 위한 특징(feature) 추출을 위한 새로운 기법을 제안하였다. 제안된 기법은 다음과 같은 의미를 가진다. 첫째 비선형 자료의 특징 추출 성능에서는 제안된 방법이 기존의 Kernel PCA와 유사한 성능을 나타내었다. 둘째 기존의 비선형 추출 기법에 비해 메모리 사용면에서 효율적이다. 특히 제안된 방법은 학습 자료의 개수 N 이 클 경우에는 매우 유용하다.

1. 서 론

정보화 사회의 진행과 더불어 최근 스마트카드(smart card) 시스템을 비롯한 임베디드(embedded) 시스템의 사용이 활발해짐에 따라 위/변조나 도용에 강건한 인증 시스템의 필요성이 그 어느 때 보다도 높아지고 있다. 이를 해결하기 위하여 PIN(Personal Identification Number)보다 안전하고 쉽게 변하지 않으며 사용자마다 고유한 특징인 생체정보를 활용하는 기술이 최근 활발히 연구되어지고 있다. 생체 인식을 이용하여 사용자를 인증하기 위해서는 등록과정에서 저장된 원본 데이터와 인증을 목적으로 새로 제시한 샘플을 비교하는 과정을 수행하는데, 이와 같은 방식은 다음과 같은 문제점을 가지고 있다. 첫째 생체 정보를 중앙 집중식 데이터베이스에 저장하는 경우에는 데이터베이스의 유지 및 관리의 어려움이 있고 해킹 위험, 프라이버시 침해 등의 문제가 있다. 둘째 인증을 요구하는 생체 샘플 정보를 보안성이 결여된 채로 서버(server)에 전달하게 되어 정보가 누출될 위험성을 가지고 있다. 따라서 보안을 강화하기 위해 생체정보 샘플과 마스트 템플릿의 비교 과정이 외부 서버에서 처리되는 방식이 아닌, 카드 내부의 프로세스에서 수행될 필요가 있다. 그러나 카드 내부의 메모리 크

기 및 프로세스의 처리 능력은 매우 제한적이어서 일반 컴퓨터 환경에서의 인증 알고리즘이 수행되지 않을 수 있다. 따라서 적은 메모리와 프로세스의 제한된 처리 능력 하에서 동작 가능한 생체인증 알고리즘의 개발이 필요하다. 논문의 구성은 다음과 같다. 2장 본론에서는 특징 추출에 관한 기존 연구의 단점을 분석하고 이를 바탕으로 제안된 특징 추출 알고리즘에 대해 설명한다. 마지막으로 3장 결론에서는 제안된 방법에 의한 실험 결과 및 향후 연구 방향에 대해 논의한다.

2. 본 론

2.1 특징 추출에 관한 기존 연구

주성분분석(Principal Component Analysis:PCA) 방법은 학습 자료의 특징추출을 위해 사용하는 대표적인 기법중의 하나이다[1]. 하지만 PCA 방법은 일괄처리 방식으로 동작하므로 새로운 학습 자료가 추가 되면 고유공간(eigenspace)을 다시 계산 하여야 하는 단점이 있다. 또 다른 문제점은 PCA의 적용범위는 학습자료 간의 선형 상관관계가 있는 경우로만 국한된다. 이러한 문제점을 해결하기 위한 기존의 연구를 살펴보면 다음과 같다. 첫째 일괄처리 방식의 문제점을 해결하기 위한 방식은 크게 두 가지로 나눌 수 있다. 하나는 새로운 고유공간을 계산하는데 있어 평균의 갱신을 허용하지 않는 방법이며[2][3], 또 다른 접근 방법은 Hall[4]에 의해 제안된 방법으로 평균의 갱신을 허용하는 방법이다. Hall은 그의 논문에서 평균을 갱신하는 방법이 고정하는 방법에 비해 분류문제에서 우수한 성

1) 본 연구는 보건복지부 보건의료기술진흥사업의

지원에 의하여 이루어진 것임 (과제번호 :

A05-0909-A80405-05N1-00000A)

능을 나타냄을 실험을 통해 보였다. 두 번째 비선형 PCA에 관한 기존의 연구를 살펴보면 다음과 같다. Tipping[5]은 선형 주성분분석 기법을 결합한 MPCA(mixture PCA)방법으로 비선형 문제를 해결하고자 하였다. 이러한 해결책들은 모두 비선형 최적화 기법을 요구할 뿐 아니라 목적함수가 종종 지역 최소화에 빠질 가능성이 있다. 최근에 제안된 방법 중 Scholkopf[6]는 커널함수(kernel function)를 이용한 커널 주성분 분석(KPCA) 방법을 제안하였다. 이는 선형 주성분분석 방법에서와 유사한 방법으로 고유치(eigenvalue)문제를 해결함으로써 비선형 특징을 추출한다. 비록 KPCA 방법이 비선형 자료에 대해 특징 추출을 수행할 수 있지만 다음과 같은 단점이 있다. KPCA는 학습 자료의 개수가 N 개 일 때 $N \times N$ 크기의 커널 행렬을 저장해야 할 뿐만 아니라 커널행렬의 고유값을 구해야 한다. 만약 N 의 값이 큰 경우에는 고유값을 계산하기 위해 많은 양의 메모리가 필요하다. 이를 해결하는 몇몇 시도가 있었다. Rosipal[7]은 계산상에 있어 단순하고 효율적인 EM알고리즘을 KPCA에 적용하는 방법을 제안하였다. Moerland[8]은 EM 알고리즘을 KPCA에 적용하는 온라인 커널 주성분 분석 방법을 제안 하였다. Scholkopf[9]에 의해 제안된 축약된 집합 선택 기법은 KPCA의 메모리 문제를 어느 정도 해결 하였으나 많은 계산량이 요구 되는 단점이 있다. Smola[10]는 희소 커널 특징 분석을 제안하여 KPCA의 계산상의 문제점을 해결하였다. 하지만 이 방법은 커널 행렬을 저장해야 하는 문제점을 해결하지는 못했다.

2.2 제안된 특징 추출 방법

온라인 PCA 방법은 학습자료 x_{N+1} 이 추가 되었을 때 이전 학습 자료의 저장 없이 갱신하는 것이다. 먼저 새로운 학습 자료가 추가 되었을 때 갱신된 평균은 식 (1)과 같이 구한다.

$$\bar{x}' = \frac{1}{N+1} (N\bar{x} + x_{N+1}) \quad (1)$$

식 (1)에 의해 새로운 평균이 구해지면 추가된 학습 자료에 의해 갱신된 고유 벡터 집합 을 구할 수 있다. 이를 위해서는 이전의 고유벡터를 회전행렬에 적용하여야 하는데 이를 위해 먼저 직교간차벡터를 구해야 한다. 직교간차벡터는 식 (2) 와 같이 계산된다.

$$\hat{h} = (U_{nk}a_{N+1} + \bar{x}) - x_{N+1} \quad (2)$$

식 (2) 에서 구한 \hat{h} 을 정규화한 것은 식 (3)과 같이 표시된다.

$$h_{N+1} = \begin{cases} \frac{h_{N+1}}{\|h_{N+1}\|_2} & \text{for } \|h_{N+1}\|_2 \text{ and } h_{N+1} \neq 0 \\ \text{otherwise} & \end{cases} \quad (3)$$

여기서 $\|h_{N+1}\|_2$ 은 노름(norm)을 의미한다.

$q = k + 1$ 라 정의하면 새로운 고유벡터 U'_{nq} 은 식 (3)에 의해 구해진 h_{N+1} 과 이전의 고유벡터 U_{nk} 에 의

해 생성된 행렬을 회전행렬 $R_{(k+1)(k+1)}$ 에 적용하여 구할 수 있으며 식 (4)와 같이 구한다.

$$U'_{nq} = [U_{nk} \ h'_{N+1}] R_{(k+1)(k+1)} \quad (4)$$

여기서 $R_{(k+1)(k+1)}$ 은 회전행렬 이며 식 (5)의 고유공간간의 해이다.

$$D_{(k+1)(k+1)} R_{(k+1)(k+1)} = R_{(k+1)(k+1)} A'_{(k+1)(k+1)} \quad (5)$$

여기서 $A'_{(k+1)(k+1)}$ 은 새로운 고유치들의 대각 행렬이다. 행렬 $D_{(k+1)(k+1)}$ 는 다음과 같이 구성 할 수 있다.

$$D_{(k+1)(k+1)} = \frac{N}{N+1} \begin{bmatrix} \Lambda_{kk} & 0 \\ 0^T & 0 \end{bmatrix} + \frac{N}{(N+1)^2} \begin{bmatrix} aa^T & \gamma a \\ \gamma a^T & \gamma^2 \end{bmatrix} \quad (6)$$

여기서 $\gamma = h_{N+1}^T (x_{N+1} - \bar{x})$, $a = U^T (x_{N+1} - \bar{x})$ 와 같이 구하며 0는 k 차원의 영벡터(zero vector) 이다. 행렬 $D_{(k+1)(k+1)}$ 를 구성하는 몇 가지 방법이 제안되었는데 Hall이 제안한 방법만이 평균을 갱신할 수 있도록 제공하는데 이 기법은 평균의 갱신을 허용하지 않는 기법에 비해 성능이 우수한 것으로 알려져 있다 [4]. 따라서 제안된 방법에서는 Hall 이 제안한 평균의 갱신을 허용하는 방법을 이용한다. 2.2절에서 설명한 온라인 PCA 방식은 그 적용 범위가 학습 자료간의 선형적인 관계가 존재할 때 적용이 가능하다. 학습 자료간에 비선형성이 존재하는 경우에 대해 온라인 PCA를 적용하고자 할 경우 Cover[11]의 정리가 하나의 해결책이 될 수 있다. 그의 정리를 간단히 요약하면 “몇 가지 조건이 만족될 때 입력 공간에서 선형적으로 분리 불가능한 문제를 비선형 사상을 적용한 고차원 특징 공간(feature space)에서는 선형적으로 분리 가능한 문제로 변환될 확률이 높다” 이다. 비록 고차원 공간으로의 사상이 학습 자료간의 비선형 문제를 해결할 수는 있지만 고차원 공간으로의 사상으로 인한 계산 비용의 증가를 초래한다. 이러한 문제는 커널 기법(kernel trick)으로 해결이 가능하다. 커널 함수 $K(x_i, x_j)$ 는 다음과 같이 표현할 수 있다.

$$K(x_i, x_j) = (\phi(x_i) \cdot \phi(x_j)) \quad (7)$$

식 (7) 이 의미하는 것은 만약 특정 사상 ϕ 를 절히 선택하면 특징 공간에서의 내적(inner product)은 입력공간에서의 커널과 동일하다는 것으로 고차원의 특징 공간에서 계산할 필요가 없어 계산상의 증가를 줄일 수 있다는 것이다. 커널 함수는 Mercer[12]의 조건을 만족하는 함수들이 사용 가능하다고 알려져 있다. 커널함수 K 가 양반정치행렬(semi positive definite)이면 특정 사상 함수 ϕ 가 존재한다는 것을 Vapnik[12]은 증명하였다. 하지만 다항식 커널 함수를 제외한 나머지 커널 함수에서 특정 사상 함수 ϕ 를 구하는 것은 상당히 어려운 문제이다. 다항식 커널함수에서 특정 사상 함수 ϕ 를 구하는 과정은 다음

과 같다. 설명을 위해 $d=2$, $x = (x_1, x_2)$, $y = (y_1, y_2)$ 라 두면,

$$(x \cdot y)^2 = (x_1^2, \sqrt{2}x_1x_2, x_2^2)(y_1^2, \sqrt{2}y_1y_2, y_2^2)^T = (\phi(x) \cdot \phi(y)) \quad (8)$$

따라서 $\phi(x) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$ 이 된다. 다항식 특징 사상 함수의 경우 차원 d 의 값에 따라 성능은 큰 차이를 보이지 않는다[13]. 따라서 본 논문에서는 KPCA의 일괄처리 방식의 문제점은 Hall이 제안한 온라인 PCA기법을 적용하여 해결하고, 선형 자료에만 적용이 가능한 문제는 다항식 커널 함수에서 구해진 특징 사상함수를 온라인 PCA에 적용하여 비선형이면서 온라인 학습 자료에 대해서도 적용할 수 있는 온라인 비선형 주성분분석 알고리즘을 제안한다.

3. 실험

제안된 방법의 메모리 효율성 및 정확도를 검증하기 위해 대용량의 학습 자료에 대한 실험을 수행하였다. 학습 자료는 기계학습에서 학습 알고리즘의 성능을 평가하는 벤치마킹 자료로 많이 사용되는 바나나 학습 데이터 이다(<http://www.first.gmd.de/~raetsch>). 이때 사용하는 다항식 특징 사상 함수의 차원 d 는 2로 하였다. 메모리 효율 면에서 보면 바나나 학습 자료를 KPCA에 적용할 경우 고유벡터를 구하기 위해 5300×5300 크기의 커널 행렬을 저장하여야 한다. 하지만 제안된 방법에서는 커널 행렬을 저장할 필요 없이 4×4 크기의 회전행렬 R 과 4×4 크기의 D 행렬 정도만이 필요하다. 표 1 에 KPCA와 제안된 방법과의 메모리 요구량을 비교하였다. 표 1 에서 메모리 효율성이란 제안된 방법이 필요로 하는 메모리를 1로 설정하였을 때 KPCA를 수행하기 위해 필요한 메모리의 상대적인 값이다. 표에서 알 수 있듯이 학습 자료의 개수가 클수록 제안된 방법이 매우 유용한 기법임을 보여주고 있다.

4. 결론

본 논문에서는 임베디드 생체인식 시스템을 위한 특징(feature) 추출을 위한 새로운 기법을 제안하였다. 제안된 기법은 다음과 같은 의미를 가진다. 첫째 비선형 자료의 특징 추출 성능에서는 제안된 방법이 기존의 일괄처리 KPCA와 유사한 성능을 나타내었다. 둘째 기존의 비선형 추출 기법에 비해 메모리 사용면에서 효율적이다. 특히 제안된 방법은 학습 자료의 개수 N 이 클 경우에는 매우 유용하다. 향후 연구 과제는 제안된 특징 추출 방법에서 추출된 특징을 입력으로 사용하는 임베디드 생체인식 시스템에 적합한 인식기를 개발하는 것이다.

표 1. 바나나자료에 대한 KPCA와 제안된 방법의 메모리 비교

	KPCA	제안된방법
Kernel 행렬	$O(5300 \times 5300)$	필요없음
R 행렬	필요없음	$O(4 \times 4)$

D 행렬	필요없음	$O(4 \times 4)$
효율성	2,247,200	1

참고문헌

[1] I.T. Jolliffe, "Principal Component Analysis," New York Springer-Verlag, 1986.
 [2] H. Murakami, B.V.K. V Kumar., "Efficient calculation of primary images from a set of images," IEEE PAMI, 4(5), pp511-515, 1982.
 [3] J. Winkeler, B.S. Manjunath and S. Chandrasekaran., "Subset selection for active object recognition," In CVPR, volume 2, pp511-516, IEEE Computer Society Press, June 1999.
 [4] P. Hall, D. Marshall, and R. Martin., "Incremental eigenanalysis for classification," In British Machine Vision Conference, volume 1, pp 286-295, September 1998.
 [5] M.E. Tipping and C.M. Bishop., "Mixtures of probabilistic principal component analysers," Neural Computation 11(2), pp443-482, 1998.
 [6] B. Scholkopf, A. Smola and K.R. Muller., "Nonlinear component analysis as a kernel eigenvalue problem," Neural Computation 10(5), pp1299-1319, 1998.
 [7] R. Rosipal and M. Girolami., "An Expectation Maximization approach to nonlinear component analysis," Neural Computation, 13(3) pp.505-510, 2001.
 [10] P. Moerland, "An on-line EM algorithm applied to kernel PCA," IDIAP Research Report. 2000.
 [8] B. Scholkopf, S. Mika, C. Burges, P. Knirsch, K.R. Miller, G. Ratsch and A.J. Smola., "Input Space versus Feature Space in Kernel-based Methods," IEEE Transactions on Neural Networks, vol. 10, pp. 1000-1017, September 1999.
 [9] A.J. Smola, O.L. Mangasarian, and B. Scholkopf., "Sparse kernel feature analysis," Technical Report 99-03, University of Wisconsin, Data Mining Institute, Madison, 1999.
 [10] Cover, "Geometrical and statistical properties of system of linear inequalities with applications in pattern recognition," IEEE Transactions on Electronic Computers, Vol. EC-14, 326-334. 1965.
 [11] J. Mercer., "Functions of positive and negative type and their connection with the theory of integral equations," Philos. Trans. Roy. Soc. London, Vol. 209, pp.415-446, 1909.
 [12] V. N. Vapnik., Statistical learning theory. John Wiley & Sons, New York, 1998.
 [13] P. Moerland, "Mixture models for unsupervised and supervised learning," Ph.D. thesis, IDIAP research report IDIAP-PR 00-18, 2000.