

## 유전자 알고리즘을 이용한 n-tuple 필터링

한우연<sup>o</sup> 남미영 이필규  
인하대학교

{parthia<sup>o</sup>, rera }@im.inha.ac.kr, pkrhee@inha.ac.kr

### N-tuple Filtering Using Genetic Algorithm

Wooyeon Han<sup>o</sup> Miyoung Nam Phillkyu Rhee  
Dept. of Computer Science & Engineering, Inha University

#### 요 약

continuous n-tuple 알고리즘은 tuple의 무작위적 추출을 기본으로 한다. 무작위적 추출의 여러 가지 장점을 감안하더라도, 무작위적 추출을 통한 인식의 성능은 가변성을 가지게 된다. 그리고 무작위적 추출은 의미 있는 정보의 선택이 불가능하다는 단점을 가진다.

본 논문에서는 무작위적 추출이 가지는 여러 가지 약점을 보완하기 위해서, 유전 알고리즘을 이용하여 얼굴 인식에 효과적인 tuple을 선택하여 사용하였다. 유전 알고리즘을 이용함으로써 얼굴 인식에 효과적이지 않은 tuple의 필터링 효과를 기대할 수 있다.

#### 1. 서론

continuous n-tuple 알고리즘[1]은 tuple의 무작위적 추출을 기본으로 하여 만들어 졌다. 이 알고리즘은 무작위적 추출로 인하여 간단한 개념과 빠른 속도, 비교적 효율적인 성능 등의 장점을 갖지만, 그에 따른 몇 가지 약점을 가지게 된다. 그 중 하나가 인식 성능의 가변성이다. 인식이 무작위적으로 추출된 n-tuple에 의해 결정되므로 인식 성능에 대한 신뢰도가 떨어지는 것이다.

본 논문에서는 유전 알고리즘[2,3]을 전처리 필터로 적용하여, 보다 얼굴 인식에 적합한 신뢰할 수 있는 n-tuple을 추출하고 얼굴 인식 성능을 향상시킬 수 있도록 하였다.

#### 2. 배경 지식

##### 2.1 유전 알고리즘

진화연산 알고리즘은 자연계의 생명체 중 환경에 잘 적응한 개체가 좀더 많은 자손을 남길 수 있다는 자연선택 과정과 자연계의 생명체의 설계도와 같은 유전자의 변화를 통해서 좋은 방향으로 발전해 나간다는 자연 진화의 과정을 모방하여 컴퓨터로 모의 수행을 하는 최적화 알고리즘의 하나이다. 즉, 실제계의 문제를 풀기 위해 잠재적인 해들을 컴퓨터상에서 코딩된 개체로 나타내고, 여러 개의 개체들을 모아 개체군을 형성한 뒤, 세대를 거듭하면서 이들의 유전 정보를 서로 교환하거나 새로운 유전 정보를 부여하면서 적자생존의 법칙에 따라 모의 진화를 시킴으로써, 주어진 문제에 대한 최적의 해를 찾는 계산 모델이다.

유전 알고리즘은 하나의 점이 아닌 개체군에 기반하며 적합도 함수를 이용한 맹목적인 검색이다. 그리고 결정

론적 변화규칙이 아닌 확률적 변화 규칙을 이용하는 것도 특징 중의 하나이다.

탐색의 방향이나 영역이 초기값에 의해서 결정되지 않고 세대마다 확률적으로 결정되므로 지역 최소 점에 빠질 가능성이 적어 전역 최적화가 가능하며, 목적 함수의 미분과정이나 특별한 수학적 연산을 필요로 하지 않는 장점을 갖는다. 그에 반해 개체 수, 선택 방법이나 교배 방법의 결정, 돌연변이의 비율 등 파라미터의 수가 많다는 단점이 있다.

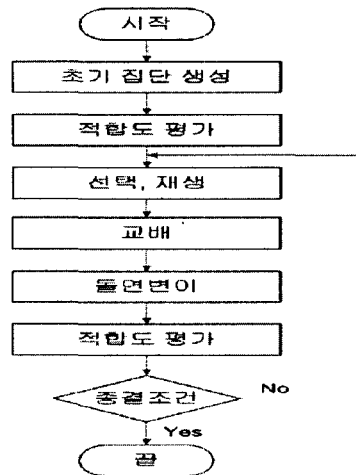


그림 1 유전 알고리즘 순서도

##### 2.2 continuous n-tuple classifier

전통적인 n-tuple classifier[4]는 입력 벡터로 이진 값을 받는다. 이것을 입력이 연속적 이거나 multi-level인 입력에 적당하도록 새롭게 제안한 continuous n-tuple

classifier는 속도와 단순성에 장점을 갖고 있다.

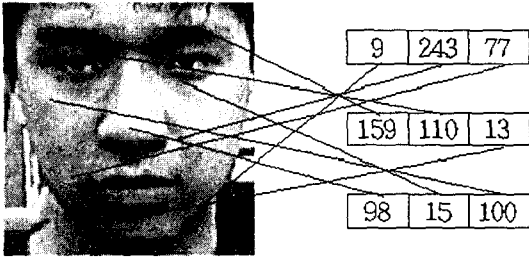


그림 2 continuous n-tuple의 샘플링 과정

그림 2는 d 차원의 입력 공간이 m개의 n-tuple로 샘플링 되는 것을 보여준다. 각 n-tuple은 입력 공간에서 고정된 위치의 집합을 정의한다. j 번째 n-tuple의 위치 집합은

$$n_j = \{a_{j1}, a_{j2}, \dots, a_{jn} \mid 1 \leq a_{ji} \leq d\}$$

학습 과정에서는, 입력 벡터  $\mathbf{x} = x(1) \dots x(d)$ 에 대하여 다음과 같은 투영을 수행하고 저장한다.

$$y_j = x(a_{j1}) \dots x(a_{jn})$$

인식 과정에서는, 각 class마다 j 번째 투영된 벡터에 대해, 각 클래스마다 가장 가까운 저장된 벡터를 찾아서 거리를 더하여 인식 점수를 구한다.

$$r_c = \sum_{j=1}^m \min_k D(y_{jk}^c, z_j)$$

$$D(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n |x_i - y_i|$$

그리고 테스트 영상은  $r_c$  값이 가장 작은 클래스에 할당 되게 된다.

### 3. 유전자 알고리즘을 이용한 n-tuple 필터링

#### 3.1 학습 데이터의 정렬

얼굴영역을 학습데이터로 정하는데 있어서 정렬은 매우 중요한 요소이다. 이에 대해 Henry A. Rowley는 두 눈과 입의 중심점을 기준으로 하여 표준이 되는 평균 좌표를 만든 후 기준 영상을 이에 맞추어 변형하는 방법을 사용하였다.[5]

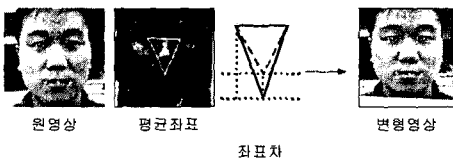


그림 3 영상의 변형

그러나 이 방법은 원 영상의 상하, 좌우 비율에 대한 변

화가 일어나게 되므로 그 신뢰성을 보장할 수 없게 된다. 따라서 본 논문에서는 영상의 정렬과 추출에 대해 다음과 같은 방법을 사용하였다.

1. 양쪽 눈의 높이가 똑같이 되도록 학습데이터의 회전 변환을 수행한다.
2. 양쪽 눈의 중심 좌표간 거리가 모두 똑같이 되도록 학습 데이터의 크기를 가로, 세로 비율을 유지한 상태로 조절한다.
3. 모든 데이터의 눈과 입사이의 중심좌표에 대한 평균을 구한다.
4. 3에서 구한 평균 중심좌표를 기준으로 눈의 거리까지를 반지름으로 하는 원을 그리고, 이 원의 외곽을 따라 영상을 다시 추출해낸다.

#### 3.2 히스토그램 평활화

히스토그램 평활화는 영상의 히스토그램이 평탄하게 되도록 영암값의 재분배를 통해 농도를 조절하는 것이다.

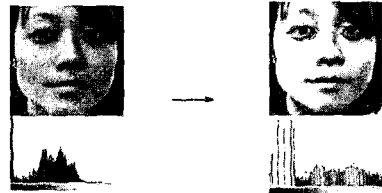


그림 4 히스토그램 평활화

#### 3.3 영색체 설계 및 적합도 평가

유전 알고리즘에서 영색체 설계와 적합도 평가는 성능의 가장 핵심적인 부분을 담당한다.

본 논문에서는 미리 무작위 적으로 추출된 n-tuple의 사용 여부를 결정하는 이진 영색체를 사용하였다.

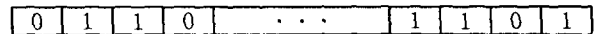


그림 5 영색체 설계

그리고 얼굴 인식에 효과적인 n-tuple을 추출하기 위해서 인식률을 이용하는 적합도 평가 함수를 이용하였다. 인식을 평가에서 테스트 영상에 할당된  $r_c$  값이 가장 작은 클래스가 테스트 영상 본래의 클래스와 같다면 인식률은 증가한다.

#### 4. 실험 결과

본 논문의 실험은 두 가지 과정을 거친다. 하나는 유전자 알고리즘을 이용해 효과적인 n-tuple을 추출해 내는 훈련 과정이고, 다른 하나는 추출되어진 n-tuple이 얼굴 인식에 효과적인지 검증하는 과정이다.

훈련 과정의 실험을 위해서, 128X128로 추출된 연구실

의 100명의 자체 보유 데이터 2190개 중 학습에서 995개, 테스트에서 1195개의 영상을 사용하였다.

그리고 검증 과정의 실험을 위해서, 128X128로 추출된 Feret DB를 이용하였다. 학습과 테스트에는 각각 1092개의 서로 다른 영상을 사용하였다.

유전 알고리즘을 이용한 n-tuple 필터링 방법의 효과를 검증하기 위해서, 무작위적으로 추출한 n-tuple과 유전 알고리즘을 사용한 n-tuple을 이용해서 인식률을 테스트 비교하였다. (n-tuple 알고리즘에서 파라메타 n=3, quantization level=256로 설정해 주었다.) 200개의 n-tuple을 유전 알고리즘으로 필터링 한 결과 161개의 n-tuple을 얻을 수 있었다.

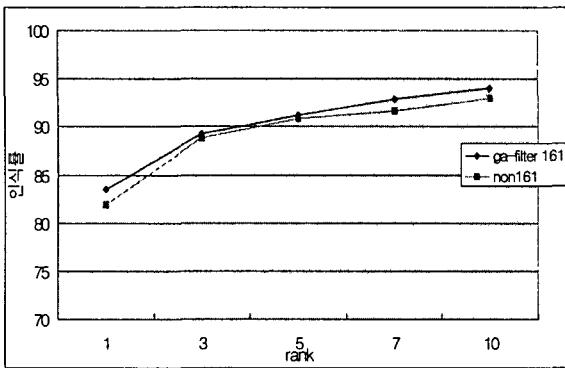


그림 6 유전 알고리즘을 사용했을 때와 하지 않았을 때 161개의 n-tuple 비교

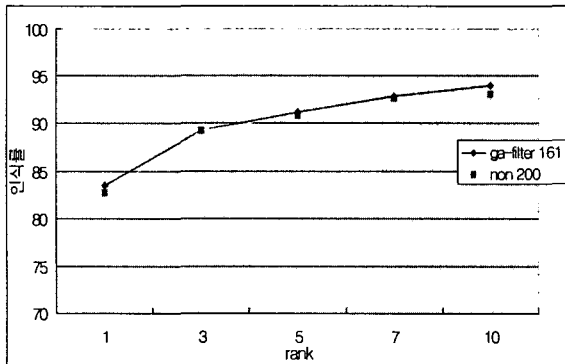


그림 7 200개의 n-tuple을 유전 알고리즘으로 필터링 했을 때와 하지 않았을 때 비교

그림 6은 유전 알고리즘을 사용해서 필터링 한 n-tuple과 무작위적으로 추출된 같은 개수의 n-tuple 비교한 결과이다. 같은 개수의 n-tuple을 사용했을 때 인식률을 비교하면 유전 알고리즘을 사용했을 때의 인식률이 향상됨을 볼 수 있다. 그림 7은 무작위적으로 추출한 n-tuple에 유전 알고리즘 적용했을 때와 그렇지 않을 때를 비교했다. 사용하는 n-tuple의 수가 줄어든다면 인식 성능은 비슷한 것을 볼 수 있다.

## 5. 결론 및 향후 연구 과제

본 논문에서는 n-tuple 알고리즘의 무작위성에 대한 극복 방법을 제안하였다. 실험 결과 유전 알고리즘을 필터로 이용하여 성능이 향상된 것을 확인할 수 있다. n-tuple을 필터링 처리함으로써 인식 성능에 대해 보다 안정적인 신뢰도를 가질 수 있으며, 필터링을 통해 인식률을 떨어뜨리지 않고 더 적은 양의 메모리를 사용해 속도를 증가시킬 수 있다.

n-tuple 알고리즘은 tuple 비교에서 명도값 매칭을 기본으로 하기 때문에 학습 영상 추출이 성능에 큰 영향을 미친다. 따라서 n-tuple 알고리즘의 성능을 최대화시킬 수 있는 학습 영상 추출 방법에 대한 연구가 필요할 것이다.

또한 n-tuple 알고리즘은 여러 개의 학습 영상 중에 가장 유사한 tuple만을 이용하는 방식이기 때문에 학습을 위해 여러 개의 영상이 필요하다. 그렇지만 이번 실험에서 검증을 위해 사용한 Feret DB 같은 경우에는 학습을 위해서 클래스 당 하나의 영상만이 존재한다. 이러한 약점을 극복하기 위해서는 전처리 방법에 대한 추가적인 연구가 모색되어야 할 것이다.

## 6. 참고 문헌

- [1] S.M Lucas, "Real-time face recognition with the continuous n-tuple classifier", Vision, Image and Signal Processing, IEE Proceedings-, Volume: 145, Issue: 5, Oct. 1998
- [2] J. Baker, "Adaptive Selection Methods for Genetic Algorithms", Proc. First ICGA, Jul. pp. 101-111, 1987
- [3] Holland, J., Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor, 1975
- [4] N.M Allinson and A Kolcz, "A Principled Approach to n-tuple Recognition Systems", the Institution of Electrical Engineers, 1997
- [5] Henry A. Rowley, Shumeet Baluja and Takeo Kanade, "Neural Network-Based Face Detection", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.20, No.1, Jan 1998