특집

생체정보 인증 및 위 · 변조 검출 알고리듬

임재혁, 이상윤(연세대학교 전기전자공학부)

I. 서 론

21세기 정보의 시대는 인터넷과 같은 글로벌 네트워크를 통해 원하는 정보를 편리하게 수집, 분석 및 가공할 수 있다. 그러나 이러한 개인의 중요한 정보가 타인에 의해 쉽게 도용되거나 파괴되는 심각한 문제가 제기되고 있다. 이로 인해 개인의 정보만이 손실되는 것이 아니라 국가의 중요 정보와 전자상거래 등의 경제 활동에 필요한 정보도 동시에 손실되는 현상이 발생되고 있다. 따라서 현재 많이 사용되고 있는 패스워드 또는 PIN(Personal Identification Number)만을 이용한 사용자 인증 방법으로는 개인, 산업, 그리고 국가의 중요 정보를 안전하게 보관할 수 없는 실정이다. 이러한 문제를 해결하기 위해 최근 들어 개인의 고유한 생체정보인 신체적 또는 행동학적 특징에 따라 사람들의 신원을 확인하는 바이오 메트릭(biometric) 즉, 생체정보 인식기술이 대두되고 있다.[1] 생체정보는 개인의 고유정보인 지문, 홍채, 음성, 얼굴 모양, 손의 형태, 서명, 손등의 정맥분포 등 아주 다양하다. 이것은 신체의 일부이거나, 개개인의 행동 특성을 반영하여 잊어버리거나 타인에게 대여 또는 도난당하지 않기 때문에 정보보안을 위한 새로운 분야로 활성화되고 있다[1, 2]. 그러나 생체정보 역시 개인의 주요정보이면서 프라이버시(privacy)와 관련이 있기 때문에 사용자 인증을 위하여 저장된 생체정보가 타인에게 도용이 된다면 패스워드나 PIN과 달리 변경이 불가능하여 심각한 문제를 야기할 수 있다. 이러한 문제를 해결하는 여러 방법 중 하나가 워터마킹 기법을 사용하는 것이다. 이는 생체정보에 부가정보를 삽입함으로써 해당 생체정보에 대한 인증 및 위 변조를 검출할 수 있다. 즉, 생체정보가 타인에게 도용되었을 경우 워터마킹 기법에 의해 인증 및 위 변조 여부를 판단하여 생체인식 시스템이 인증을 요구한 생체정보에 대하여 이를 수용 및 거부할 수 있다 본 고에서는 우선 생체정보를 이용한 인증 시스템의 취약점을 분석하고, 이에 대한 해결 방안 중 하나인 생체정보 인증 및 위 변조 검출 기술의 최근 연구 동향을 설명한다.

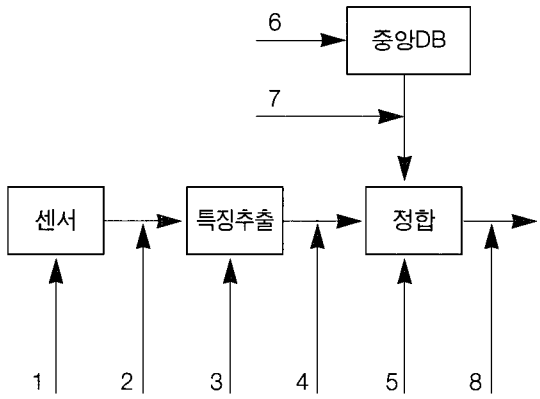
1) 반성범, 이경희, 안도성, 이남일, 생체정보 보호연구 동향, 전자통신동향분석, 제18권, 제5호, 2003년 10월.

〈표 1〉 생체인식 시스템의 공격 포인트 내용 및 대처방안

공격 포인트	공격 내용	대처방안
1	센서에 가짜 지문이나 복사한 서명, 얼굴 마스크를 이용하는 경우	Liveness 검출
2	미리 저장해 둔 생체 신호를 다시 사용하는 경우	통합시스템 설계 (센서 + 특징추출 + 정합과정)
3	침입자가 원하는 특징을 생성하도록 특징 추출단을 변경하는 경우	
4	생체인식 시스템의 인식 알고리듬을 알고 있을 때 이를 침입자가 원하는 특징을 추출할 수 있도록 변경하는 경우	
5	데이터베이스 안에 있는 템플릿을 변경하는 경우	
6	미리 선택된 정합 결과가 나오도록 시스템을 공격하는 경우	워터마크 사용
7	채널에서 전송 중인 템플릿을 수정 또는 다른 것으로 대체하여 정합 결과를 변경하는 경우	
8	최종 판결을 변경하는 경우	암호화

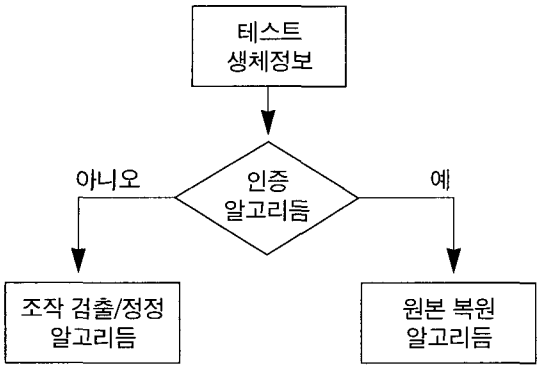
II. 생체인식 시스템의 취약점 분석 및 대처 방안

일반적인 사용자 인증 시스템에서 생체정보를 중앙 DB(Database)에 저장하고, 사용자가 인증을 요구할 경우 중앙 DB에 등록된 데이터를 인증 시스템으로 전송한 후, 인증을 요구한 사용자의 생체정보와 비교하여 이를 인증하는 형식이다. 그림 1은 이러한 전형적인 생체인식 시스템에서 가능한 공격 포인트를 나타낸 것이다[3]. 이와 같이 여러 공격 포인트가 존재하며 또한 이러한 공격을 막을 수 있는 방안도 계속 연구되고 있다. 예를 들어, 공격 포인트 1의 경우 인증을 요구한 생체정보가 인위적으로 만들어 진 허위가 아니고 실제 사용자 자신의 저장된 생체정보인지를 판단하는 liveness 검출을 통하여 방지할 수 있다. 또한 공격 포인트 2, 3, 4, 그리고 5의 경우 센서, 특징 추출, 그리고 정합 과정을 하나의 통합 시스템으로 견고하게 설계함으로써 공격을 막을 수 있다. 공격 포인트 6과 7의 경우 워터



〈그림 1〉 MPEG-21의 디지털 아이템

마킹 기법을 이용한 인증 및 위 변조 검출을 통하여 방어할 수 있다. 마지막으로 공격 포인트 8의 경우 암호화를 이용하여 해결할 수 있다. 이러한 공격 포인트들에 대한 대처 방안들은 여러 종류가 존재하며, 그 중 일반적인 예를 간략하게 정리하면 표 1과 같다. 본 고에서는 여러 대처방안 중 워터마킹을 이용한 생체정보의 인증 및 위 변조 검출 기법에 대하여 자세히 알아본다.



〈그림 2〉 생체정보 인증의 흐름도

III. 생체정보 인증 및 위·변조 검출 기술

생체인식 기술이 네트워크를 이용한 비대면 응용에 적용되기 위해서는 개인의 프라이버시 보호를 위한 생체정보의 안전한 저장 및 전송 처리 기술이 필요하다. 이러한 문제는 생체정보에 워터마킹 기법을 이용하여 인증 및 위 변조 방지를 위한 부가정보 삽입으로 해결할 수 있다. 이러한 생체정보의 인증 및 위 변조 검출 단계는 그림 2와 같다. 우선 침입자에 의한 조작 여부를 판단하는 인증 단계이다[4]. 다음은 침입자에 의해서 생체정보가 조작되었다면 어떠한 부분이 조작되었는지 검출하는 단계이다. 이러한 단계는 오류정정코드(error correction code)를 사용하여 조작된 부분을 원본으로 정정하는 연구로 최근 발전하고 있다[5]. 마지막으로 워터마크의 특성상 부가정보를 삽입할 경우 원본 생체정보의 데이터를 수정해야 하기 때문에 기존 데이터의 변경이 필수적이다. 이러한 변경은 생체인식 시스템에 있어서 인식률 저하 등과 같은 문제를 야기할 수 있다. 따라서 현재 진행되고 있는 생체정보 보호의 또 다른 방향은 만약 생체정보가 인증이 되었다면 삽입된 워터마크를 완전히 제거함으로써 원본 데이터로의 복원이 가능한 방법들이 연구되고 있다. 이러한 세 가지 방법들에 대하여 다음 절에 상세히 설명한다. 본 고에서 언급하는 생체정보란 정합단계에서 사용되는 특징 벡터가 아닌 얼굴, 지문 및 홍채 영상 또는 음성 등의 로우(raw) 데이터를 의미한다.

1. 생체정보 인증 알고리듬

그림 1의 공격 포인트 6과 7의 경우 중앙 DB에 저장된 생체정보가 침입자에 의해 변경 및 대체되었는지를 판단하기 위하여 정합단계 이전에 중앙 DB로부터 전송된 생체정보의 인증이 필수적이다. 일반적으로 생체정보의 인증을 위해서는 해쉬(hash) 알고리듬을 사용한다. 해쉬 알고리듬은 데이터의 무결성 및 메시지의 인증 등을 위하여 사용될 수 있는 함수로써 임의의 입력 데이터를 "0"과 "1"로 구성된 고정 길이(128비트)의 해쉬 코드로 압축시키는 함수이며 다음과 같은 특징을 가지고 있다.

- 다양한 가변 길이 입력 데이터에 적용 가능
- 고정된 길이의 해쉬 코드 출력
- 해쉬 코드로 입력 데이터 계산 불가능
- 동일한 해쉬 코드를 갖는 서로 다른 입력 쌍이 존재하지 않음

이러한 해쉬 알고리듬을 이용하여 생체정보 인증의 대표적인 예는 다음과 같다.

〈삽입과정〉

(1) 생체정보 화소값들에 대한 최하위 비트를 모두 "0"으로 변경한다. 이는 인증 코드를 삽입할 공간을 생성하는 것이다. 즉, "0"으

생체정보 화소값	125	135	128	231	214	175	139	140	.....	142
최하위 비트 제거	124	134	128	230	240	174	138	140	.....	142
사용자 정의 코드	1	1	0	0	0	1	1	0	.....	1
부가정보 삽입	125	135	128	230	214	175	139	140	.....	143

(a) 부가정보 삽입과정

조작된 생체정보	125	135	128	240	215	180	139	140	.....	143
최하위 비트 추출	1	1	0	0	1	0	1	0	.....	1
사용자 정의 코드	1	1	0	0	0	1	1	0	.....	1
조작된 화소 검출					X	X				

(b) 부가정보 추출과정

〈그림 3〉 위 · 변조 검출 알고리듬의 예

로 변경된 최하위 비트 위치에 생체정보의 인증과 관련된 코드를 손실 없이 삽입 및 추출할 수 있다.

(2) 해쉬 알고리듬을 이용하여 최하위 비트가 제거된 생체정보의 해쉬 코드를 계산한다.

(3) 이러한 해쉬 코드를 삽입과정 (1)에서 생성된 공간에 삽입한다. 해당 생체정보의 크기가 128비트의 해쉬 코드보다 클 경우 해쉬 코드를 삽입하고 남는 공간에 대해서 생체정보에 대한 부가 정보(저작권자, 생성날짜, 중요도 등)들을 삽입할 수 있다. 예를 들면, 100 100 크기의 얼굴영상에 대해서 해쉬 코드와 부가정보를 합하여 10,000비트를 삽입할 수 있다. 더욱 더 많은 부가정보를 삽입하기 위해서 해당 화소의 하위 비트 중 여러 비트를 "0"으로 변경하여 사용 가능하나 이에 의한 생체정보의 화질 및 음질 열화를 고려해야 한다.

〈추출과정〉

(1) 인증을 요구하는 생체정보의 화소값에 대하여 최하위 비트를 추출한 후, 이를 삽입된 해쉬 코드와 부가 정보들로 분리한다.

(2) 삽입과정과 동일한 해쉬 알고리듬을 이용하여 최하위 비트가 제거된 생체정보의 새로운 해쉬 코드를 계산한다.

(3) 두 해쉬 코드의 128비트 전체를 비교하여 동일할 경우 아무런 조작이 없었다고 판단할 수 있으며, 두 코드가 상이할 경우 침입자에 의한 조작이 있었다고 판단한다.

하지만 이러한 방법은 알고리듬이 공개될 경우 침입자에 의해서 쉽게 공격이 될 수 있다. 예를 들면, 침입자가 삽입방법과 동일하게 허위의 생체정보를 생성한 후, 중앙 DB에 저장된 원본 생체정보와 교체를 한다면 생체인식 시스템은 이러한 오류를 발견할 수 없다. 이러한 문제를 해결하기 위해서는 해쉬 알고리듬의 입력 값으로 원본 생체정보 뿐만 아니라 해당 사용자의 정보도 추가로 사용함으로써 해결할 수 있다.

2. 위 변조 검출 알고리듬

생체정보의 위 변조 검출을 위한 워터마킹 기법은 조작이 가해졌을 경우 해당 위치를 파악할 수 있는 국부적인 정보를 삽입한다. 일반적인 삽입 및 추출 과정은 다음과 같다.

<삽입과정>

(1) 생체정보의 화소값들에 대하여 최하위 비트를 모두 "0"으로 변경한다.

(2) 생체정보와 동일한 길이의 "0"과 "1"로 구성된 코드를 생성한다. 예를 들면, 100 100 크기의 생체정보에 대해서 키(key)를 사용하여 10,000비트 길이의 코드를 생성한다. 키는 생체인식 시스템에서 신뢰성 있게 보관 및 관리한다.

(3) 이러한 코드를 과정 (1)에서 생성된 최하위 비트 위치에 삽입한다.

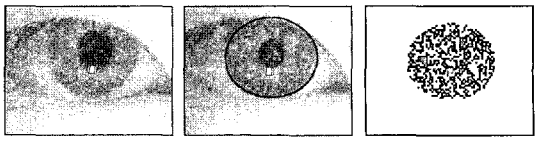
<추출과정>

(1) 인증을 요구하는 생체정보의 화소값에 대하여 최하위 비트를 추출한다.

(2) 삽입과정과 동일한 키를 사용하여 코드를 생성한 후 추출된 최하위 비트와 비교한다.

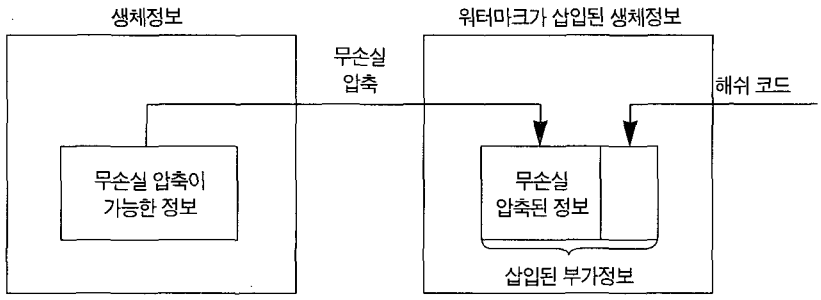
(3) 동일한 비트일 경우 조작이 없는 화소라고 판단하고, 비트가 서로 상이할 경우 조작이 있는 화소라고 판단한다. 즉 생체정보에 조작이 있었을 경우 해당 부분의 화소값에 변화가 생기고 이러한 변화는 삽입된 비트를 변화시킨다. 그림 3은 위 변조 검출 알고리듬의 예를 나타낸 것이다.

하지만 생체정보 조작에 의해서 화소값이 짝(홀)수에서 짝(홀)수로 변경되었다면 해당 화소값의 최하위 비트의 변화가 없기 때문에 조작이 있었음에도 불구하고 조작된 위치를 파악할 수가 없다. 그림 4는 이러한 결과를 나타낸 것이다. 즉 조작된 부분을 모두 찾아내는 것이 아니라 1/2 확률로 조작된 위치를 찾아낼 수 있다. 이러한 미검출 확률을 낮추기 위해서는 생성되는 사용자 정의 코드 길이를 더욱 더 길게 생성하여 해당 화소에 여러 비트 삽입하여 해결할 수 있다. 예를 들면, 100 100 크기의 생체정보에 대하여 20,000비트의 코드를 생성한 후, 해당 코드를 생체정보 화소의 하위 두 비트에 삽입한다. 이러한 경우, 조작여부를 판단하기 위하여 두 비트를 사용하기 때문에 조작이 있었을 경우 찾지 못하는 미검출 확률을 1/4로 낮출 수 있다. 하지만 이는 한 비트만 삽입한 경우에 비해서 생체정보의 화질 및 음질 열화가 더욱 더 많이 발생한다.

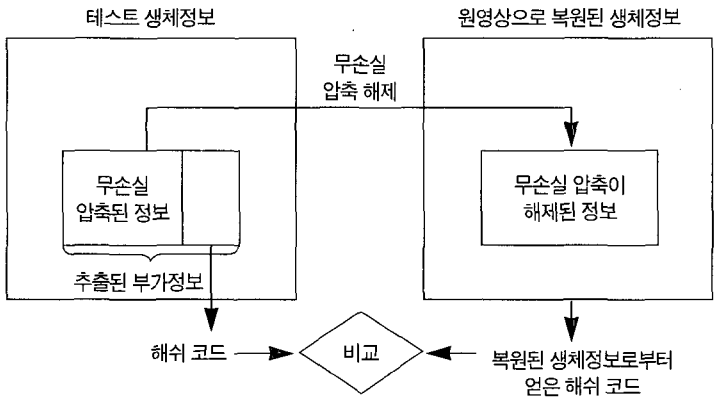


〈그림 4〉 위 · 변조 미검출의 예

최근 위 변조 검출뿐만 아니라 에러정정코드를 사용하여 검출된 부분을 정정하는 방법이 연구되고 있다[6]. 이러한 방법의 개념[9]은 블록 코드 또는 컨벌루션(convolution) 코드를 사용하여 에러정정코드를 생성하고 이렇게 생성된 코드를 워터마킹 기법으로 삽입한다. 하지만 생체정보 전체에 대한 에러정정코드를 생성하여 해당 코드를 부가정보로 삽입하기에는 데이터의 용량이 너무 크기 때문에 모두 삽입할 수 없다. 따라서 최신 위 변조 정정 알고리듬의 경우 생체정보의 중요 부분인 화소의 최상위 비트 또는 생체정보를 일정 크기로 나눈 블록의 평균값에 대



(a) 부가정보 삽입과정



(b) 부가정보 추출과정

〈그림 5〉 원본 복원이 가능한 워터마킹 알고리듬

해서만 에러정정코드를 생성하고 해당 코드를 워터마킹 기법을 이용하여 삽입하고 있다.

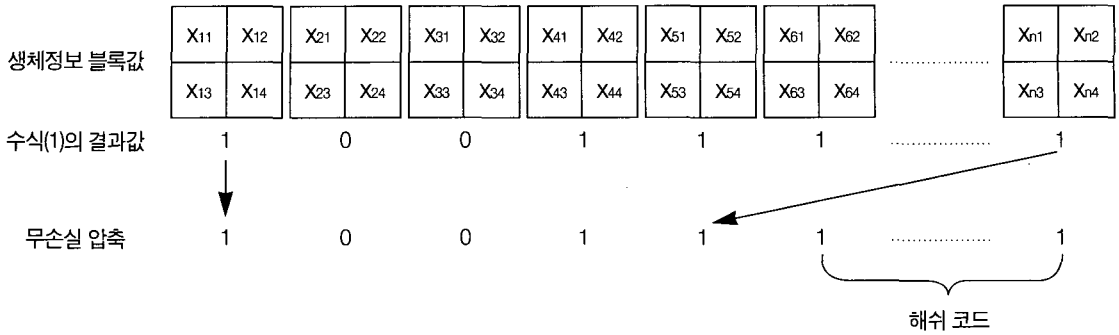
3. 원본 복원 가능 알고리듬

일반적으로 워터마킹 기법에서 사용하고 있는 부가정보의 삽입 방법은 기존의 데이터를 변경하는 것이다. 이렇게 데이터를 변경할 경우 워터마크 추출단에서 기존의 원본 데이터가 어떠한 것인지 복원할 방법이 없다. 따라서 원본 복원이 가능하도록 부가정보를 삽입하기 위해서는 그림 5와 같이 생체정보로부터 데이터의 손실 없이 부가정보를 삽입할 수 있는 공간을 먼저 생성하는 것이다. 이렇게 생성된 공간에 인증을 위한 코드를 삽입 및 추출할 수 있다. 삽입과정과 추출과정은 다음과 같다.

〈삽입과정〉

(1) 원본 생체정보에 대하여 해쉬 코드를 계산한다.

(2) 생체정보로부터 그림 5-(a)와 같이 무손실 압축이 가능한 특징벡터를 추출한다. 예를 들면, 그림 6과 같이 생체정보를 일정 크기로 분할한 후, 해당 블록에 수식 (1)을 적용하여 "0"과 "1"로 구성된 비트스트림을 생성한다. 일반적으로 영상의 인접한 화소들



〈그림 6〉 원본 복원이 가능한 워터마킹 알고리듬의 예

은 상관관계가 매우 높기 때문에 수식 (1)에 의해서 많은 수의 "0"이 발생한다. 이는 비트스트림에 대한 압축이 매우 용이하다는 것을 의미한다.

(3) 생성된 비트스트림을 무손실 압축을 한다. 무손실 압축을 할 경우 해당 비스트림의 크기가 줄어들기 때문에 해쉬 코드를 삽입할 공간이 생성된다.

(4) 생성된 공간에 과정 (1)에서 계산된 원본 생체정보의 해쉬 코드를 삽입한다.

〈추출과정〉

(1) 삽입과정의 수식 (1)을 사용하여 인증을 요구한 생체정보로부터 비트스트림을 추출한 후 압축된 비스트림과 해쉬 코드로 분리한다. 임계값 T의 경우 생체정보의 화질 및 음질을 결정하는 파라미터로써 값이 커질수록 화소값의 변화량이 크기 때문에 생체정보의 화질 및 음질 열화가 더욱 더 발생한다.

(2) 압축된 비트스트림에 대하여 그림 5-(b)와 같이 압축을 해제한 후 인증을 요구한 생체정보에 삽입하여 원본 생체정보로 복원한다.

(3) 복원된 생체정보에 대하여 새로운 해쉬 코드를 계산한다.

(4) 두 해쉬 코드를 비교하여 값이 상이한 경우 조작이 있었다고 판단하고, 값이 동일할 경우 조작이 없었다고 판단하는 동시에 복원된 생체정보가 원본 생체정보임이 판명된다.

$$f(x_n) = |x_{n1} - x_{n4}| + |x_{n2} - x_{n4}| + |x_{n3} - x_{n4}|$$
$$\begin{cases} 0 & if\ (p-1)T < f(x) \le pT \\ 1 & if\ pT < f(x) < (p+1)T \end{cases} \quad p = 1, 3, 5, .. \qquad (1)$$

이러한 방법은 생체정보에 조작이 없을 경우 삽입된 부가정보를 완전히 제거함으로써 원본 생체정보로의 복원이 가능하다. 이는 부가정보 삽입으로 인해 발생할 수 있는 생체인식 시스템의 인식률 저하를 사전에 방지할 수 있다.

앞에서 언급한 세 가지의 방법들은 독립적으로 사용되는 경우가 일반적이었으나, 최근 연구 동향은 이러한 방법들을 결합하여 동시에 인증, 위 변조 검출 및 원복 복원 가능성까지 모두 갖춘 알고리듬의 개발로 진행 중이다.

Ⅳ. 결 론

최근 사용자 인증의 정확성을 높이기 위하여

생체인식기술에 대한 관심이 고조되면서 생체정보 및 그 응용에 대한 연구 개발이 국내 외에서 활발히 전개되고 있다. 하지만 생체인식 기술이 인간의 생리학적 혹은 행동학적 특징을 추출하고 이를 분석하여 인증시스템을 생성하는 것이므로, 생체정보 그 자체가 인간의 특성을 나타낸다는 점에서 생체정보 보호는 대단히 중요하다. 따라서 본 고에서는 생체정보 보호를 위하여 우선 생체인식 시스템의 취약점을 분석하고, 이의 해결 방안 중 하나인 생체정보 인증 및 위 변조 검출 기술의 최근 연구 동향을 살펴보았으며 간략한 특징은 다음과 같다.

- 생체정보 인증 알고리듬
  : 침입자에 의한 조작 여부 판단
- 위 변조 검출 알고리듬
  : 조작된 생체정보의 위치 파악
- 원본 복원 알고리듬
  : 조작이 없을 경우 원본으로 복원 가능

위와 같은 방법을 생체인식 시스템에 효과적으로 적용한다면 개인의 생체정보 침해에 대한 우려를 완화시킬 수 있을 것으로 기대된다.

참고문헌

[1] N. Ratha, J. Connell, and R. Bolle, "An analysis of minutiae matching strength," Proc. of AVBPA 2001 (LNCS 2091), pp. 223-228, 2001.

[2] http://www.biometrics.or.kr/

[3] A. Jain, and U. Uludag "Hiding biometric data," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 25, issue 11, pp. 1494-1498, Nov. 2003.

[4] J. Fridrich, M. Goljan, and R. Du, "Lossless data embedding for all image formats," Proceeding of SPIE Photonics West, vol. 4675, pp. 572-583, Jan., 2002.

[5] Y. Shengsheng, and Z. Jingli, "Content-based watermarking scheme for image authentication," Proc. of ICARCV 2004 vol. 2, pp. 1083-1087, 6-9 Dec. 2004.

[6] L. Phen-Lan, H. Po-Whei, and P. An-Wei, "A fragile watermarking scheme for image authentication with localization and recovery," Proceedings of IEEE Sixth International Symposium on Multimedia Software Engineering, 2004. pp. 146-153, 13-15 Dec. 2004.

[7] J. H. Lim, H. B. Lee, S. Y. Lee, and J. H. Kim, "Invertible Watermarking Algorithm with Detecting Locations of Malicious Manipulation for Biometric Image Authentication," International Conference Advances in Biometrics 2006, Hong Kong, China, pp. 763-769, Jan. 2006.

[8] E. Abdel-Azeem, R. Seireg, and S. I. Shaheen, "Cryptographic security evaluation of MD4 hash function," Radio Science Conference, 1996. NRSC '96., Thirteenth National pp. 345-354, 19-21 March 1996.

[9] J. G. Proakis, "Digital Communication," Forth edition, McGraw-Hill Series in Electrical and Computer Engineering, 2001.

저자소개

임 재 혁

1997년 동국대학교 전자공학과 학사
1999년 동국대학교 전자공학과 석사
2003년 동국대학교 전자공학과 박사
2003년-현 재 연세대학교 전기전자공학부 연구교수
주관심분야 영상분할 및 추적, 워터마킹, 비디오 코덱 등

이 상 윤

1987년 연세대학교 전자공학과 학사
1989년 연세대학교 전자공학과 석사
1999년 Georgia Institute of Technology 박사
2004년-현 재 연세대학교 전기전자공학부 조교수
1989년-2004년 KT 선임연구원,
주관심분야 신호 및 영상처리, 컴퓨터 비젼, 생체인식 등