Journal of Korean Institute of Intelligent Systems, Vol. 23, No. 2, April 2013, pp. 113-119
http://dx.doi.org/10.5391/JKIIS.2013.23.2.113

얼굴인식 템플릿 보호를 위한 Real Fuzzy Vault

Real Fuzzy Vault for Protecting Face Template

이대종* · 송창규* · 박성무** · 전명근*†
Dae-Jong Lee* · Chang-Kyu Song* · Sung-Moo Park** · Myung-Geun Chun*†

*충북대학교 전자공학부, **한국폴리텍대학 홍성 캠퍼스 전기과
* School of Electronics Engineering, Chugbuk National University
Department of Electricity, Hong Seong Campus, Korea Polytechnic College

요 약

얼굴인식 시스템은 사용자 인터페이스의 편리함과 용이한 구현성에 기반하여 범죄 수사를 포함한 다양한 분야에서 널리 사용되고 있다. 그러나 얼굴인식정보가 불법 사용자에게 누설되었을 때 개인의 프라이버시를 침해 할 수 있는 문제점을 지니고 있다. 본 논문에서는 실수형 오류정보 부호 코드화를 이용하여 얼굴인식 정보를 보호하기 위한 Real fuzzy vault 방법을 제안한다. 제안된 방법은 분실 시 재생성할 수 없는 얼굴영상 정보와 달리 개인 키값을 수시로 변경할 수 있으므로 사용자의 프라이버시를 보호할 수 있는 장점이 있다. 제안된 방법을 실제 얼굴인식에 응용하기 위하여 구현하여 실험함으로서 논문에서 제안된 방법의 타당성과 유용성을 보였다.

키워드 : 퍼지 볼트, RN-ECC, 얼굴인식, 바이오인식

Abstract

Face recognition techniques have been widely used for various areas including criminal identification due to their capability of easy implementing and user friendly interface. However, they have some drawbacks related to individual's privacy in case that his or her face information is divulged to illegal users. So, this paper proposed a novel method for protecting face template based on the real fuzzy vault. This proposed method has some advantages of regenerating a new face template when a registered face template is disclosed. Through implementing and testing the proposed method, we showed its validity and usefulness.

Key Word : Fuzzy Vault, RN-ECC, Face Verification, Biometrics

1. 서 론

바이오인식 기술은 개인을 인증하는 긍정적 역할을 함으로서 신분확인과 관련된 다양한 불법행위를 적발하거나 예방할 수 있다. 그러나 바이오인식 기술은 본인의 동의 없이 개인의 신상정보나 그와 관련된 모든 거래나 데이터베이스상의 변동 자료를 추적할 수 있으며, 다양한 개인적인 정보를 특정 개인과 관련하여 지속적으로 축적할 수 있다. 이러한 문제점을 해결하기 위하여 바이오인식 템플릿을 보호하기 위한 기법들이 최근 들어 국내외적으로 활발하게 이루어지고 있다[1-2]. 이들 기법들은 특징벡터 변환(feature transformation)과 바이오 암호시스템(Biometric Cryptosystem)으로 나누어진다. 특징벡터 변환 방법은 역변환 가능한 변환 함수를 쓰는 경우와 역변환이 가능하지 않은 방법을 사용하는지에 따라서 다시, BioHashing과 Robust Hashing 방법으로 나누어진다. 한편, 바이오 암호시스템(Biometric Cryptosystem)에 있어서는 암호화 키를 직접 바이오정보로부터 만들어 내는 방법(Key Generation)과 암호화 키를 바이오정보와 엮어서 보관한 후, 이를 필요한 경우에 바이오정보를 이용하여 다시 추출해 낼 수 있도록 하는 Key Binding 방법이 있다[3].

특징벡터 변환 방법은 사용자가 지정한 키 값이나 패스워드로부터 정의되는 함수에 기초하게 된다. 특히, 역변환 가능한 경우에 있어서는 사용되는 키 값을 안전하게 보호해야 되는 부가적인 요구사항이 생기는 반면에, 공격자에게는 바이오인식 시스템을 해킹하기 위해서는 키 값과 함께 바이오정보도 요구하게 되어 시스템의 안전도는 높아질 수 있는 구조이다. 현재, 가장 많이 연구 되고 있는 형태는 랜덤한 직교좌표에 기반한 BioHashing 방법이 있다[4]. 입력된 특징 벡터는 사전에 랜덤하게 생성되어 토큰에 저장되어 있는 직교행렬들과 내적을 통하여 특정한 값을 산출하게 되고, 이를 지정된 임계값을 이용하여 이진화 하여 원하는 BioHash 값을 갖는 구조로 이루어 졌다. 실험 결과 낮은 오수락율(False Accept

접수일자: 2013년 1월 18일
심사(수정)일자: 2013년 2월 19일
게재확정일자 : 2013년 3월 4일
† Corresponding author
이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임 (2010-0024037)

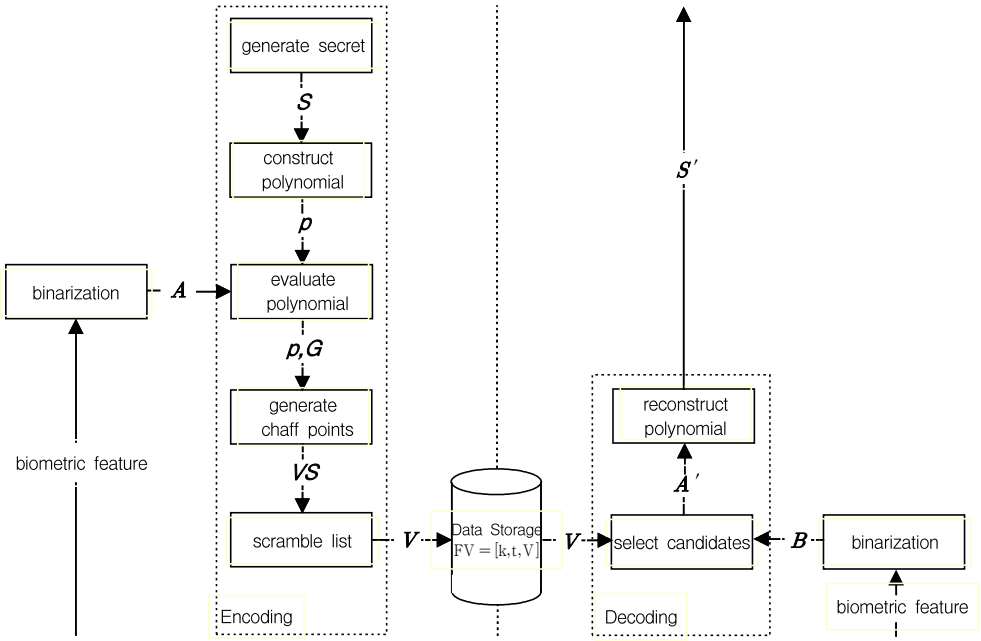


그림 1. 퍼지볼트를 이용한 바이오인식
Fig. 1. Biometrics by fuzzy vault scheme

Rate) 등의 장점이 있으나, 다음과 같은 단점과 한계가 지적되고 있다.

첫 번째로, 변환함수에 사용된 사용자 지정 키값이 유출되었을 때, 역변환 가능한 함수일 경우에 원래의 원본 바이오 템플릿 정보를 얻을 수 있다는 점이다. 두 번째로, 매칭이 변환된 영역에서 이루어지므로, 인식률의 저하가 없기 위해서는, 변환 전 템플릿간의 유사도가 변환 후에도 유지 될 수 있어야 한다. 따라서 변환 함수의 선택에 따라 인식 성능이 크게 변할 수 있다는 점이다. 마지막 문제점으로 인식률을 높이기 위하여 크기가 큰 랜덤 행렬을 사용하는 경우, 바이오 정보에 의한 변별력 보다, 행렬 자체에 의한 변별력이 커져서 바이오 인식 알고리즘을 사용하는 장점자체가 없어지고, 따라서 타인 수락율이 커질 수 있는 위험이 있다는 점이다.

본 논문에서는 바이오정보에 대한 프라이버시 및 보안을 보장하기 위하여, 실수형 오류정정 부호화 코드(Real number error correcting code)를 이용한 실수형 퍼지볼트에 기반한 얼굴 템플릿의 보호기법을 제안한다. 기존에 단순한 유한체 다항식에 기초하여, 다항식에 계수에 키를 Bind하여 이를 바이오 정보를 이용하여 추출하는 경우, 바이오 정보의 변동을 극복하기 위하여 에러 수정 코드나 CRC(Cyclic redundancy check)를 이용하였다. 그러나 이럴 경우 구현이 복잡할 뿐더러, 오차 수정 코드로 매칭을 하는 경우에 바이오 인식 정보의 매칭을 위해 개발 된 실수형의 매칭기를 사용할 수 없는 단점이 있다. 따라서 이를 극복하고자 실수형 다항식의 근사화 특성을 가지는 RN(Real number) ECC(오류정정코드)가 제안되고 지문인식에 적용되었다[5]. 본 논문에서는 이러한 실수형 오류정정기법을 얼굴인식에 적용하여 실험함으로서 그 결과의 타당함을 보이고자 한다.

2. 퍼지 볼트를 이용한 바이오정보 보호기법

퍼지 볼트는 중요한 정보를 보호하기 위해 Jules와 Sudan에 의해 제안된 방법으로 인코딩(Encoding)과 디코딩(Decoding)과정으로 이루어진다[6]. 먼저 인코딩 과정을 살펴보자. Alice는 집합 A를 이용하여 비밀정보 S를 은닉하려고 한다. 그러면 비밀정보 S를 이용하여 변수 x에 대한 다항식 $P(x)$를 생성한다. 그리고 집합 A의 원소들을 다항식의 변수 x에 각각 대입하여 다항식 $P(x)$위에 존재하는 점들을 생성한다. 그리고 나서 공격자로부터 다항식 $P(x)$를 은닉하기 위해서 $P(x)$와 무관한 임의의 점들(chaff points)을 다항식 위의 점들에 삽입하여 볼트 V(Vault)를 생성함으로 인코딩 과정을 마친다.

이번에는 Bob이 집합 B를 이용하여 금고에 있는 S를 얻고자 한다. 만약 B의 원소들과 A의 원소들 사이에 일치하는 부분이 많다면, B는 V에서 다항식 $P(x)$위에 존재하는 대부분의 점들을 식별해내게 된다. 그러나 집합 B를 이용해 찾아낸 점들에는 다항식과 무관한 점들도 포함되어 있을 수 있다. 이러한 오류 점들은 오류정정부호(Error correcting code)를 통해 정정되고 Bob 은 정확하게 S를 얻어내게 된다. 반면에 B와 A가 거의 일치하지 않는다면 $P(x)$를 알아낼 수 없도록 방해하는 많은 점들(chaff points) 때문에 Bob 은 S에 접근할 수 없다.

바이오인식의 특징 중 하나는 등록과정에서 추출된 바이오특징 정보와 인증을 위해서 새롭게 취득하는 바이오 인식 정보는 비록 본인일 지라도 항상 동일하지 않다는 점이다. 얼굴인식의 경우는 표정이나 조명과 같은 변수에 있어서 항상 동일한 특징값이 없어지기가 어렵다. 따라서, 위에서 언급된 개념 중에 fuzzy(애매한) 라는 개념이 도입된 퍼지 볼트는 바이오인식 시스템과 잘 결합될 수 있다. 따라서 퍼지 볼트를 이용한 지문[7], 얼굴[8][9], 홍채[10] 등의 연구가 활발하게 이루어지고 있다.

그림 1은 전형적인 퍼지볼트 방법과 바이오 인식 시스템을 결합한 바이오인식 시스템을 나타냈다. 퍼지 볼트를 만드는 인코딩과정을 간략히 설명하며 다음과 같다. 사용자의 식별

자로 쓰일 K-비트로 구성된 PI (Pseudo identity)인 S를 랜덤하게 생성한다. S는 균등하게 l-비트단위로 구성된 $(\{s_1,s_2,...,s_{K/l}\})$로 분할되었고, 이것들로 (K/l)개의 계수를 갖는 다항식 P의 계수로 구성한다. 계수들이 l-비트 값인 통상 퍼지볼트 시스템의 모든 수학적 연산이 유한계 $GF(2^l)$에 근거하기 때문이다. 결론적으로, 다항식 P는 차수 $d=((K/l)-1)$, $P(x)=s_1+s_2x+\cdots+s_{(K/l)}x^{(K/l)-1}$로 표현된다. 퍼지 볼트를 위해서는 세 집합 (T,N,R)을 생성하는 것이 필요하다.

첫 번째 집합은 양자화된 바이오인식 특징값인 $A=\{a_1,a_2,...,a_A\}$을 사용하여 다항식 $P(x)$를 계산함으로써 형성된 사용자 볼트 집합 T $(T=\{(a_1,P(a_1)),(a_2,P(a_2)),...,(a_A,P(a_A))\}$이다. 사용자 볼트 집합 T의 모든 성분들은 볼트를 여는(unlock) 과정에서 다항식을 재구성하기 위해 사용된다. 따라서 생성되는 점들은 최소한 $(d+1)$이어야 한다. 여기서 집합 T는 A개의 점들로 구성된다고 본다.

두 번째 집합은 몇 개의 chaff 점들로 구성된 한 집합 N이다. 이러한 점들은 사용자 볼트 집합 T를 숨기는데 중요한 역할을 한다. $N=\{(v_1,w_1),(v_2,w_2),...,(v_g,w_g)\}$(여기서 v_i는 다항식의 입력에 해당되는 x값이고, w_i는 y 값으로 여기서 사용되는 x값은 집합 T 를 구성하는 x값들과 겹치지 말아야하며 즉, $(v_i \neq a_j, i=1,2,...,g, j=1,2,...,A)$이고, 다항식 $P(x)(w_i \neq P(v_i), i=1,2,...,g)$에 위치하지 않는다는 조건을 만족하면서 유한계의 범위에서 임의로 생성되었다. 이렇게 구성되는 집합 A와 집합 N을 임의로 섞어서 최종적인 볼트인 V를 만들게 된다. 이렇게 하면 등록 과정은 모두 마치게 되며, 구해진 볼트는 스마트카드나 USB 저장장체에 저장됨으로서 개인의 프라이버시를 보호 할 수 있다.

볼트를 여는 과정은 두 개의 입력, 볼트(V)와 테스트 B를 필요로 한다. 볼트(V)는 저장장치로부터 얻어지고, 테스트 B는 위에서 보인 바와 같이, B개의 바이오특징을 취득하여 얻어진다. B를 이용하여 주어진 다항식의 계수를 Lagrange 보간법을 이용하여 추정하게 된다. 그러나 위와 같은 과정은 실제적인 바이오인식 시스템에 적용할 때 주요 문제가 있었다. 첫 번째로, 지문이나 얼굴과 같은 바이오정보는 취득할 때 마다, 외부적 변동요인에 의해서 특징값이 달라지므로 위의 과정에서와 같이 집합 A와 B가 정확하게 일치하는 경우는 드물다. 또한, 등록과정에서 만들어진 유사식별자 S에 대해서 Lagrange 보간에 의해서 만들어지는 재구성된 유사식별자 S^*와의 오차검출이나 수정을 위한 오류검출 및 정정코드(error correcting code)가 반드시 필요하다.

오류정정부호코드 중에서 퍼지 볼트 연구 분야에 가장 많이 쓰이는 것이 Reed-Solomon 코드이다[11]. Reed-Solomon(RS)코드는 1960년 Irving Reed와 Gus Solomon에 의해 개발된 오류정정코드로써 다원 BCH(Bose-Chaudhuri-Hoquenghem) 코드의 한 범주에 속한다. RS코드는 GaloisField $GF(q)$의 원소로 코드워드가 구성되고 $GF(q)$상의 심볼 단위로 인코딩 되고 디코딩되기 때문에 통신 선로 상에서 발생하는 산발오류(random error)와 연집오류(burst error)를 정정할 수 있기 때문에 각종 디지털 통신 시스템 및 데이터 저장 시스템의 신뢰성 향상을 위해 광범위하게 사용되는 오류정정 코드이다.

지금까지의 모든 대부분의 퍼지 볼트 연구에서는 RS 코드를 사용하였다. 그러나 이를 적용하기 위해서는 모든 변수나 입력값들이 GaloisField $GF(q)$상의 값들로 표현 되어야 하고 따라서, 지문이나 얼굴에서 추출되는 실수(real number)값을 갖는 특징값을 바로 사용 할 수 없다. 더욱이 다항식의 역변환을 구함에 있어서도, 기존의 대수학에서 널리 사용하고 있는 일반적인 역변화(genealized inverse)를 사용할 수 없는 제약이 있다. 이에 본 연구에서는 퍼지 볼트가 모든 실수형 연산을 기반으로 하는 실수형 퍼지 볼트를 제안하고자 한다. 이를 위래 필요한 것이 실수형 오류정보 부호화 기법(RN-ECC; real number error correcting code)이다. RN-ECC는 1984년 Marshall[12]에 의해 소개된 후 다양한 분야에서 연구되고 있다[13].

3. 얼굴 템플릿 보호를 위한 실수형 퍼지 볼트 기법

3.1 주성분 분석기법을 이용한 얼굴특징 추출

주성분 분석 기법(PCA)은 다변수 선형 데이터 분석에서 잘 알려진 기법이며 주된 개념은 데이터의 최대 분산 방향을 나타내는 상호직교 기저 벡터의 집합을 찾는 것을 목적으로 하며, 이때, PCA를 이용한 eigenface 방법은 다음과 같이 간략히 기술되어진다[14].

하나의 얼굴영상이 각각의 화소 값을 갖는 2차원 배열 $(n\times n)$이라고 가정하자. 영상 $\mathbf{z_i}$는 연속적인 행들의 연결로 $n^2\times 1$의 벡터로서 고려되어진다. 그러면 N개 얼굴영상의 학습 집합은 $Z=(\mathbf{z_1},\mathbf{z_2},...,\mathbf{z_N})$로 표시된다. 공분산 행렬을 정의하면 다음과 같다.

$$R=\frac{1}{N}\sum_{i=1}^{N}(\mathbf{z_i}-\bar{\mathbf{z}})(\mathbf{z_i}-\bar{\mathbf{z}})^T=\Phi\Phi^T \quad (1)$$

$$\bar{\mathbf{z}}=\frac{1}{N}\sum_{i=1}^{N}\mathbf{z_i} \quad (2)$$

그러면, 공분산 행렬의 고유치와 고유벡터가 계산되어진다. 여기서 r개의 가장 큰 고유치에 대응하는 r개의 고유벡터를 $E=(\mathbf{e_1},\mathbf{e_2},...,\mathbf{e_r})$라 하자. 얼굴영상의 집합 Z에 대해서, 그것들의 대응되는 특징벡터 $F=(\mathbf{f_1},\mathbf{f_2},...,\mathbf{f_N})$는 다음과 같이 Z를 PCA-변환된 공간으로 다음과 같이 투영함으로서 얻어진다.

$$\mathbf{f_i}=E^T(\mathbf{z_i}-\bar{\mathbf{z}}) \quad (3)$$

윗 식에서 $\mathbf{f_i}$는 i번째 $r\times 1$벡터이다.

3.2 RN-ECC에 기반한 실수형 퍼지 볼트

본 논문에서는 얼굴 템플릿의 특징값에 대해서 실수형 오류정보 부호 코드화를 수행하는 실수형 퍼지 볼트 방법을 적용하고자 한다[6]. 기존에 단순한 유한체 다항식에 기초하여, 다항식에 계수에 키를 Bind하여 이를 바이오 정보를 이용하여 추출하는 경우, 바이오 정보의 변동을 극복하기 위하여 에러 수정 코드나 CRC(Cyclic redundancy check)를 이용하였다. 그러나 앞에서 지적 하였다시피, 이럴 경우 구현이 복잡할 뿐더러, 오차 수정 코드로 매칭을 하는 경우에 바이오 인

식 정보의 매칭을 위해 개발 된 실수형의 매칭기를 사용할 수 없는 단점이 있다. 따라서 이를 극복하고자 실수형 다항식의 근사화 특성을 가지는 실수형오류정정코드를 적용하였다. 여기서는 간략히 중요한 개념만을 언급하고자 한다.

랜덤 키 값인 실수형 데이터 K개로 구성된 $\{x_i\}$는 길이 K인 열벡터 $\boldsymbol{x}=[x_0,x_1,\cdots,x_{K-1}]$로 표현된다. 그리고 길이 N인 코드벡터는 $\boldsymbol{y}=[y_0,y_1,\cdots,y_{N-1}]$로 표시하기로 하면, 두 벡터 간의 관계는 다음과 같이 표현할 수 있다.

$$\boldsymbol{y}=\boldsymbol{x}\boldsymbol{G} \tag{4}$$

여기서, $\boldsymbol{G}$는 rank가 K인 $K\times N$ 발생행렬(generator matrix)이다. 이러한 관계는 블록코드(block code)는 (N,K) 코드로 표시할 수 있다. 양자화잡음과 채널오차를 각각 $\boldsymbol{q}$와 $\boldsymbol{e}$로 표기하자. 행렬 $\boldsymbol{G}$는 양자화 잡음과 정정할 수 없는 오차들이 없을 경우 $\boldsymbol{x}$가 정확히 복원될 수 있도록 하는 다음의 식을 만족하는 $\boldsymbol{G}$의 $N\times K$인 오른쪽 역행렬(right inverse)이다.

$$\boldsymbol{G}\boldsymbol{G}^{-1}=\boldsymbol{I}_{\boldsymbol{K}} \tag{5}$$

Rank $N-K$인 $(N-K)\times N$ 패리티 검사 행렬 $\boldsymbol{H}$는 켤레전치(conjugate transpose) $*$를 이용하여 다음과 같이 정의된다.

$$\boldsymbol{G}\boldsymbol{H}^{*}=0 \tag{6}$$

추정하고자 하는 코드벡터 $\boldsymbol{r}$은 식 (7)과 같다.

$$\boldsymbol{r}=\boldsymbol{y}+\boldsymbol{e} \tag{7}$$

그러면 수신된 벡터 $\boldsymbol{r}$의 신드롬(syndrome) $\boldsymbol{s}$는 다음가 같이 계산된다.

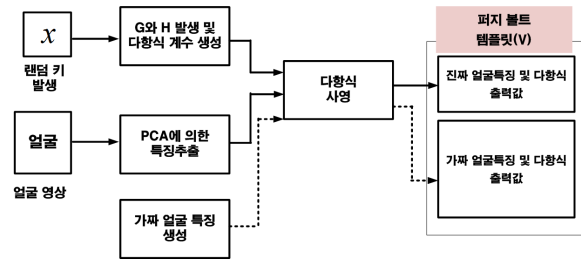
$$\boldsymbol{s}=\boldsymbol{r}\boldsymbol{H}^{*}=(\boldsymbol{y}+\boldsymbol{e})\boldsymbol{H}^{*}=\boldsymbol{e}\boldsymbol{H}^{*} \tag{8}$$

여기서 $\boldsymbol{e}$는 차원 N의 알려지지 않은 오차 패턴이며, $\boldsymbol{e}=\boldsymbol{r}-\boldsymbol{y}$ 와 같다. 따라서 오차값 $\boldsymbol{e}$가 매우 적다면 신드롬 $\boldsymbol{s}$는 매우 작은 값을 갖게 된다. 전송된 코드벡터 $\boldsymbol{y}$와 추정하고자 하는 코드벡터 $\boldsymbol{r}$이 동일한 이상적인 경우에 신드롬 $\boldsymbol{s}$는 0이 된다. 신드롬 $\boldsymbol{s}$가 0에 근접할 경우 최종적으로 구하고자 하는 열벡터 $\hat{\boldsymbol{x}}$는 다음식에 의해 추정된다.

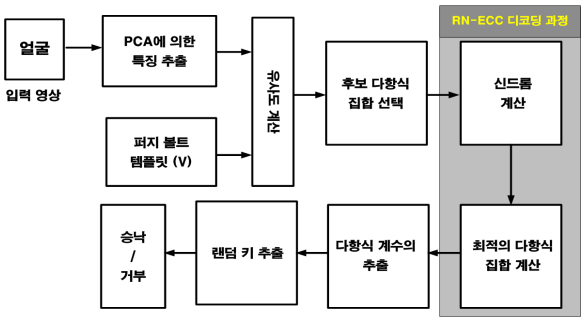
$$\hat{\boldsymbol{x}}=\boldsymbol{r}\,\boldsymbol{G}^{\boldsymbol{T}}(\boldsymbol{G}\boldsymbol{G}^{\boldsymbol{T}})^{-1} \tag{9}$$

3.3 RN-ECC 기반의 실수형 얼굴 퍼지 볼트 방법

그림 2에서는 본 논문에서 제안한 RN-ECC에 기반한 실수형 퍼지 볼트 기법을 나타내었다. 우선 등록과정에서는 등록 개인에 랜덤하게 발생된 랜덤 키 $\boldsymbol{x}$를 부여하고, 3.2에서 설명할 발생행렬(generator matrix) $\boldsymbol{G}$와 패리티 검사 행렬 $\boldsymbol{H}$를 다양한 기저함수를 이용하여 생성하고, 생성된 발생행렬 $\boldsymbol{G}$와 키 값인 $\boldsymbol{x}$ 이를 이용하여 코드벡터 $\boldsymbol{y}$를 생성한다. 본 논문에서 코드벡터 $\boldsymbol{y}$는 다항식 $p(x)$를 발생시키기 위한 다항식 계수값을 의미한다. 다음 단계로 생성된 다항식 계수값과 얼굴정보에 대한 특징값을 이용하여 다항식 $p(x)$를 산출한다. 마지막 단계에서 얼굴 특징을 보호하기 위하여 거짓 특징점을 이용한 거짓 다항식값인 chaff point를 생성하여 퍼지 볼트 템플릿을 구축한다. 퍼지 볼트 템플릿에서는 계수의 차수에 해당하는 입력부분과 출력항인 다항식으로 구성되어 있다.



(a) 등록 과정



(b)인증 과정

그림 2. 제안된 RN-ECC에 기반한 실수형 얼굴 퍼지 볼트 방법
Fig. 2. A proposed real face fuzzy vault method based on RN-ECC
(a) Registration process (b)Verification process

인증과정에서는 우선, 검증 영상에 대한 얼굴의 특징을 추출하고 추출된 특징값과 등록과정에서 생성된 퍼지 볼트 템플릿을 비교하여 유사도가 높은 후보 다항식 집합을 선택한다. 선택된 후보 다항식 집합에는 등록시 사용된 얼굴특징에 해당하는 다항식 집합뿐만 아니라 거짓 얼굴 특징에 의해 생성된 다항식 집합도 포함되어 있다. 이러한 후보 다항식 집합에서 본인의 얼굴특징에 해당하는 집합들을 선택하기 위하여 본 논문에서 제안된 RN-ECC 디코딩 과정이 수행된다. RN-ECC 디코딩 과정은 선택된 후보 다항식 집합에서 $\boldsymbol{r}$로 표현된 다항식 계수값을 추출하기 위해 필요한 모든 집합을 구성한 후, 구성된 모든 조합에 대하여 다항식 계수값 $\hat{\boldsymbol{r}}$ 값을 추정하고 식 (4)에 의해 신드롬 $\boldsymbol{s}$값을 계산한다. 최종적으로 계산된 신드롬 $\boldsymbol{s}$값 중에서 가장 최소값을 갖는 집합에서 계산된 계수값 $\hat{\boldsymbol{r}}$값을 선택하고, 선택된 계수값 $\hat{\boldsymbol{r}}$ 과 등록과정에서 생성한 발생행렬(generator matrix) $\boldsymbol{G}$를 이용한 식 (9)에 의해 개인 키 값인 $\hat{\boldsymbol{x}}$를 복원하고, 등록과정에서 생성된 키값인 $\boldsymbol{x}$와 비교하여 최종 개인 인증을 하게 된다.

제안된 방법의 장점으로는 분실 시 재생성할 수 없는 얼굴영상특징점 정보와 달리 개인 키 값인 $\boldsymbol{x}$값을 수시로 변경시킬 수 있음으로 주기적으로 키값을 변경할 수 있다는 점이다. 또한, 인증시 사용되는 퍼지 볼트 템플릿에는 개인의 얼굴인식 정보뿐만 아니라 거짓 정보를 담고 있는 얼굴인식정보도 포함되어 있음으로 등록시 사용된 개인의 얼굴

인식정보가 없는 한 퍼지볼트 템플릿이 분실되었다 하더라도 등록된 개인의 얼굴인식정보를 추출하기 어렵다는 장점이 있다.

4. 실험결과 및 분석

본 논문에서 제안된 RN-ECC를 이용한 얼굴 템플릿 보호 기법의 성능을 평가하기 위하여 ORL 얼굴 데이터 베이스를 적용하였다[15]. ORL 얼굴 데이터베이스는 서로 다른 환경에서 40명으로부터 400개의 얼굴영상을 포함하고 있다. 각 개인에 대해서 얼굴 영상의 수는 10이며, 이 영상들은 위치, 회전, 스케일, 얼굴 감정에서 변화를 주고 있다. 회전에서 변화는 최대 20도 회전하였고, 스케일에서 변화는 사람과 비디오 카메라 사이에 거리를 변화하였다. 각 영상들은 디지털화되고 0에서 255까지 그레이 값을 가진 112×92 영상 크기에 의해 나타내어진다. ORL 얼굴 데이터베이스의 일부를 그림 3에서 나타내었으며, 학습을 위해 개인당 5장의 사진을 사용하였고 나머지 얼굴영상을 이용하여 검증을 하였다.

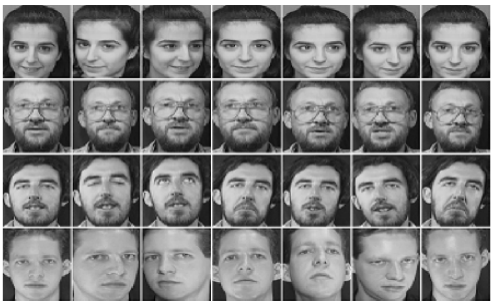


그림 3. ORL 얼굴 DB에서의 일부 얼굴영상
Fig. 3. Some face images in ORL DB

주성분 분석기법을 이용하여 36차를 갖는 얼굴특징을 추출하였다. 그러나 36차의 얼굴 특징 한 개만으로는 다항식에 기반한 퍼지볼트를 적용할 수 없기 때문에 36차의 벡터를 3차원을 갖는 12개의 벡터로 재구성하였다. 또한 얼굴 특징 템플릿은 얼굴의 실제 특징값 뿐만 아니라 얼굴 특징정보를 보호하기 위한 거짓 특징점도 포함되어 있다. 거짓 특징점의 삽입 위치가 얼굴의 실제 특징점의 위치와 매우 근접해 있다면 얼굴 인식률의 성능이 저하된다. 따라서 본 연구에서는 실제 얼굴 특징점과 유클리디언 거리로 0.02 이상인 값을 갖는 특징점들을 랜덤하게 선택하여 거짓 얼굴 특징점을 구축하였다.

그림 4에서는 주성분분석기법에 의해 얻어진 36개의 얼굴 특징을 3차원으로 재구성하여 얻어진 한 사람의 얼굴 특징값과 얼굴 특징값을 보호하기 위해 발생시킨 거짓 특징값을 나타냈다. 퍼지 볼트 템플릿을 구축하기 위해 한 사람당 5개의 얼굴 영상을 이용하였음으로 한 사람당 3차원의 특징벡터의 수는 66개이다. 또한, 얼굴 특징값을 보호하기 위하여 200개의 거짓 특징값을 삽입하여 퍼지 볼트를 구성하였다.

퍼지 볼트 템플릿에는 얼굴의 특징점 뿐만 아니라 인증과정에서 필요한 키 값을 복원하는데 필요한 다항식 출력값이 필요하다. 이를 위해 3차원으로 재구성된 얼굴의 특징점 (f_{i1},f_{i2},f_{i3})과 식 (10)에 나타낸 다항식을 이용하여 다항식 출력값 $p_i(x)$를 생성하게 된다.

$$p_{i(}x) = a_0 + a_1 f_{i1} + a_2 f_{i2} + a_3 f_{i3} + a_4 f_{i1} f_{i2} + a_5 f_{i1} f_{i3} + a_6 f_{i2} f_{i3} + a_7 f_{i1}^2 + a_8 f_{i2}^2 + a_9 f_{i3}^2 \quad (10)$$

식 (10)에 나타낸 다항식 계수 a_i를 계산하기 위해서는 랜덤 키 $\boldsymbol{x}$와 발생행렬 $\boldsymbol{G}$, 그리고 인증과정에서 계수값을 복원하는데 사용될 패리티 검사행렬 $\boldsymbol{H}$가 필요하다. 본 논문에서는 랜덤 키 $\boldsymbol{x}$= [11, 22, 33, 44, 55, 44, 33, 22]로 임의로 설정하였고, 발생행렬 $\boldsymbol{G}$와 패리티 검사행렬 $\boldsymbol{H}$는 DCT(Discrete cosine transform)행렬에 기반을 둔 기저함수를 이용하여 생성하였다.

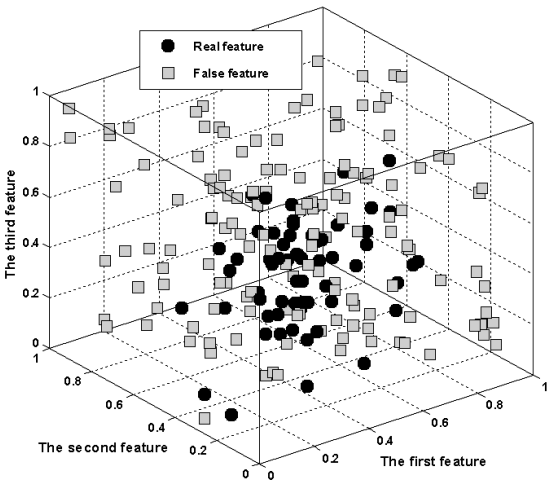


그림 4. 거짓 특징값이 포함된 얼굴 특징
Fig. 4. Face features including false features

얼굴 특징정보를 보호하기 위하여 퍼지 볼트 템플릿에는 66개의 얼굴 특징점과 랜덤하게 발생시킨 거짓 얼굴 특징점 뿐만 아니라 인증과정에서 키 값을 복원하는데 필요한 다항식 $p_i(x)$값이 포함되어 있다. 66개의 얼굴 특징점에 대해서는 식 (10)에 의해 다항식값을 산출하지만, 랜덤하게 발생시킨 거짓 특징점에 대해서는 식 (10)에 의해 산출된 다항식 $p_i(x)$ 대신에 임의의 거짓 특징정보(Chaff Point)를 삽입하게 된다. 이를 위해 거짓 특징점에 대한 다항식 $p_i(x)$값은 식 (10)에 의해 계산된 다항식 출력값에 $\pm 25\%$ 내에 존재하는 랜덤 값을 합산하여 거짓 특징정보를 발생시켰다. 그림 5에서는 200개의 거짓 특징점에 대한 거짓 특징정보를 나타냈다. 그림 4에서 알 수 있는 바와 같이 가짜 특징점에 해당하는 다항식은 식 (10)에 의해 산출된 다항식 출력값과 상이한 값이 퍼지 볼트 내에 포함되어 있음을 알 수 있다.

제안된 방법에 의한 얼굴정보를 이용한 인식과정은 다음과 같다. 우선 인증하고자 하는 얼굴의 특징과 거짓 특징점이 포함된 템플릿과의 비교를 통해 유사도가 높은 순서대로 퍼지 볼트 템플릿에 서 n개의 다항식 집합을 선택한다. 저장된 키값을 복원하기 위해서는 다항식 계수값들이 요구된다. 따라서, 다항식 계수값의 완벽한 복원은 시스템 성능을 좌우한다. 이러한 다항식 계수값을 구하기 위해서는 최소한 10개의 진짜 특징점을 가지고 있는 다항식 집합이 필요하

다. 따라서 선택된 n개의 다항식 집합은 10이상이어야 한다.

인식하고자 하는 얼굴정보에 대한 특징값의 차원은 36차원이며, 이를 퍼지 템플릿에 저장된 특징벡터와 비교하기 위하여 3차원을 갖는 12개의 특징벡터로 재구성한다. 따라서 인식하고자 하는 12개의 특징벡터와 퍼지 볼트 템플릿에 있는 특징벡터와 비교하여 유사도가 가장 높은 10개 이상의 다항식 집합을 퍼지 볼트 템플릿으로부터 구한다. 이 때 선택된 다항식 집합의 수는 최소 10개에서 최대 12개를 선택할 수 있다. 따라서 본 논문에서는 선택된 다항식 집합의 수인 n를 10에서 12까지 1씩 증가하면서 성능을 평가하였다. 선택된 다항식 집합의 수 n개 중에서 다항식 계수값을 복원하기 위해서는 10개의 다항식 집합을 선택해야 한다. 다항식 집합을 10개로 선택한 경우는 그 집합에 대해서만 고려하여 키 값을 복원하지만, 다항식 집합의 수를 11개 이상으로 할 경우 선택된 다항식 집합들 중에서 최적의 다항식 집합 10개를 선택하여야 한다. 본 논문에서는 모든 조합에 대한 전수조사를 실시하여 계산된 신드롬값을 이용하여 최적의 다항식 집합을 선택하였다. 예를 들어 n값이 12인 경우 신드롬값 계산을 위해 $C(12,10)=66$의 다항식 집합에 대한 신드롬값을 계산하고 신드롬 값이 가장 낮은 다항식 집합을 선택하여 계수값을 복원한 후, 복원된 계수값을 이용하여 키 값을 추정하게 된다.

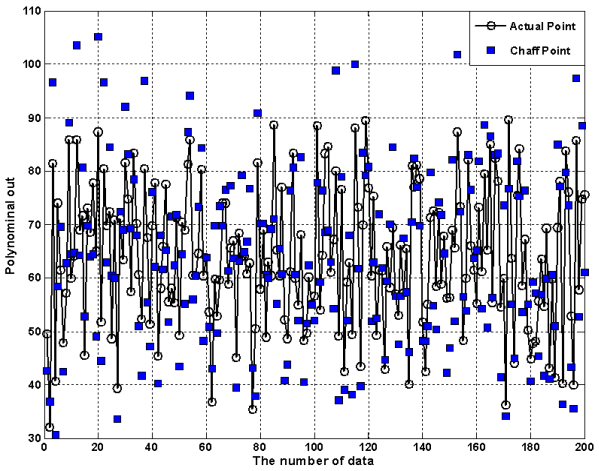


그림 5. 거짓 특징값에 해당되는 chaff point 삽입
Fig. 5. Chaff point injection corresponding with false features

표 1에서는 선택된 n값에 대한 인식결과를 나타냈다. 표 1에서 FAR은 오인식률(타인을 본인으로 인식), FRR은 오거부율(본인을 거부)을 의미한다. FRR을 위해 사용된 얼굴영상의 수는 200개(40명×5개)이고, FAR을 위해 사용된 얼굴영상의 수는 7800개(39명×5개×40set)이다. 또한 거짓 특징점의 개수는 200으로 설정하여 실험하였다. 표 1에서 보는 바와 같이 선택된 후보 다항식 집합의 수는 오거부율에는 영향을 미치지는 않으나, 오인식률 측면에서는 가장 유사도가 높은 10개만을 선택하는 것이 오차가 가장 낮게 나타났다. 따라서 본 연구에서는 선택된 후보 다항식 집합의 수를 10으로 설정하여 제안 방법의 성능을 다양하게 분석하였다.

표 1. 선택된 후보 다항식 집합의 수에 따른 인식률
Table 1. Verification rate according to the number of selected polynominal set

n	조합의 수	FAR(%)	FRR(%)
10	1	0.06	1.5
11	11	0.15	1.5
12	66	0.24	1.5

표 2에서는 선택된 후보 다항식 집합의 수를 10로 고정한 상태에서 제안방법과 PCA 기반의 얼굴인식 방법을 비교하여 나타냈다. 우선 제안방법의 경우 표 2에서 보는 바와 같이 오거부율은 거짓 특징의 수에 크게 상관없이 1.5%로 나타났다. 오인식률은 거짓 특징의 수가 증가할수록 감소하였다. 즉 거짓 특징점의 수가 100일때 오인식율은 0.46%, 거짓 특징점의 수가 1000일 때 오인식율은 0%로 나타났다. 거짓 특징점의 수가 증가할수록 보안성 측면에서는 잇점을 갖는다. 즉 거짓 특징점의 수가 100일 때 진짜 특징점을 찾기 위해서는 약 $C(112,10)=5.6594\times10^{13}$번의 조합이 필요하다. 거짓 특징점의 수가 200이상인 경우에는 조합의 수를 수치적으로 계산하기 힘들 정도로 많은 연산이 필요하다. PCA기반의 얼굴인식방법의 경우는 임계값의 증가에 따라 오인식율은 증가하고 오거부율은 감소한다.

표 2. 거짓 특징점의 수에 따른 인식률
Table 2. Verification rate according to the number of false features

제안방법			PCA 기반 얼굴인식		
거짓 특징점의 수	FAR(%)	FRR(%)	임계값	FAR(%)	FRR(%)
100	0.46	1.5	0.1	2.34	1.5
200	0.06	1.5	0.2	2.34	1.5
300	0.06	1.5	0.3	2.34	1.5
400	0.06	2	0.4	2.34	1.5
500	0.01	2	0.5	2.34	1.5
600	0.04	3	0.6	2.34	1.5
700	0	1.5	0.7	2.44	1.5
800	0	2.5	0.8	2.87	1.5
900	0	1.5	0.9	4.01	1.5
1000	0	1.5	1.0	5.6	1.5
			1.1	8.19	1.5
			1.2	11.69	1.5
			1.3	17.42	0.5
			1.4	24.20	0.5
			1.5	33.16	0

두 가지 방법을 비교하기 위하여 제안방법에서 인식 성능에 중점을 두어 거짓 특징점의 수를 700으로 설정할 때 오인식률은 0%, 오거부율은 1.5%로 나타났으며, PCA기반의 얼굴인식 방법은 임계값이 0.2일 때 오인식률은 2.34%, 오거부율은 1.5%로 나타났다. 두 방법 모두 오거부율은 동일한 결과를 나타냈지만 시스템 성능의 중요한 척도인 오인식률 측면에서는 제안방법이 매우 우수하다. 또한, 인식성능을 비교하지 않더라도 기존의 PCA기반의 얼굴인식방법은 개인의 얼굴정보를 보호 없이 사용하기 때문에 분실시 다른 대책이 없는 반면에 제안방법은 퍼지 볼트 템플릿에 의해 진짜 얼굴특징 정보를 보호할 수 있을 뿐만 아니라 키 값의

변경에 의해 퍼지 볼트 템플릿에 포함된 다항식 집합을 수시로 변경할 수 있음으로 보호측면에서도 우수한 성능을 보임을 확인할 수 있다.

5. 결 론

본 논문에서는 얼굴인식 템플릿의 특징값에 대해서 실수형 오류정보 부호 코드화를 수행하는 실수형 퍼지 볼트 방법을 새롭게 제안한다. 기존에 단순한 유한체 다항식에 기초하여, 다항식에 계수에 키를 Bind하여 이를 바이오 정보를 이용하여 추출하는 경우, 바이오 정보의 변동을 극복하기 위하여 에러 수정 코드나 CRC(Cyclic redundancy check)를 이용하였다. 그러나 앞에서 지적 하였다시피, 이럴 경우 구현이 복잡할 뿐더러, 오차 수정 코드로 매칭을 하는 경우에 바이오 인식 정보의 매칭을 위해 개발 된 실수형의 매칭기를 사용할 수 없는 단점이 있다. 따라서 이를 극복하고자 실수형 다항식의 근사화 특성을 가지는 RN(Real number) ECC(오류정정코드)를 적용하였다. 제안된 방법의 타당성을 검증하기 위하여 다양한 얼굴에 적용한 결과 기존 방법에 비하여 우수한 결과를 보임을 확인할 수 있었다.

References

[1] M. K. Muhammad, N. Marsono, Rabia Bakhteri. “Biometric encryption based on a fuzzy vault scheme with a fast chaff generation algorithm,” *Future Generation Computer Systems*, Vol. 29, No. 3, pp. 800-810, 2013.

[2] A. Marinoa, F. H. Alvarezb, L. H. Encinasb, "A crypto-biometric scheme based on iris-templates with fuzzy extractors," *Information Sciences*, Vol. 195, 15, pp. 91-102, 2012.

[3] E. C. Chang and S. Roy, “Robust Extraction of Secret Bits From Minutiae,” *Proceedings of Second International Conference on Biometrics*, pp. 750 - 759, 2007,

[4] A. B. J. Teoh, Y. W. Kuan, S. LEE, "Cancellable Biometrics and Annotations on BioHash," *Pattern Recognition*, Vol. 41, No. 6, pp.2034-2044, 2008.

[5] Dae Jong Lee, Yong-Nyuo Shin, Seon-Hong Park, Myung-Geun Chun, "RN-ECC Based Fuzzy Vault for Protecting Fingerprint Templates," *International Journal of Fuzzy Logic and Intelligent Systems*, vol. 11. no. 4, pp. 286-292, 2011.

[6] Juels A and Sudan M, "A fuzzy vault scheme," *Proceeding of IEEE Int. Symposium on Information Theory*, pp. 408, 2002.

[7] P. Li, X. Yang, K. Cao, X. Tao, R. Wang, and J. Tian, "An alignment-free fingerprint cryptosystem based on fuzzy vault scheme," *Journal of Network and Computer Applications*, Vol. 33, No. 3, pp. 207-220, 2010.

[8] Y. Wang and K. N. Plataniotis, “fuzzy vault for face based cryptographic key generation,” *Proceedings of Biometrics Symposium*, pp. 1-6, 2007.

[9] F. Thomas, Z. Xuebing, and B. Christoph, "Fuzzy Vault for 3D face recognition systems," *Int. Conf. on Intelligent information hiding and multimedia signal processing*, pp. 1069-1074, 2008.

[10] L. Yiun Joo, P. Kang Ryong, L. Sung Joo, B. Kwanghyuk, K. Jaihie, "A new method for generating an invariant iris private key based on the fuzzy vault system," *IEEE Trans. on System, Man, Cybernetics*, Vol. 38, No. 5, pp.1302-1313, 2008.

[11] K. Moon, Error Correcting Code: Mathematical Methods and Algorithm, *Wiley-Interscience*, 2005

[12] T. Marshall, "Coding of real-number sequences for error correction: a digital signal processing problem," *IEEE Journal of Selected Areas in Communication*, Vol. 2, No. 2, pp. 381-392, 1984.

[13] A. Kumar and A. Makur, "Improved coding-theoretic and subspace-based decoding algorithms for a wider class of DCT and DST codes", *IEEE Trans. on Signal Processing*, Vol. 58, No. 2, pp. 695-708, 2010.

[14] M. Turk, A. Pentland, “Face recognition using eigenfaces”, *Proc. IEEE Conf. On Computer Vision and Pattern Recognition*, pp. 586-591, 1991.

[15] ORL face database, http://www.uk.research.att.com/facedatabase.html

저 자 소 개

이대종(Dae-Jong Lee)
제22권 1호(2012년 2월호) 참조

송창규(Myung-Geun Chun)
제18권 1호(2008년 2월호) 참조

박 성 무 (Sung-Moo Park)
1980년 충북대학교 전기공학과(학사)
1982년 충북대학교 전기전자공학과 (공학석사)
2006년 충북대학교 전기전자공학과 (공학박사)
1994년～현재 한국폴리텍대학/대전, 청주,홍성캠퍼스 전기과 교수
관심분야 : 전기기기 고장진단, 인버터 속도제어, 패턴분류
Phone : +82-41-630-0594
E-mail : smp@kopo.ac.kr

전명근(Myung-Geun Chun)
제22권 1호(2012년 6월호) 참조