

# 유전자 알고리즘을 이용한 영상의 특징 및 외곽선 추출

정회원 김 대 극\*, 이 재 곤\*

## Feature and Contour Extraction of an Images using Genetic Algorithms

Dae-Keuk Kim\*, Jae-Kon Lee\* *Regular Members*

### 요 약

본 논문은 유전자 알고리즘을 이용한 영상의 특징 및 외곽선 추출 기법을 제안하고 있다. 영상으로부터 물체의 특징 및 외곽선을 추출하는 기법은 영상 분석 문제를 해결하는데 매우 중요한 방법이다. 현재까지의 연구에서는 외곽선 및 에지를 에너지 상태로 해석하여 목적 함수를 최소화하는 모델들이 제시되었다. 그러나 기존의 방법들은 국부적인 에너지 최소 상태에 빠지기 쉬운 문제점을 가지고 있다.

본 논문에서는 특징 및 외곽선 추출 문제에 대하여 최적의 에너지 상태를 갖는 경로를 탐색하는 문제로 규정하고, 유전자 알고리즘을 적용하여 기존 모델의 단점인 국부적인 최소 상태를 벗어나고자 하였다. 먼저 적합도를 평가하기 위하여 영상 영역간의 인접 에너지 강도와 에지 강도를 이용한 새로운 목적 함수를 정의하였으며, 이 함수를 유전자 알고리즘내의 적합도 와 특징 및 외곽선을 추출하는 과정에서 영역들의 분할 또는 결합의 기준으로 사용하였다. 그리고 유전자 알고리즘의 기본 연산자들은 새로운 경로를 만들어 내도록 수정되었다.

### ABSTRACT

In this paper, new feature and contour extraction technique based on genetic algorithm is presented. In image analysis, feature and contour extraction is an important preprocessing to obtain high level information. In previous researches, contour and edge are regarded as energy states and the solutions were found by minimizing defined object function. But limitation is that previous methods tend to find local optimal position. First of all, we define an objective function based on genetic algorithm for evaluating the fitness function, which is used to measure a degree of separation and compactness within fine-segmented regions and an edge strength along boundaries of all regions. And we modified the simple genetic algorithm in such a way that new paths are created every iterations.

### I. 서 론

유전자 알고리즘(Genetic Algorithms)은 1970년대 미국의 John Holland에 의해 제안된 적응 탐색 기법의 하나로 자연환경에서 종(species)들이 적자 생존의 경쟁을 통해 개체집단(population)을 환경에 적응하도록 진화시켜 나가는 것에 기초하여 제안된 알고리즘이다[1][2]. 유전자 알고리즘은 다른 종류의 탐색 알고리즘과는 달리 방대하고 복잡한 함수에 대하여 전역적인 최적화(global optimization)를 할 수 있는 장점을 가지고 있

기 때문에 최근에 크게 주목을 받고 있으며, 근본적으로 탐색을 위한 알고리즘으로 함수의 최적화(function optimization), 파라미터 조정(parameter tuning), 스케줄링(scheduling), 패턴인식(pattern recognition) 및 기존의 알고리즘으로 해결하기 힘든 여러 분야에 적용되고 있다[3]. 이에 변경 가능한 물체의 특징 및 외곽선을 추출하기 위하여 유전자 알고리즘을 적용하여 기존의 방법들이 갖고 있던 초기 위치의 결정 문제와 국부적인 에너지 최소 상태에 빠지는 문제를 해결하고자 하였다. 본 논문은 유전자 알고리즘을 이용한 영상의 특징 및

\* 한림정보산업대학 전자통신과, 정회원  
논문번호: 98019-0925, 접수일자: 1998년 9월 25일

외곽선 추출 기법을 제안한다. 먼저 영상분할 정도를 평가하기 위하여 영역간의 인접 에너지 강도와 에지 강도를 이용하여 새로운 목적함수를 정의하였으며, 이러한 함수를 유전자 알고리즘내의 적합도와 특징 및 외곽선 추출 과정에서 영역들을 분할할 것인지 혹은 결합할 것인지를 결정하는 기준으로 사용하였다. 또한 실험에 사용될 이미지의 세선화를 위해 유전자 알고리즘을 적용한 필터를 구현하여 잡음을 제거하게 된다. 그리고 임의 탐색 및 동시 처리가 가능한 유전자 알고리즘을 통해 특별한 결정 파라미터 및 분류기 구성을 가능케 함으로써 영상의 특징 및 외곽선을 효과적으로 추출함을 최종 목적으로 한다. 제안된 알고리즘을 실험하기 위해 그레이 영상을 사용하며 영상을 이진코드로 변환하여 입력 값으로 사용한다. 제안된 알고리즘의 실험은 원 영상과 원 영상에 잡음을 첨가시킨 영상에 대해 기존의 알고리즘과 비교 분석함으로써 효과적인 특징 및 외곽선 추출 과정을 보인다. 본 논문의 주된 목적은 언급한 바와 같이 특징 및 외곽선 추출 시스템을 구현하기 위해 유전자 알고리즘을 전처리 단계에 적용함으로써 효과적인 추출을 행하며 나아가서 기존에 사용된 알고리즘과 비교 분석을 통해 더 나은 알고리즘 개발 향상에 기여하는데 있다. 본 논문은 II장에서 유전자 알고리즘에 대하여 언급하고, III장에서는 기존의 특징 및 외곽선 추출 방법들과 제안된 유전자 알고리즘을 이용한 특징 및 외곽선 추출 과정에 관하여 설명한다. IV에서 유전자 알고리즘을 이용한 실험결과를 언급하고 결론을 맺는다.

## II. 유전자 알고리즘(Genetic Algorithms)

유전자 알고리즘은 자연선택과 자연계 진화과정의 적자생존을 원칙으로 하는 적응적인 탐색 기법으로서 최적해에 대한 효과적인 탐색 도구로 널리 사용되고 있다. 이것은 다른 탐색방법 및 최적화 방법과 다른 점은 파라미터를 코딩한 것을 직접 이용하며, 점(point)이 아닌 다점(multi points) 탐색 방법을 취하는 것이다. 또한 탐색에 적합도를 이용하고 다른 부가적인 지식을 요구하지 않으며 결정론적인 규칙이 없이 확률적 연산자를 사용하여 수행된다. 이와 같은 특징으로 인해 다른 탐색 또는 최적화 방법 중 하나인 계산에 의존한 방법(calculus-based method)에 비하여 전역적 해를 구할 가능성이 높으며 다른 여러 탐색방법에 비하여 효율적이다. 유전자 알고리즘을 실제 응용에서 사용하기 위해 먼저 문제를 유전자형에 대응하는 문자열로 변환한다. 그리고 이진수 문자열의 나열인

염색체를 표현하고, 이 염색체들의 모임인 개체집단을 생성한다. 문자열을 평가하여 평가치가 높은 집단을 선택하여 남도록 하는데, 이것은 자연계에 있어서 도태에 대응하는 것이다. 이렇게 선택된 집단에 대하여 연산자를 적용함으로써 새로운 문자열을 생성한다. 기본적인 연산자 재생은 문자열을 적합도에 따라 개체 집단에서 두 개의 염색체를 선택하는 역할을 하는데 개체집단내의 모든 염색체의 적합도의 합과 각 염색체의 적합도의 상대적 값, 즉 선택될 확률을 구하게 된다. 교차 연산자는 재생에 의해 선택된 두 염색체의 인자값을 서로 맞바꾸어 새로운 염색체를 생성한다. 이때 무작위로 선택된 교차 위치와 개수에 따라 교차 알고리즘이 달라지게 된다. 또한 돌연변이 연산자는 염색체내의 인자를 무작위로 선택하고, 그 값을 임의대로 바꾸어 새로운 염색체를 만드는 기능을 담당한다. 이것은 개체 집단의 특성을 다양하게 변화시키며 국부적인 최소 상태를 벗어 날 수 있게 한다. 위에서 설명한 연산자의 사이클을 반복함으로써 환경에 대응하는 평가치가 높은 문자열을 만들어 내어 문자열의 집단 전체의 평가치를 향상 시켜 나간다. GA 연산자를 이용하여 새로운 세대를 생성하였을 때 부모 세대보다 더 낮은 적합도를 갖는 염색체가 생성될 수도 있는데, 세대를 거듭하면서 적합도가 낮은 것들은 염색체끼리의 경쟁에서 도태되어 사라지게 되므로 낮은 적합도를 갖는 염색체가 다음 세대에 생기더라도 문제

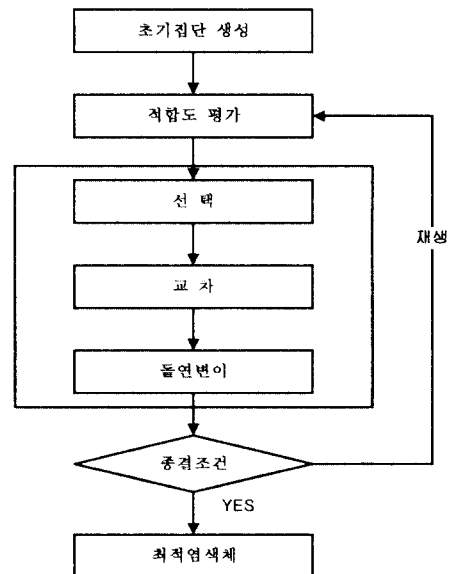


그림 1. 단순 유전 알고리즘 및 흐름도

가 되지 않는다. 기본적인 단순 유전 알고리즘(Simple Genetic Algorithm : SGA)의 흐름도는 다음과 같다[4].

### III. 유전자 알고리즘을 이용한 특징 및 외곽선 추출

유전자 알고리즘을 이용하여 영상의 특징 및 외곽선을 추출하기 위해 제안된 방법은, 찾고자 하는 특징 및 외곽선을 최적 상태의 에너지 값을 갖는 하나의 경로로 정의하고 기본적인 유전자 알고리즘을 수정하여 최적의 경로를 탐색한다. 이 방법으로 기존의 특징 및 외곽선 추출 모델이 갖고 있는 초기 위치의 결정 문제와 국부적인 에너지 최소 상태에 빠지는 문제를 개선하는데 목적이 있다.

#### 3.1 특징 및 외곽선 추출을 위한 전처리 과정

수정된 알고리즘에서 영상의 외곽선 에너지와 예지 강도를 이용하기 위해 추출의 전 단계로 양자화와 잡음 제거를 한다. 양자화는 영상 입력의 첫 단계로 영상을 스캐너 및 디지털 카메라를 사용하여 입력 데이터로 변환한다. 입력된 영상은 어느 정도의 잡음이 포함되는 경우가 많으므로 잡음이 들어간 영상은 인식의 전처리로서 잡음을 제거해 주어야 한다. 일반적으로 잡음은 고립된 한 점으로 나타나므로 고립점을 제거하여 잡음을 없애는 처리과정이 필요하다. 즉, 한 화소 주위의 8근방의 화소값이 모두 같으면 중심값도 주변 화소값으로 대체한다.

#### 3.2 유전자 알고리즘

기존의 특징 및 외곽선 추출 방법은 모델을 기초로 하는 정적인 것으로 예지 영상에 대해 물체의 특성을 나타내는 특징치를 정의하고 template 정합을 이용하는 방법[5][6]과 모델을 수학적으로 정확히 기술하여 위치를 찾아내는 Hough 변환[7]을 이용하는 방법이 많이 이용되었다. 이 방법은 모델의 형태가 위치와 시간의 흐름에 따라 변화하거나 물체의 밝기가 변화할 경우에는 특징 및 외곽선 추출에 사용할 수 없는 단점이 있다. 또한 이들 방법들은 에너지 함수를 정의한 다음 수학적 방법을 적용하여 이를 최소화하는 장점이 있으나 내부 에너지의 매개변수를 영상의 종류에 따라 경험적으로 정해야 하는 점과 이전 정보로부터 새로운 정보를 얻기 위해 바로 인접한 영역의 정보를 참조하게 되므로 초기 위치가 잘못 선정되면 에너지의 최소화 방향을 찾을 수 있을 만큼 외곽선 근처에 이웃하는 잡음에 의한 영향으로 국부적인 에

너지 최소 상태에 빠지는 문제를 지니고 있다.

본 논문에서는 영상의 특징 및 외곽선을 추출하기 위해 기존의 모델들의 단점을 보완한 수정된 유전자 알고리즘의 적용을 제안하고 있다. 각 염색체들은 하나의 개체집단을 형성하고, 이 개체집단은 진화 조건에 따라 새로운 개체집단을 형성한다. 문제의 해를 찾기 위해 수정된 유전 연산자를 사용하여 목적함수가 최대가 되는 염색체를 찾는다

##### 3.2.1 유전 연산자

선택 연산자는 개체집단의 적합도를 비교하고 적합도 이하의 개체는 버리며 적합도가 높은 개체를 복제하는 방식으로 우수한 개체만을 남긴다. 적합도에 의해 선택하는 방법도 경쟁자를 설정하는 방법에 따라 다양하다. 본 논문에서는 높은 에너지 값을 가질 수록 다음 세대에 살아 남을 확률이 높아지게 하는 Roulette Wheel 선택 방법을 이용한다. 이 방법은 에너지  $f$ 의 새로운 에너지  $f'$ 는 식(1)과 같이 정의된다.

$$f' = f - (f_{avg} - c \cdot \sigma) \quad (1)$$

$f$ 에 대한 표준편차  $\sigma$ 는  $f$  값의 변화량을 줄여 조기 수렴을 예방하기 위해 사용된다. 평균 에너지  $f_{avg}$ 와 표준편차  $\sigma$ 는 식(2),(3)과 같다.

$$f_{avg} = \sum_{i=0}^{N-1} f_i / N \quad (2)$$

$$\sigma^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} (f_i^2 - f_{avg}^2) \quad (3)$$

음수가 되는 에너지 값은  $f'$ 를 0으로 하여 다음 세대에 도태시킨다. 수정된 알고리즘의 교차는 염색체를 구성하는 점의 수가 일정하지 않으므로 선택 과정을 거친 두 염색체를 선택한 후 겹쳐지는 두 점 사이의 세그먼트를 서로 교환하여 새로운 염색체를 형성한다. 두 염색체가 여러 곳에서 겹치게 되면 각 세그먼트의 에너지 값을 비교하여 에너지 값이 낮은 쪽은 도태 시켜 버린다. 또한 돌연변이는 국부적인 최소 상태를 벗어나기 위한 중요한 연산자로 외곽선을 추출하기 위해 염색체를 구성하는 각 좌표 값을 새로운 위치로 이동시킨다.

##### 3.2.2 예지 강도 측정 및 적합도 함수

예지의 정도는 픽셀의 밝기 정보의 기울기(gradient)

정도를 이용한다. 그러나 그 크기는 정량적으로 정의할 수 없는 퍼지 성질을 내포하고 있다. 픽셀의 위치가  $(i, j)$  일때의 기울기의 크기는 식(4)와 같이 정의된다[8].

$$|\nabla G(i, j)| = \left[ \left( \frac{\rho G(i, j)}{\rho x} \right)^2 + \left( \frac{\rho G(i, j)}{\rho y} \right)^2 \right]^{1/2} \quad (4)$$

기울기 크기를 이용한 영상 전체의 상대적 에지 정도 E는 식(5)와 같다.

$$E = \sum_{(i, j)} \frac{\mu_g(|\nabla G|)}{\mu_g(|\nabla G'|)} \quad (0 \leq E \leq 1, \quad |\nabla G| \leq |\nabla G'|) \quad (5)$$

식(5)는 외곽선 추출시 에지 정도가 가장 낮은, 즉 기울기 크기가 적은 에지를 갖는 영역이 추출 되도록 하는 역할 담당하게 된다. 또한 에너지는 경사로 상의 모든 점들에 대해서 계산되며 한 점의 Gradient 크기의 제곱에 비례한다. 에너지를 계산 할 때 경로 위에 존재하는 모든 점들이 에너지 계산에 같은 정도로 기여하게 되므로 한 점의 Gradient 값이 지장된 임계치 보다 낮을 경우, 즉 외곽선과 거리가 먼 위치에 있는 점들에 대해서는 염색체의 에너지 기여도가 낮게 된다. 이것은 에지들이 연결되어 있지 않은 경우, 가장 길이가 짧은 곳을 외곽선의 위치로 취하게 되면 임계치 보다 낮은 Gradient 값을 갖는 점들의 수가 많아지기 때문에 전체 에너지를 떨어뜨리게 된다. 영상에서의 한 화소의 에너지  $E_g$ 는 식(6)에 의해 얻을 수 있다.

$$E_g(r, c) = (\nabla I(r, c) * G_\sigma)^2 \quad (6)$$

잡음을 줄이기 위한 평활화의 정도는  $\sigma$ 에 의해 조절되며, 염색체의 에너지는 경로를 구성하는 각 점의 에너지  $E_g$ 를 합한 다음 점의 개수로 나누어 정규화된다. 또한 유전자 알고리즘에서 문제의 해는 적합도에 의해 결정된다. 본 논문에서는 영상의 특징 및 외곽선을 추출하기 위한 적합도는 영상의 외곽선 에너지와 에지 강도에 의해 결정된다. 적합도 함수는  $F = G \cdot E$ 로 정의된다. 최적의 특징 및 외곽선 추출은 적합도를 최대로 하는 영역들의 구성으로 정의되며 유전자 알고리즘에서는 최대 적합도를 갖도록 진화 과정을 통해 최적 염색체 구조를 찾는다.

#### IV. 실험 결과 및 분석

모의 실험에 앞서 유전자 알고리즘을 이용할 때 임의성이 내부에 중요한 역할을 하기 때문에 난수값에 따른 초기값을 이용한 실험에 있어 다른 세부적인 현상이 발생된다. 따라서 개체 집단의 크기와 교배 및 돌연변이 확률은 알고리즘에 영향을 주게된다. 본 실험에서 사용된 유전자 알고리즘의 파라미터를 표 1에 나타내었다.

표 1. 유전자 알고리즘에서의 초기 파라미터

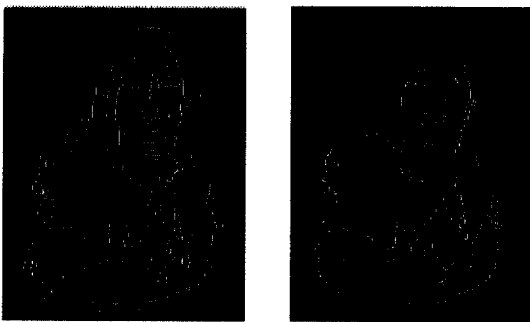
GA Operator	초기 설정값
Chromosome Length	10
Population Size	20
Number of Generations	10
Mutation Rate	0.100
Crossover Rate	0.950
Crossover Type	Single
Elitist Operator	On
Function	Simple
Gray Coding	On

제안된 알고리즘을 이용한 영상의 외곽선 추출에 대한 성능평가를 위하여 256 그레이 레벨의 256 \* 256 크기의 영상을 사용하였다. 영상에 대한 외곽선 추출 시험은 인텔 펜티엄 프로 233MHz상에서 Matlab 4.0으로 알고리즘을 구현하여 실험하였다. 그림 2(a)는 실험에 사용된 원 영상이며, 그림 2(b)는 원영상에 가우시안 잡음을 첨가한 영상이다. 이 두 개의 영상에 대한 제안된 방법과 양자화와 잡음제거를 거쳐 sobel 연산자를 이용했을 때의 결과를 비교하였다. 그림 3(a)과 그림 4(a)는 원영상과 가우시안 잡음이 포함된 영상에 대하여 유전자 알고리즘을 통해 얻은 외곽선 추출 영상이며, 그림 3(b)와 그림 4(b)는 기존 추출 방법인 sobel 연산자를 이용하였을 때의 결과이다. 그림에서 볼 수 있듯이 원영상과 잡음이 첨가된 영상에 대해 기존의 방법을 사용하였을 때의 결과 보다 제안된 방법이 효과적인 특징과 외곽선 추출이 되었음을 볼 수 있다. 이러한 결과는 3 \* 3 마스크로 sobel 연산자로 이용된 방법은 이웃화소의 그레이 레벨값과 비교 연산하게 될 때 이웃화소들이 별 차이를 가지고 있지 않기 때문에 원영상 및 잡음이 첨가된 영상에서 효과적인 외곽선을 추출 할 수 없었음을 알 수 있었



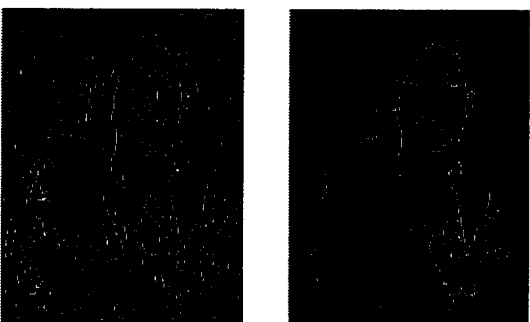
(a) 원영상 (b) 잡음이 첨가된 영상

그림 2. 실험 영상



(a) 유전자 알고리즘을 이용한 추출 (b)소벨(sovel) 연산자 모델을 이용한 추출

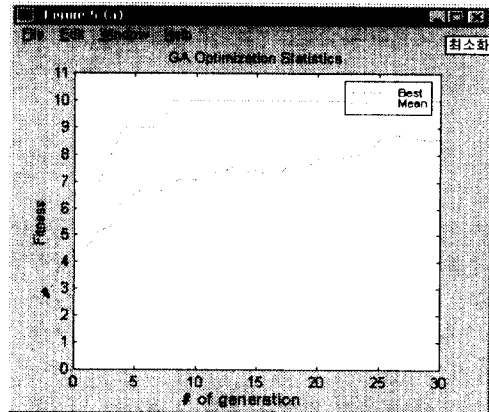
그림 3. 원 영상에 대한 각 방법에 따른 외곽선 추출



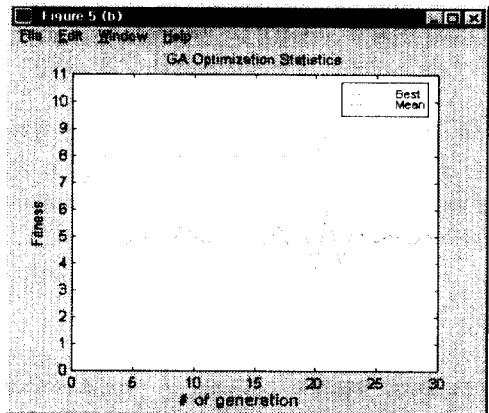
(a) 유전자 알고리즘을 이용한 추출 (b)소벨(sovel) 연산자 모델을 이용한 추출

그림 4. 잡음이 첨가된 영상에 대한 외곽선 추출

다. 그러나 유전자 알고리즘은 특징인 교차 및 돌연변이로 인하여 초기 결정 문제나 국부적인 에너지 최소 상태에서 벗어날 수 있어서 더 효과적인 외곽선



(a) 초기영상



(b) 가우시안 잡음이 첨가된 영상

그림 5. 각 실험 영상에 대한 mean과 fitness 분포

추출 성능을 보이고 있다. 그림 5는 두 영상에 대한 세대 변화에 따른 mean과 fitness의 분포를 나타내었다. 그림 5(a)와 (b)에서 볼 수 있듯이 영상에 대한 잡음의 첨가가 적을수록 세대가 지남에 따라 에너지 값이 빠른 속도로 증가함을 알 수 있었다. 마지막으로 유전자 알고리즘의 성능은 진술한 바와 같이 파라미터에 의해 영향을 받는다. 교차와 돌연변이는 수렴속도를 결정하게 된다. 특히 수행 결과에 대한 신뢰도는 개체 집단 내의 개체수와 진화정도에 따라 달라지게 된다. 영상의 특징 및 외곽선 추출에 있어 교차 및 돌연변이는 영역들의 결합 및 분할정도를 결정한다. 만약 결합의 빈도가 높게되면, 영상의 외곽선 추출에 적합하도록 최종 영역의 수가 줄어들게되며 정교하거나

잡음이 첨가된 영상의 외곽선을 추출하기 위해서는 분할 빈도수를 증가시켜 예기치 않는 영역의 결합을 예방 할 수 있다.

### V. 결 론

본 논문에서는 영상에서 효율적인 특징 및 외곽선을 추출하기 위해 유전자 알고리즘을 이용한 외곽선을 에너지 값이 가장 높은 상태로 정의하였고, 기존의 모델이 가지고 있던 잡음으로 인한 초기 위치 결정 문제와 국부적인 에너지 최소 상태를 탈출 하고자 하였다. 실험에서 초기 영상과 잡음이 첨가된 영상의 특징 및 외곽선 추출을 비교함으로써, 유전자 알고리즘이 초기 위치 및 국부적인 에너지 최소 상태를 벗어날 수 있음을 보였다.

제안된 유전자 알고리즘은 실험에서 볼 수 있듯이 초기 위치에 대하여 효과적인 외곽선 추출 성능을 보인다. 그러나 속도 저하의 문제점을 안고 있어 앞으로 병렬처리가 필요하리라 본다. 또한 영상에 잡음이 많이 첨가될 경우 외곽선을 찾기 어려워지므로 확실적인 변화에 효과적인 추출을 가능하게 하는 수정 연산자의 개발이 필요하겠다.

### 참 고 문 헌

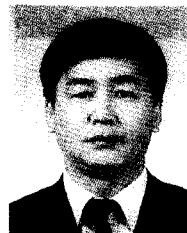
1. Aha, D.W, "Machine Learning," A tutorial presented at the 5th International Workshop on Artificial Intelligence & Statistics, p1-67, 1995.
2. Goldberg, D. E., "Genetic Algorithms in Search, Optimizations & Machine Learning," Addison Wesley, 1989.
3. C. A. Ankenbrandt, B. P. Buckles and F.E.Petry, "Scence recognition using genetic algorithms with semantic nets," Pattern Recognition Letters, VOL. 11, pp.285-293, Apr. 1990.
4. Lawrence Davis, Handbook of genetic algorithms, Van Nostrand Reinhold, New York, 1991.
5. S.A Dudani, K. J. Breeding and R. B. McGhee, "Aircraft identification by moment invariants," IEEE Trans. Comput. C-26, pp. 39-46, 1977.
6. Sidhartha Maitra, "Moment Invariants," Proceedings of the IEEE, VOL. 67, No. 4, pp. 697-699, 1979.
7. R.O Duda and P.E Hart, "Use of the Hough Transforms to detect Lines and Curves in Picture,"

Communications of the ACM, Vol. pp.11-15, 1972.

8. William K. Pratt. Digital image processing, Jone Wiley and Sons. Inc, New York, 1991.



**김 대 극 (Dae-Keuk Kim)** 정회원  
 1966년 7월 25일생  
 1992년 : 강원대학교 전자공학과 졸업(공학사)  
 1994년 : 강원대학교 대학원 전자공학과 졸업(공학석사)  
 1997년 : 강원대학교 대학원 전자공학과(박사과정 수료)  
 1997년~현재 : 한림정보산업대학 전자통신과 전임강사  
 <연구분야> 신경회로망, 음성신호처리, 패턴인식



**이 재 곤 (Jae-Kon Lee)** 정회원  
 1951년 9월 19일생  
 1985년 : 숭실대학교 대학원 전자공학과 졸업(공학석사)  
 1996년 : 강원대학교 대학원 전자공학과(박사과정 수료)  
 1989년~현재 : 한림정보산업대학 전자통신과 부교수  
 <연구분야> 음성신호처리, 신경회로망