

# 인식률을 향상시키는 효과적인 Rank-level fusion 방법

안정호<sup>1\*</sup>, 권태연<sup>2</sup>, 노건태<sup>3</sup>, 정익래<sup>4</sup>

<sup>1\*,2,4</sup>고려대학교 정보보호대학원 정보보호학과, <sup>3</sup>서울사이버대학교 정보보호학과

e-mail : {pumphelp<sup>1\*</sup>, kte4567<sup>2</sup>, irjeong<sup>4</sup>}@korea.ac.kr

gnoh<sup>3</sup>@iscu.ac.kr

## A efficient Rank-level fusion method improving recognition rate

Jung-Ho Ahn<sup>1\*</sup>, Taeyean Kwon<sup>2</sup>, Geontae Noh<sup>3</sup>, Ik Rae Jeong<sup>4</sup>

<sup>1\*,2,4</sup>Department of Information Security, Graduate School of Information Security, Korea University

<sup>3</sup>Department of Information Security, Seoul Cyber University

### 요 약

생체정보를 이용한 사용자 인증은 차세대 인증 방법으로서 기존의 인증 시스템에서 급진적으로 사용되고 있는 인증 방법이다. 현재 대부분의 생체인증 시스템은 단일 생체정보를 이용하고 있는데, 단일 생체인증 시스템은 노이즈로 인한 문제, 데이터의 질에 대한 문제, 인식률의 한계 등 많은 문제점들을 가지고 있다. 이를 해결하기 위한 방법으로 다중 생체정보를 이용하는 사용자 인증 방법이 있다. 다중 생체인증 시스템은 각각의 정보에 대한 information fusion을 적용하여 새로운 정보를 생성한 뒤, 그 정보를 기반으로 사용자를 인증한다. information fusion 방법들 중에서도 Rank-level fusion 방법은 표준화 작업이 필요하고 높은 계산 복잡도를 갖는 Score-level fusion 방법의 대안으로 선택되고 있다. 따라서 본 논문에서는 기존 방법보다 정확도가 높게 향상된 Rank-level fusion 방법을 제안한다. 또한, 본 논문에서 제안하는 방법은 낮은 정확도를 갖는 matcher를 사용하더라도 정확도를 향상시킬 수 있음을 실험을 통해 보이고자 한다.

### 1. 서론

현재 일상에서 사용되고 있는 대부분의 사용자 인증 방법은 패스워드나 PIN을 사용하고 있다. 하지만 망각, 분실 등의 이유로 높은 보안 성능을 제공하지 못한다. 반면에 생체정보를 이용한 사용자 인증은 사용자의 고유한 생체정보를 이용하기 때문에, 망각 및 분실의 문제가 없어서 언급한 방법에 비해 높은 보안 성능을 제공할 수 있다.[8]

현재 대부분의 생체인증 시스템은 단일 생체정보를 이용하고 있는데, 노이즈로 인한 문제, 데이터의 질에 대한 문제, 인식률의 한계 등 많은 문제점들을 가지고 있다. 이를 해결하기 위한 방법으로 다중 생체정보를 이용하는 사용자 인증 방법이 있다.

다중 생체정보를 이용하여 사용자를 인증하는 경우, information fusion 과정이 필요하다. information fusion은 아래와 같이 5가지 level에서 수행될 수 있다[1].

- Sensor-level
- Feature-level
- Score-level
- Rank-level
- Decision-level

Sensor-level과 Feature-level에서 information fusion을 수행할 경우, 각 생체정보는 고유한 특성을 가지기 때문에 생체정보간의 호환성이 낮고 호환성이 높은 생체정보라도 information fusion을 하는데 있어서 높은 계산 복잡도를 갖기 때문에 비효율적이다.

Score-level에서 information fusion을 수행할 경우, 높은 정확도를 가질 수 있지만, 각 matcher들의 score값들이 서로 이질적이기 때문에 정규화 작업을 수행해야 한다. 하지만 score를 정규화하기 위한 방법들은 높은 계산 복잡도를 갖기 때문에 비효율적이다.

Decision-level에서 information fusion을 수행할 경우, 계산 복잡도가 앞선 방법들에 비해 낮지만 fusion에 사용할 수 있는 information이 심각하게 적기 때문에 fusion을 하더라도 정확하지 않은 information을 얻게 된다.

Rank-level에서 information fusion을 수행할 경우, Score-level에서 fusion을 하는 것만큼의 정확도를 갖는 information을 얻을 수 있고, Decision-level에서 fusion을 하는 것만큼의 낮은 계산 복잡도를 갖는다. 또한 정규화 작업이 필요 없기 때문에, 여러 가지 생체정보를 보다 쉽게 결합할 수 있다.

본 논문에서는 기존에 존재하는 Rank-level fusion 방법

보다 정확하고, 인식률이 낮은 matcher를 사용하더라도 정확도를 향상시킬 수 있는 효과적인 Rank-level fusion 방법을 소개한다.

## 2. 배경지식

### 2.1. 알고리즘

생체정보를 이용해 사용자를 인증하기 위해서 먼저 이미지에 대한 특징점들을 추출해야한다. 이 논문에서는 특징점을 추출하는 알고리즘으로 Daugman's iris recognition[2] 방법과 주성분 분석(Principal Component Analysis)[3]을 사용한다.

#### 2.1.1. Daugman's iris recognition

Daugman이 제안하는 홍채인식 알고리즘은 먼저 노이즈를 제거하기 위해 이미지에 대한 Gaussian blurring을 적용한다. 노이즈를 제거한 이미지에 아래의 수식과 같은 Integro-differential operator를 이용하여 홍채의 바깥쪽 경계와 안쪽 경계를 찾는다.

$$\max_{(r, x_0, y_0)} \left| G_{\sigma}(r) * \frac{\partial}{\partial r} \oint_{(r, x_0, y_0)} \frac{I(x, y)}{2\pi r} ds \right|$$

홍채의 경계를 찾기 위해 중심과 반지름을 바꿔가면서 원주에 해당하는 픽셀들의 단위길이 당 합의 변화량이 가장 큰 곳으로 동공 및 홍채의 경계를 찾는다. 경계를 찾은 뒤 rubber sheet 방법을 이용하여 원 모양의 홍채를 직사각형의 모양으로 정규화 시킨다. 정규화된 이미지에 2D Gabor wavelet을 사용하여 홍채 코드를 생성한다. 마지막으로 생성된 홍채 코드와 저장된 홍채 코드의 유사도를 측정하는데, 해밍 거리측정 방법을 사용한다. 두 홍채 정보들의 유사도 값이 사전에 정해진 임계값 이상이 되면 동일인으로 판단하고, 그렇지 않으면 다른 사람이라고 판단한다.

#### 2.1.2. 주성분 분석

주성분 분석은 n개의 주어진 이미지들의 일차결합으로 주성분을 구성하고, 분산이 큰 순서대로 제 1 주성분, 제 2 주성분, ..., 제 m 주성분을 구하여 m개의 주성분으로 전체 이미지들을 나타내고자 하는 것이 목적이다. 전체 이미지 개수가 n개라 하자.

$$X = (x_1, x_2, \dots, x_n)$$

$$S = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T$$

여기서 S는 공분산행렬,  $\mu$ 는 전체 이미지에 대한 평균이다.  $W = (w_1, w_2, \dots, w_m)$ 를 m개의 가장 큰 고유값에 대응하는 m개의 고유벡터라 하자. 그러면 원래 이미지 집합 X에 대해서 그에 대응하는 특징벡터  $Y = (y_1, y_2, \dots, y_m)$ 는 다음과 같이 구할 수 있다.

$$y_i = W^T(x_i - \mu)$$

## 3. 제안하는 방법

본 논문에서 제안하는 방법은 다음과 같다.

$$C_p = \sum_{i=1}^N (\log(w_i r_i(p) + 1)) + w_i \log(r_i(p) + 1)$$

여기서  $w_i$ 는 i번째 matcher에 할당된 가중치를,  $r_i(p)$ 는 p번째 사람에 대한 i번째 matcher의 우선순위 값을,  $C_p$ 는 p번째 사람에게 부여된 새로운 우선순위를 나타낸다.

로그함수(log(x+1))의 그래프는 0으로 가까이 갈수록 함수 값이 가파르게 감소하고, 점점 멀어질수록 함수 값이 부드럽게 증가하는 성질을 갖는다. 이 성질을 이용하여 일부 범위에서 우선순위 값을 확대 또는 축소시켜 효과적인 새로운 우선순위를 생성할 수 있다. 또한 가중치를 우선순위에 곱하여 보다 정확한 새로운 우선순위를 생성할 수 있다.

## 4. 실험 및 결과

### 4.1. 실험 환경 및 방법

<표 1> 실험 환경 및 실험 데이터 정보

Desktop	
운영체제	Window 8.1K 64bit
CPU	Intel Core i7-4790 3.60GHz
RAM	8.00GB
Matlab	
버전	R2015a(Win_64bit)
Dataset	
이름	CASIA
버전	V3

Rank-level fusion 방법의 정확도를 측정하기 위해서 먼저 Daugman의 알고리즘과 주성분 분석을 이용하여 다음과 같은 방법으로 우선순위를 부여한다.

- ① 임의의 홍채 이미지를 선택한 뒤, Daugman의 알고리즘을 이용하여 미리 생성한 템플릿과 비교하였을 때 유사한 홍채 이미지에 대해서 유사한 순서대로 우선순위를 부여한다.
- ② 주성분 분석을 통해 Daugman의 알고리즘을 수행할 때 생성되는 홍채 코드에 대한 특징벡터들을 추출해낸다. 추출된 특징벡터들을 유클리디안 거리측정 방법을 이용해 임의의 홍채 코드들과의 거리를 잰다. 유클리디안 거리가 가까운 홍채 코드들에 대해 거리가 가까운 순서대로 우선순위를 부여한다.

①과 ②에서 생성한 우선순위를 이용하여 기존 Rank-level fusion 방법들과 본 논문에서 제안한 Rank-level fusion 방법의 정확도를 비교실험 하였다. 단, D는 Daugman의 알고리즘을 의미하고 P는 주성분 분석을 의미한다.

본 논문에서 제안한 방법의 가중치는 각각 0.95, 0.05로 설정하였다. 즉  $w_1=0.95, w_2=0.05$ 로 설정하였다. 이렇게 설

정한 이유는 가중치를 변경하면서 많은 실험을 한 결과 0.95, 0.05가 가장 정확하기 때문이다. 또한 Logistic Regression, [7] 방법의 가중치도 각각 0.95, 0.05로 설정하였고, [5] 방법의 가중치는 각각 0.42, 0.05, 0.53으로 설정하여 실험을 진행하였다.

- Case 1) D
- Case 2) P
- Case 3) D + P + Highest Rank
- Case 4) D + P + Borda Count
- Case 5) D + P + Logistic Regression
- Case 6) D + P + Mixed Group Rank
- Case 7) D + P + Inverse Rank Position
- Case 8) D + P + Nonlinear Weighted
- Case 9) D + P + Proposed method

<표 2> 알고리즘 및 fusion 방법의 인식률

알고리즘 및 fusion 방법	인식률 (%)
Daugman의 알고리즘[1]	94.59
주성분 분석[3]	27.77
Highest Rank[4]	65.54
Borda Count[4]	66.89
Logistic Regression[4]	94.59
Mixed Group Rank[5]	94.59
Inverse Rank Position[6]	91.22
Nonlinear Weighted[7]	91.89
Proposed method	95.27

<표 3> fusion 방법에 따른 소요시간

fusion 방법	소요시간 (s)
Highest Rank	$1.86 \times 10^{-5}$
Borda Count	$1.26 \times 10^{-5}$
Logistic Regression	$1.18 \times 10^{-5}$
Mixed Group Rank	$1.32 \times 10^{-5}$
Inverse Rank Position	$1.29 \times 10^{-5}$
Nonlinear Weighted	$1.27 \times 10^{-5}$
Proposed method	$1.94 \times 10^{-5}$

기존에 제안된 Rank-level fusion 방법을 사용하여 새로운 우선순위를 생성한 뒤 Daugman의 알고리즘과 비교하였을 때, 인식률이 약 3%에서 30%정도 낮다. 이는 fusion을 수행하기 위해 사용되는 matcher들 중에서 인식률이 상대적으로 너무 낮은 matcher를 사용하면, fusion을 하더라도 인식률을 향상시킬 수 없다는 의미 있는 결과이다.

하지만 본 논문에서 제안하는 fusion 방법은 인식률이 낮은 matcher를 사용함에도 불구하고 인식률을 향상시키

는 결과를 보였다.

### 5. 결론

본 논문에서는 생체정보를 이용해 사용자를 인증할 때, 인식률을 향상시키기 위해 사용하는 information fusion 방법들 중 하나인 Rank-level fusion 방법을 소개하였다.

일반적으로 생체정보를 이용해 사용자를 인증할 경우, information fusion을 하면 인식률이 향상된다. 하지만, 본 논문과 같이 사용되는 matcher 중 인식률이 낮은 matcher를 사용하게 되면 information fusion을 하더라도 인식률이 낮아진다. 따라서 본 논문에서는 이러한 환경에서도 인식률을 향상시킬 수 있는 Rank-level fusion 방법을 제안하였으며, 이는 기존의 Rank-level fusion 방법들의 계산복잡도와 차이가 크지 않기 때문에 인식률이 낮은 matcher를 사용하는 환경에서 유용하게 사용할 수 있다.

### ACKNOWLEDGMENT

본 연구는 미래창조과학부 및 정보통신기술진흥센터의 SW 중심대학지원사업의 연구결과로 수행되었음 (2015-0-00936)

### 참고문헌

[1] A. Ross, K. Nandakumar, and A. Jain, "Handbook of multibiometrics.", Springer Science & Business Media, Vol. 6, 2006

[2] J. Daugman, "How iris recognition works.", IEEE Transactions on circuits and systems for video technology, Vol 14, pp.21-30, 2004

[3] I. Jolliffe, "Principal component analysis.", John Wiley & Sons, Ltd, 2002

[4] T. K. Ho, J. J. Hull, and S. N. Srihari, "Decision combination in multiple classifier systems.", IEEE transactions on pattern analysis and machine intelligence, Vol. 16, pp.66-75, 1994

[5] O. Melnik, Y. Vardi, and C. H. Zhang, "Mixed group ranks: Preference and confidence in classifier combination.", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, pp.973-981, 2004

[6] M. Jović, Y. Hatakeyama, F. Dong, K. Hirota, "Image retrieval based on similarity score fusion from feature similarity ranking lists.", International conference on Fuzzy Systems and Knowledge Discovery, Vol. 4223, pp. 461-470, 2006

[7] A. Kumar, and S. Shekhar, "Personal identification using multibiometrics rank-level fusion.", IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), Vol. 41, pp.743-752, 2011

[8] 문기영, "특집 생체인식 기술현황 및 전망", TTA저널/정보통신표준화소식, pp.38-47, 2005