특집

프라이버시 보호를 위한 생체정보의 변환

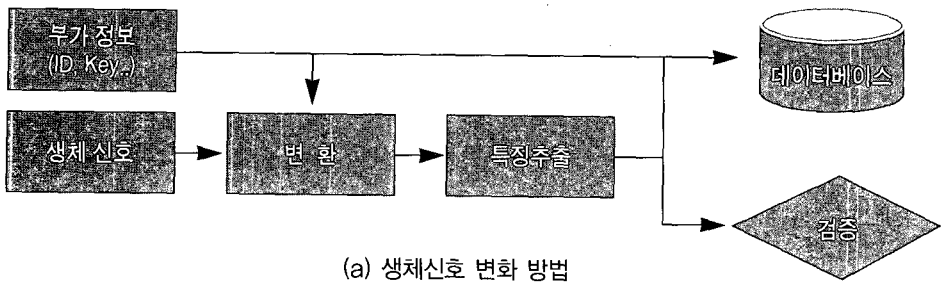
이철한, 최정윤, 김재희(연세대학교 전기전자공학부, 생체인식 연구센터),
박강령(상명대학교 미디어학부, 생체인식 연구센터)

I. 서 론

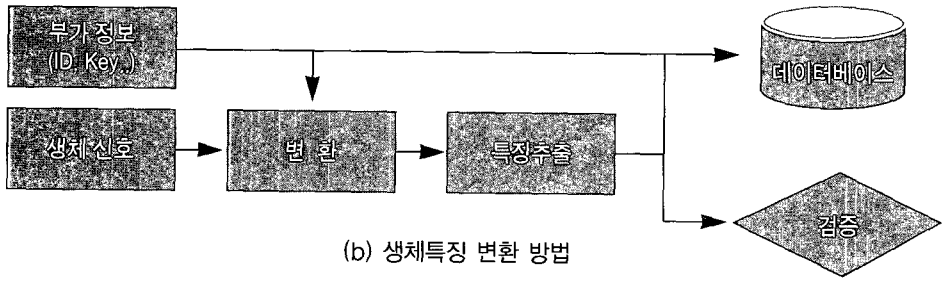
생체인식은 개인의 생체정보를 이용하여 개인을 인증하는 방법이다. 생체정보가 개인 인증의 수단으로써 사용될 수 있는 가장 큰 이유는 생체정보는 개인마다 다르다는 고유성(Uniqueness)과, 시간의 흐름에도 크게 변하지 않는다는 불변성(Permanence) 때문이다. 하지만 이러한 생체정보의 고유성과 불변성은 개인 생체정보의 도난이나 도용 시 심각한 프라이버시 침해의 문제를 야기시킬 수 있다. 전통적인 개인 인증 방법인 신분증이나 패스워드의 경우 도난이나 도용 시 새로운 신분증이나 패스워드를 재 발급하면 문제를 해결 할 수 있으나, 생체정보의 경우 새로운 생체정보를 재 생성하는 것은 불가능하다는 문제점이 있다. 또한 생체인식을 이용한 개인인증방법이 활성화 되면서 개인의 생체정보가 범죄수사와 같은 수사기관이나 은행 또는 기타 인터넷을 이용하는 다른 상업적인 기업체와 공유될 수 있고 이로 인해 개인의 생체정보가 도용될 수 있는 문제점이 있다. 이러한 문제점을 해결하기 위해 최근에 생체정보를 변환 후 변환된 생체정보를 이용해 개인을 인증하는 "Changeable (Cancelable) Biometrics"의 개념이 소개 되었다[1]. 이 개념은 특정 변환방법을 이용하여 원 생체정보를 새로운 생체정보로 변환하는 것으로, 변환방법과 변환된 생체정보를 알더라도 원 생체정보로의 복원은 불가능하게 함으로써 개인의 생체정보를 보호 할 수 있고, 또한 생체정보의 도난이나 도용 시에 변환방법을 변경하여 새로운 생체정보를 생성함으로써 개인의 프라이버시 침해를 막을 수 있는 방법이다.

II. 생체정보의 변환 방법

생체정보를 변환하여 새로운 생체정보를 생성하는 방법은 크게 생체신호를 변환하는 방법(Changeable biometrics in signal domain), 생체신호에서 특징을 추출 후 추출된 생체특징을 변환하는 방법(Changeable biometrics in feature domain)으로 나눌 수 있다[1]. 이들 방법들은 그림 1과 같이 정리될 수 있는데 생체정보의 변환은 입력 생체신호나 생체특징을 변환시키는 방법으로 변환된 생체의 인증을 위한 일관성 및 생체정보 보호를 위한 생체신호의 재 생성을 위해



(a) 생체신호 변화 방법



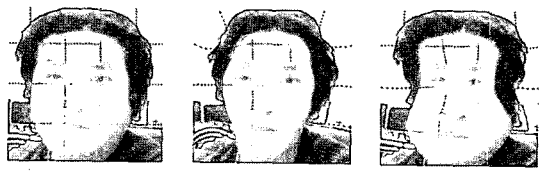
(b) 생체특징 변환 방법

〈그림 1〉 생체정보 변환 방법

ID나 Key와 같은 부과적인 정보를 이용하여 변환 방법이 결정된다.

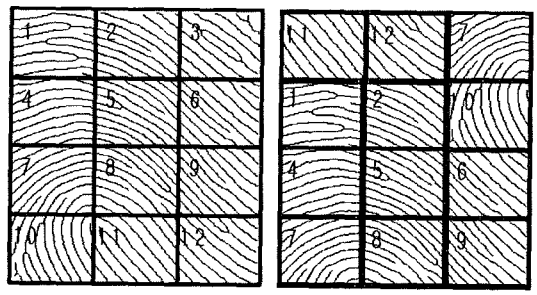
1. 생체신호 변환 방법(Changeable biometrics in signal domain)

프라이버시 보호를 위한 생체정보 변환 방법을 소개한 Ratha[1, 2]는 생체신호 변환 방법(Changeable biometrics in signal domain)으로 모핑(Morphing)을 이용한 방법과 블록 변환(Block Permutation)방법을 예로 들었다. 모핑을 이용한 방법은 입력 영상을 특정 모핑 함수로 새로운 생체신호를 생성하는 방법이다(그림 2). 변환된 생체정보가 도난이나 도용 시에는 새로운 모핑 함수를 이용하여 또 다른 변환된 생체정보를 획득 할 수 있다.



(a) 원영상 (b) 변환 영상 (c) 변환 영상

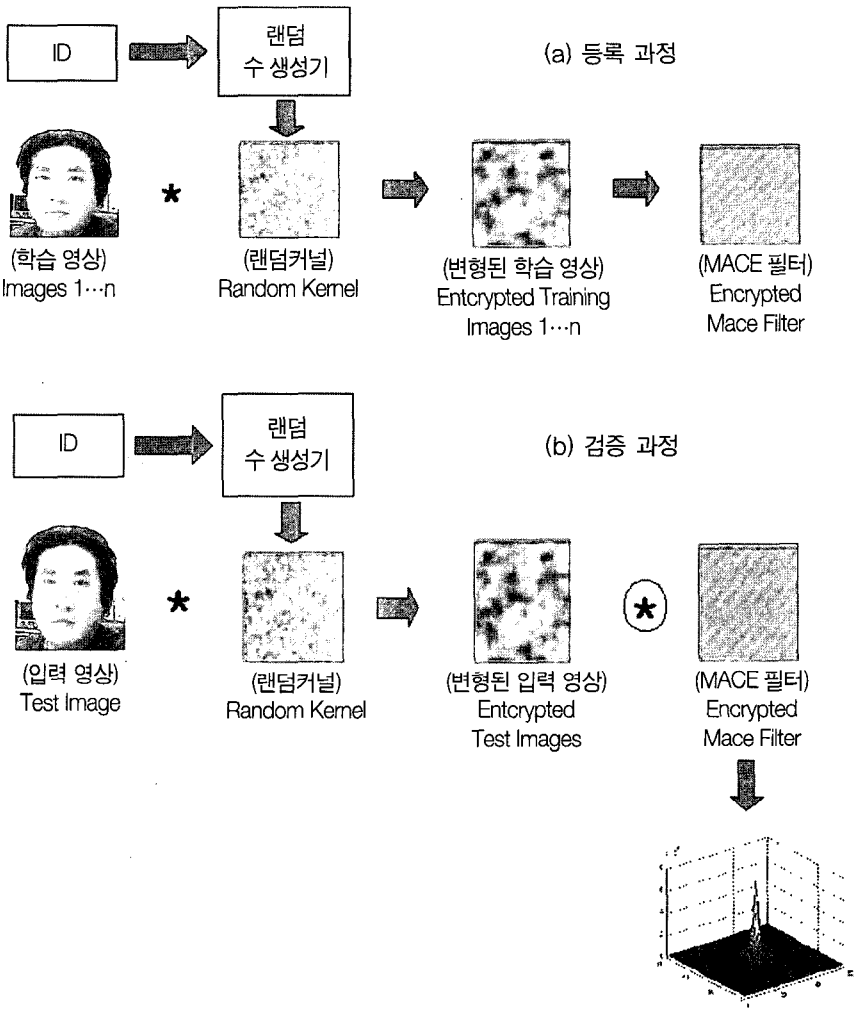
〈그림 2〉 모핑(Morphing)을 이용한 생체신호 변환 방법



(a) 원영상 (b) 변환 영상

〈그림 3〉 블록 변환(Block Permutation)을 이용한 생체신호 변환 방법

블록 변환(Block Permutation)을 이용한 방법은 입력 영상을 블록으로 나눈 후 그 순서를 그



〈그림 4〉 MACE 필터를 이용한 생체신호 변환

림 3와 같이 섞는 방법이다. 이 방법은 단순하게 블록의 순서를 섞는 것뿐 아니라 각 블록을 회전시키는 것까지 포함된다. 이러한 모핑 방법과 블록 변환 방법을 이용할 경우에는 동일인의 생체신호에 대해 동일한 변환을 적용하여 동일하게 변환된 생체신호를 얻을 수 있어야 하며, 이를 위해서는 변환 전에 생체신호의 정렬(Alignment)이 반드시 요구된다. [2]에서 생체신호의 정렬 방법으로 얼굴의 경우 눈이나 코의 위치를 이용한 방법과 지문의 경우 특이점(Singular Point)인 중심점(Core)과 삼각주(Delta)를 이용하는 방법, 홍채의 경우 눈의 양끝 점을 이용하는 방법에 대해 언급하였으나 지문의 경우 많은 영상이 중심점과 삼각주를 동시에 추출 못하여 정렬이 불가능하고, 또한 이러한 경우 정렬의 오차가 성능에 크게 영향을 미칠 것으로 예상된다. 그리고 블록 변환의 경우 악의적인 공격자가 섞는(Scrambling) 방법만 안다면 원 생체신호로의 복원이 가능하다는 심각한 문제점이 있다.

생체신호를 변환하는 또 다른 방법으로 그림 4과 같이 [3]에서 제안한 방법이 있다.

이 방법은 MACE(Minimum Average Correlation Energy) 필터를 이용하여 얼굴인식을 하는 방법에서 적용하는 것으로 랜덤 수 생성기(Random Number Generator)에서 생성된 랜덤 커널(Random Kernel)로 영상들을 변형 후 MACE 필터를 생성하여 인증하는 방법이다. 자신의 생체정보가 도난시에는 새로운 랜덤 커널을 생성하여 생체정보를 재 생성 시킬 수 있다. 그림 4는 n개의 학습영상으로 개인의 MACE필터를 생성하는 등록 과정과 인증과정을 보여준다. 등록 과정에서 n개의 학습영상은 랜덤커널과 콘볼루션(convolution)후 변형된 학습영상을 생성하고, 이 영상들에서 MACE필터를 생성한다. 인증과정에서 등록 시 사용한 램덤커널과 같은 커널을 사용해 입력된 영상에서 변형된 영상을 생성 후 그 영상과 등록과정에서 생성된 MACE필터와의 콘볼루션으로 인증을 실시한다. 이렇게 랜덤커널이 적용되는 방법은 랜덤커널을 적용하지 않는 방법과 동일한 인증성능을 나타냄을 이론적으로, 실험적으로 보였다. 그러나 이 방법의 단점은 랜덤커널을 통해 입력 얼굴영상이 완전히 다른 영상(얼굴영상 아닌 것)으로 변화되므로 MACE필터를 이용하는 얼굴인식 시스템에서만 적용이 가능하다는 것이다.

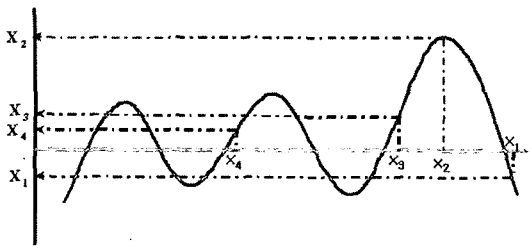
2. 생체특징 변환 방법 (Changeable Biometrics in feature domain)

생체인식 시스템의 특징추출(Feature Extraction) 과정에서 추출된 특징을 변형하여 새로운 변환된 특징을 생성하는 방법으로, [1]에서는 지문의 특징(Minutiae)을 이용한 방법에 대해 예를 들었다. 지문의 특징은 $S=\{(x_i,y_i,\theta_i), i=1,.....,M\}$로 구성된다. 여기서 x_i, y_i는 특징점의 x, y위치이고 θ_i는 특징점의 각도 정보를 나타낸다. 특징점 S는 새로운 특징점 $S'=\{(X_i, Y_i,\Theta_i\}, \ i=1,.....,M\}$로 변형 되는데 이때 변형을 위한 함수로 아래 식 1과 같은 고차 다항식(High order polynomial) $X=F(x)$, $Y=G(y)$, $\Theta=H(\theta)$를 이용하는 방법이다.

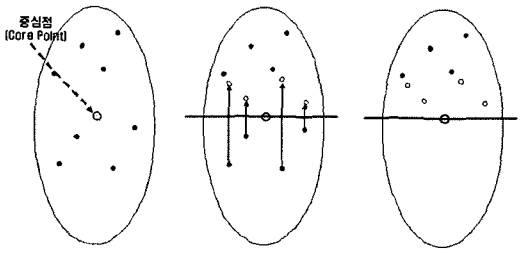
$$X=F(x)=\sum_{n=0}^{N}\alpha_n x^n=\prod_{n=0}^{N}(x-\beta_n) \qquad (식\ 1)$$

그림 5는 이러한 고차 다항식을 이용한 생체특징 변환 방법에 대해 보여준다. (가로축)에서 (세로축)로의 변환은 일대일(one-to-one) 변환이며 X에서 x로의 변환은 일대다(one-to-many)가 되어 변환 함수와 변환된 특징 값(X)을 알더라도 원래 특징점(x)으로 복원이 불가능하게 된다. 이 방법은 적용면에서 매우 간단하지만 입력 특징점의 정확한 정렬(Alignment)이 반드시 요구된다.

D. Ngo[4]는 특징벡터(Feature Vector)를 랜덤벡터와 내적 후 내적 값을 이진화 시켜 복원이 불가능하게 하는 PalmHashing 방법을 제안 하였다. 입력된 손바닥 영상(Palmprint image)에서 ROI(Region of interest)를 추출 후 FDA(Fisher Discriminant Analysis)를 이용하여 특징벡터를 계산하고 계산된 특징 벡터를 랜덤 패턴과 내적을 시킨다. 여기서 랜덤 패턴은 랜덤 생성기로 랜덤 행렬을 생성 후 Gram-Schmidt 방법을 적용하여 직교행렬(Orthonormal Metrics)로 만든 행렬이다. 이러한 랜덤패턴과의 내적 후 특정 값(Threshold)보다 크면 1 작으면 0으로 이진화 시켜 복원이 불가능한 장문(Palm) 코드를 생성하게 된다. 여기서 새로운 장문 코드



〈그림 5〉 다차원 함수를 이용한 생체특징 변환 방법



〈그림 6〉 지문 특징점 위치 변환

는 새로운 랜덤패턴을 생성 함으로써 생성될 수 있다.

R. Ang[5]는 중심점을 기준으로 지문 특징점 위치를 변형시켜 변환된 지문 특징점을 생성하는 방법을 제안하였다. 이 방법은 입력된 지문에서 중심점을 찾은 후 중심점(Core)을 지나는 각도가 K(key)인 하나의 직선을 가상적으로 생성하고, 이 가상의 직선보다 아래의 특징점들을 가상의 직선위로 이동(reflection) 시켜 새로운 특징점을 생성하는 방법이다. (그림 6)[5]에서는 $\pi/10$ 간격으로 중심점을 지나는 직선으로부터 변환된 특징점을 생성하였다. 실험 결과에 따르면 변환전의 성능은 EER 4%가 나오는 반면 변환 후 성능은 EER 16.8%로 성능이 저하 되었고, 약 70% 지문의 특징점이 변환 후에도 변환전의 특징점과 일치한다는 결과를 보였다. 이 방법의 단점은 중심점을 찾아야 한다는 문제점과 생성될 수 있는 변환된 생체특징의 수에 제한이 있다는 것이며, 또한 실험 결과에서 보여주듯이 변환이 뚜렷하게 발생하지 못하는 문제점이 있다.

III. 생체정보 변환 방법의 평가 기준

생체정보의 변환 방법에 대해 현재까지 몇 가지 방법이 제안되었지만, 생체정보 변환 방법에 대한 명확한 성능 평가 기준은 정립되어 있지 않았다. 생체정보 변환 방법이 프라이버시 보호를 위한 방법으로 사용되기 위해서는 아래와 같은 5가지 사항을 고려해야 생체 정보 변환 방법의 성능을 평가해야 한다.

- ▷ 비 가역성(Non-Invertability): 변환 후 원 생체정보의 복원 가능성 평가
- ▷ 분리 가능성(Separability): 변환 후와 변환전의 생체정보를 이용한 인증성능 비교 평가
- ▷ 임의성 (Randomness): 변환 후와 변환전의 생체정보량의 변화 평가
- ▷ 계산 복잡도(Computational Complexity): 생체정보 변환에 필요한 연산량 측정
- ▷ 매개변수화(Parameterized Formulation): 얼마나 많은 변환된 생체정보를 재생산할 수 있는가에 대한 평가

비 가역성은 변환방법과 변환된 생체정보를 알더라도 원 생체정보의 복원 가능성에 대한 평가 지표로 변환함수의 비 가역성이나 특징 공간의 크기 변화로써 판단할 수 있다. 분리 가능성은 생체정보 변경 후와 변경 전의 성능에 대한 평가로 자기자신의 정합결과의 분포(Genuine distribution)와 다른 사람과의 정합결과의 분포(Imposter distribution)의 변화로 평가 할 수 있

다. 임의성은 변환 후의 생체정보가 가질 수 있는 정보량을 계산하여 변환 전과 변화 후의 정보량의 변화를 평가하는 지표로 생체특징의 엔트로피(Entropy)로써 평가할 수 있으며, 계산복잡도는 생체정보를 변화하는데 필요한 연산량을 평가하는 지표로 연산량이 적을 수로 좋은 변환 방법이라 할 수 있다. 매개변수화는 생체정보의 변환에 사용된 함수가 매개변수화가 가능하다면 매개 변수를 변화시켜가며 다수의 변환된 생체정보를 재 생성할 수 있으므로, 변환 생체정보 수에 대한 지표가 된다.

Ⅳ. 결론 및 향후 과제

생체정보의 변환(Changeable Biometrics)은 생체정보의 도난이나 도용 시 개인의 프라이버시를 보호하기 위해 도입된 개념이다. 생체정보의 변환 방법으로는 입력 생체신호를 변환하는 방법(Changeable Biometrics in Signal Domain)과 생체인식 시스템에서 추출된 생체 특징을 변환하는 방법(Changeable Biometrics in Feature Domain)으로 크게 나눌 수 있다. 생체신호를 변환하는 방법에는 모핑, 블록 변환, 랜덤커널을 이용하는 방법이 있으며 생체특징을 변환하는 방법으로는 고차원 함수를 이용하여 복원이 불가능하게 변환하는 방법, 생체특징을 기하학적으로 변환하는 방법, 랜덤 생성함수에서 생성된 랜덤 행렬을 통해 변환하는 방법 있다. 이러한 생체정보의 변환 방법은 개인의 프라이버시 보호측면에서 그 중요성이 계속 언급되고 있으나 아직까지 실제 사용될 수 있는 명확한 방법과 성능을 제시한 것은 없다. 현재까지 제안된 방법의 문제점으로는 첫째, 생체정보는 입력되는 상태에 따라 변화가 발생되는데(회전, 이동, 크기 변화 등) 이러한 변화에 강인하지 못하다는 것과, 둘째, 변환 후 새로운 생체정보를 생성 시 기존의 생체정보와 형태가 달라 기존의 인식 알고리즘을 적용 못하는 경우가 있다는 것이다. 생체입력 상태의 변화에 따른 문제점을 해결하기 위해 생체정보를 변환 전에 생체정보를 정렬 시키는 방법이 제안 되었으나 이 방법은 현실적으로 적용하는데 문제점이 있다. 생체정보의 변환 방법이 실제 생체 인식 시스템에 사용되려면 최우선적으로 이러한 생체신호의 입력 변화에 강인한 변환 방법이 개발 되어야 한다. 그리고 현재의 생체정보 변환 방법은 검증(Verification)에 대해서만 고려되고 있는데 인증(Identification)에서도 사용될 수 있는 방법도 연구 되어야 한다.

<Acknowledgements>
본 연구는 한국과학재단 지정 생체인식 연구센터(BERC)의 지원을 받아 이루어 졌습니다.

참고 문헌

[1] N. K. Ratha, J. H. Connell, and R. M. Bolle, "Enhancing security and privacy in biometrics-based authentication systems," IBM Systems Journal, Vol. 40, No. 3, 2001.

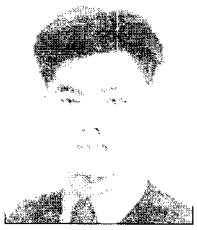
[2] R. M. Bolle, J. H. Connell, and N. K. Ratha SYSTEM AND METHOD FOR DISTORTING A BIOMETRIC FOR TRANSACTIONS WITH ENHANCED SECURITY AND PRIVACY, US Patent 6,836,554 B1.

[3] Savvides, M.; Vijaya Kumar, B.V.K.; Khosla, P.K., "Cancelable biometric filters for face recognition", ICPR 2004. Proceedings of the 17th International Conference on Vol 3, Page 992-925, Aug. 2004.

[4] Tee Connie, Andrew Teoh Beng Jin, Michael Goh Kah Ong, David Ngo Chek Ling, "PalmHashing: a novel approach for cancelable biometrics." Inf. Process. Lett. 93(1): 1-5, 2005.

[5] Russell Ang, Rei Safavi-Naini, and Luke McAven, "Cancelable Key-Based Fingerprint Templates", Information Security and Privacy: 10th Australasian Conference, ACISP 2005.

저자소개



이 철 한

2000년 명지대학교 전자공학과 학사
2002년 연세대학교 전기전자공학과 석사
현 재 연세대학교 전기전자공학과 박사 과정
주관심분야 생체인식, 패턴인식, 컴퓨터 비전

최 정 윤

1992년 2월 연세대학교 전기전자공학과 졸업
1992년 연세대학교 전자공학과 학사
1994년 연세대학교 본대학원 전자공학과 석사
1999년 메사추세츠공대 전기컴퓨터과 박사
1999년-2001년 메사추세츠공대 전기컴퓨터과 박사 후 과정
2001년-2004년 일리노이주립대 전기컴퓨터과 박사 후 과정
2005년-현 재 연세대학교 생체인식연구센터 연구교수
주관심분야 신호처리, 음성인식, 생체인식, 인지과학.

저자소개

김 재 희

1979년 연세대학교 전자공학과 졸업
1982년 미국 Case Western Reserve University 전기 공학 석사
1984년 미국 Case Western Reserve University 전기 공학 박사
현 재 연세대학교 전기전자공학부 교수
한국생체인식포럼 기술분과 위원장
(과학기술부 지정) 생체인식 연구센터 소장
주관심분야 생체 인식, 패턴 인식, 컴퓨터 비전

박 강 령

1994년 연세대학교 전자공학과 졸업
1996년 연세대학교 전자공학과 석사
2000년 연세대학교 전기 · 컴퓨터공학과 박사
2000년-2003년 LG전자기술원 홍채인식팀 선임연구원
2003년-현 재 상명대학교 미디어학부 조교수
2003년-현 재 생체인식연구센터 제2총괄과제 책임자
주관심분야 영상처리, 생체인식, 컴퓨터비전