

유전자 알고리즘을 이용한 홍채 특징 추출

원현석⁰, 손병준, 이일병
⁰연세대학교 컴퓨터정보과학부
{basket98⁰, sonjun, yblee}@csai.yonsei.ac.kr

Korea Information Science Society

Hyunsuk Won⁰, Byungjun Son, Yillbyung Lee
⁰Division of Computer and Information Engineering,
Yonsei University

요약

홍채인식 시스템은 영상획득, 전처리, 특징 추출, 패턴 정합의 단계로 이루어져 있다. 이 중 특징 추출은 특징 차원의 감소뿐만 아니라 분류 정확도의 증가를 위한 필수적인 과정이다. 본 논문에서는 특징을 추출하는데 있어서, 홍채데이터에 웨이블릿 변환의 다해상도 분석 기법을 시도하여 일정 영역을 추출한 후, 그 영역에 유전자 알고리즘(Genetic Algorithm)을 적용하여 가장 본별력있는 특징들만을 추출 및 사용하는 홍채인식 시스템을 제안한다. 유전자 알고리즘의 선택연산자로는 적용도 비례 방식과 전역 엘리트 방식을 사용하였으며, 적합도 함수로는 Gaussian Kernel을 사용하는 Support Vector Machine(SVM)을 사용하였다. 본 시스템을 통해 나온 최적의 특징집합을 이용한 SVM분류기로 인식률을 알아본 결과 웨이블릿만을 사용했을 때 보다 대략 1.5%정도 더 좋은 인식률을 얻을 수 있었다.

1. 서론

생체인식은 측정 가능한 생체 정보를 이용하여 특정 개인을 확인 및 식별하는 것으로써 물리적, 생리학적(정적)정보로는 지문, 홍채, 얼굴, 장문, 정맥 등을 사용하고, 행동적(동적)정보로는 서명, 음성 등을 사용한다. 이 중 홍채는 만 3세 이전에 대부분 생성이 되고, 평생 동안 변하지 않으며, 개인에 대한 변별력이 가장 높다. 또한, 신체 기관인 눈 속에 있음으로써 고도의 보안성을 확보할 수 있으며, 비접촉 방식으로 특징을 추출할 수 있다. 이러한 특징으로 인해 홍채인식 시스템은 큰 주목을 받고 있으며 계속적인 연구가 진행 중에 있다.

홍채인식 시스템은 영상획득, 전처리, 특징 추출, 패턴 정합의 단계로 이루어져 있다. 이 중 특징 추출 단계는 분류의 정확도를 유지 혹은 증가시키면서, 특징의 차원을 감소시키는데 목적이 있다. 즉, 분류능력이 떨어지는 특징들을 제거하며, 클래스들을 잘 분류하는 효과적인 특징들만을 남기는 것이다.[1]

본 논문에서 제안하는 시스템은 특징 추출을 위해 웨이블릿과 유전자 알고리즘을 사용하였다. 웨이블릿 변환을 통해 원영상의 고유정보들을 최대한 보존하면서 특징

차원을 감소시켰으며, 유전자 알고리즘을 통해 그 감소된 특징들 중 클래스들을 가장 잘 분류할 수 있도록 하는 최적의 특징집합을 추출하였다. 패턴 분류기로는 최근 생체인식 분야에 많이 쓰이고 있는 Support Vector Machine(SVM)을 사용하였다.

2. 본론

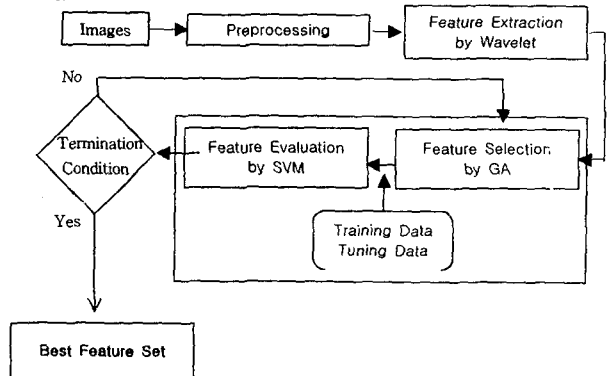


그림 1. 전체 시스템 구성도

본 논문에서 연구되어진 전체 시스템 구성도는 위와 같다. 2.1절에서는 연구에 쓰여진 웨이블릿에 대해서, 2.2 절에서는 유전자 알고리즘에 대해서 알아보기로 하겠다.

“본 연구는 과기부 뇌신경정보화사업으로부터 부분적인 지원을 받아 수행되었음”

2.1 웨이블릿을 통한 특징추출

본 연구에서는 특징을 추출하는 첫 번째 단계로 웨이블릿 변환을 시도하였다. 웨이블릿은 최근 이미지압축, 영상분석, 패턴인식, 그리고 컴퓨터비전 등에 많이 쓰이고 있는 이론으로서 여기에선 다해상도 분석(multiresolution analysis)기법을 적용하였다.[2] 이는 행방향, 열방향으로 scaling signals와 wavelets를 저주파영역에 대해서만 반복적으로 적용하는 기법이다.

특히 본 연구에서는 Daubechies 웨이블릿을 사용하여, 그림 2, 그림 3에서 보는 바와 같이 LL3 영역만을 취하였다. 이는 원영상의 정보를 크게 손실함이 없이 특징차원을 1/64크기로 줄여준다. 또한 유전자 알고리즘의 입력 벡터차원이 줄어들어서 계산복잡도가 감소하며, 특징들의 분별력은 증가시키는 효과를 얻을 수 있다.

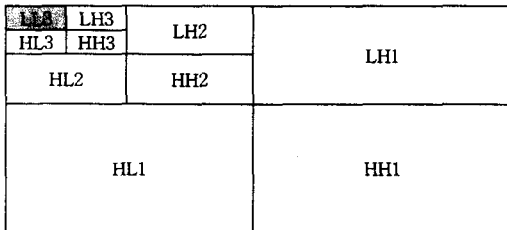


그림 2. 웨이블릿을 통한 다해상도 분석



그림 3. 3단계 분할된 홍채 영상

2.2 유전자 알고리즘을 통한 특징 추출

다음 과정은 위에서 축소된 특징들 중에서 유전자 알고리즘을 이용하여 가장 식별력이 좋은 특징들을 추출하는 것이다. 유전자 알고리즘은 자연세계의 유전과 진화 메커니즘에 기반한 계산 모델이다.[3] 풀고자 하는 문제에 대한 가능한 해들을 정해진 형태의 자료구조로 표현한 다음, 이들을 점차적으로 변형함으로써 점점 더 좋은 해들을 생성해 나간다. 각각의 가능한 해를 하나의 유기체 또는 개체(Indivisual)로 보며 이들의 집합을 개체군(Population)이라 한다. 하나의 개체는 보통 한 개의 염색체로 구성되며, 염색체를 변형하는 연산자를 유전연산자라 한다. 유전연산자로는 선택(Selection)연산자, 교배(Crossover)연산자, 돌연변이(Mutation)연산자가 있다.

본 논문에서는 선택연산자로는 적응도 비례(룰렛 휠)방식과 전역 엘리트(Global Elitism)방식을 사용하였다.[4]

그리고 교배연산자로는 하나의 교배포인트만을 사용하는 단순 교배를 사용하였다.

다음은 유전자 알고리즘의 일반적인 구조이다. 초기 개체군을 랜덤하게 '1'과 '0'로 구성한다. 여기서 '1'은 유전자 알고리즘의 입력으로 들어온 특징벡터 중에서 그 위치의 특징을 사용한다는 의미이고, '0'은 사용하지 않는다는 의미이다. 이렇게 결정된 개체군을 개체가 좋은 유전자를 갖고 있는지를 알아보기 위해 적합도 함수를 통해 평가한다. 적합도 함수는 각 개체들이 다음 세대에 선택될 만큼 훌륭한 유전자를 가지고 있는지를 평가하는 함수이다. 본 논문에서는 적합도 함수로 Gaussian Kernel을 사용하는 Support Vector Machine을 사용하였으며, 이것을 통해 나온 인식률을 적합도의 평가값으로 사용하였다. 이렇게 적합도 함수를 통해 평가되어진 개체들을 선택, 교배, 돌연변이 연산자를 통해 진화시켜나간다. 진화된 개체들을 다시 적합도 함수를 통해 평가한다. 위의 진화과정을 진화 세대수 혹은 평가값이 일정수치에 이를 때까지 반복한다.

```

generate initial population G(0);
evaluate G(0);
t := 0;
repeat
    t := t + 1;
    generate G(t) using G(t-1);
    evaluate G(t);
until termination condition has reached;
    
```

그림 4. 유전자 알고리즘의 진화 구조

진화과정이 끝난 후에 나온 개체는 시스템이 찾고자 하는 가장 식별력 있는 특징들이며, 이를 이용해 최적의 SVM분류기를 만들 수 있다.

3. 실험 환경 및 결과

실험은 CPU PENTIUM IV 1.5 MHz, MEMORY 512MB, OS WINDOWS 2000 PROFESSIONAL, MS Visual C++ 6.0 에서 수행되었다. 사용 데이터는 고정 초점 CCD 카메라를 통해 그레이 레벨의 320×240의 해상도로 획득된 100명에 대한 900개의 데이터를 사용하였고, 이를 300개(100*3)의 학습데이터, 300개(100*3)의 튜닝데이터, 300개(100*3)의 테스트데이터로 나누어 사용하

였다. 튜닝데이터를 생성한 이유는 적합도 함수의 값으로 Support Vector Machine의 분류능력을 사용하기 때문에 인식률을 내기 위한 실험데이터가 아닌 중간단계의 새로운 데이터가 필요했기 때문이다.

홍채데이터에 Daubechies 웨이블릿 변환을 사용하여 LL3만을 취한 결과 7200차원에서 113차원으로 특징차원을 줄일 수 있었다. 유전자 알고리즘에서는 개체군(population)의 크기를 15, 교배확률을 0.4, 돌연변이 확률을 0.01으로 하여 실험하였다. 또한 진화세대는 11세대로 하여 실험하였다. 개체들은 초기에 랜덤하게 '1'과 '0'으로 구성되기 때문에 평균적으로 113차원의 반 정도되는(55 ~ 58) 특징차원을 갖게 되어 메모리가 현저하게 감소되는 효과를 얻을 수 있었다.

다음 그림은 각 세대에서의 적합도함수의 평균평가값(평균인식률)을 비교한 것이다. 그림 5에서 보는 바와 같이 세대가 지날수록 인식률이 커지다가 어느 세대부터는 일정하게 수렴하는 현상을 볼 수 있다.

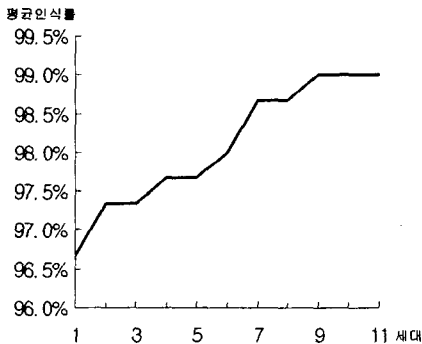


그림 5. 세대간 평균인식률 비교

다음은 웨이블릿만 사용했을 때의 특징추출방법과 제안하는 유전자 알고리즘을 함께 사용했을 때의 특징추출방법을 적용해서 얻은 인식률을 비교한 표이다. 표 1에서 보는 바와 같이 웨이블릿만을 사용했을 때보다 유전자 알고리즘으로 보다 분별력 있는 특징들만을 뽑아낸 것이 더욱 좋은 인식률을 보였다.

특징추출법	W	W + GA
인식률	97.50%	99.00%

표 1. 타 특징추출법과의 인식률 비교
(W : 웨이블릿, GA : 유전자 알고리즘)

4. 결론 및 향후 연구 계획

본 논문에서는 홍채데이터에 웨이블릿 변환과 유전자 알고리즘을 이용하여 특징을 추출하는 방법을 제안하였다. 홍채데이터에 웨이블릿을 적용하여 정보를 최대한 보존하면서 상당량의 차원이 축소된 LL3영역을 뽑아낸 후 여기에 유전자 알고리즘을 적용하여 가장 식별력 있는 특징들만을 뽑아내었다. 유전자 알고리즘은 복수 개의 개체들 간의 상호 협력에 의해 해를 탐색하였으며, 특징차원의 크기면에서 볼 때 웨이블릿 변환으로 뽑아낸 LL3영역의 반 정도 되는 결과를 보였다. 한편, 유전자 알고리즘의 적용 시 초기 개체군의 크기와 종류를 설정하는데 어려움이 있었고, 어떠한 선택연산자와 교배연산자를 사용하느냐에 따라 성능이 큰 차이를 보였다.

향후연구에서는 시스템의 최적화를 위하여 유전자 알고리즘과 유전자 프로그래밍 혹은 시뮬레이티드 어닐링 기법을 조합함으로써 더 좋은 인식률을 얻을 수 있음을 기대해 본다.

참고 문헌

- [1] Michael L.Raymer, William F.Punch, Erik D.Goodman, Leslie A, Anil K. Jain, "Dimensionality Reduction Using Genetic Algorithms", *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*, VOL. 4, NO.2, 7 2000
- [2] 기관도, 변영철, 이관용, 이일병, "웨이블릿 변환을 기반으로 하는 홍채 특징 추출", 한국정보과학회 추계 학술대회, 2002
- [3] David E. Goldberg, "GENETIC ALGORITHMS in Search, optimization & Machine Learning", *Addison Wesley*, 1989
- [4] J. A. Vasconcelos, J. A. Ramirez, R. H. C. Takahashi, and R. R. Saldanha, "Improvements in Genetic Algorithms", *IEEE TRANSACTIONS ON MAGNETICS*, VOL. 37, NO. 5, SEPTEMBER 2001
- [5] Haleh Vafaie, Kenneth De Jong, "Feature Space Transformation Using Genetic Algorithms", *IEEE Intelligent System*, March/April, pp.57 - 65, 1998
- [6] Les M. Howard, Donna J. D'Angelo, "The GA-P : A Genetic Algorithm and Genetic Programming Hybrid", *IEEE Evolutionary programming*, June, 1995