論文97-34S-8-11

적응 군집화 기법과 유전 알고리즘을 이용한 영상 영역화

(Image Segmentation Using Adaptive Clustering Algorithm and Genetic Algorithm)

河 星 旭 *, 姜 大 星 *

(Seong Wook Ha and Dae-Seong Kang)

요 약

본 논문에서는 적응 군집화 알고리즘(ACA: adaptive clustering algorithm)과 유전 알고리즘(GA: genetic algorithm)을 사용하는 새로운 영상 영역화 기법을 제안한다. 일반적인 GA는 확률적인 재구성으로 인하여 해가 유동적일 수 있으며, 유전자 집단(population) 개체의 개수 부족으로 인하여 최적해에 도달하기 전에 조기 포화 현상이 나타난다. 이들의 문제점을 해결하고 수행 시간 단축을 위하여 새로운 최상 선택 재구성 방법(GBR: great best reinsertion)을 제안하고, 이때 일어날 수 있는 조기 포화 문제를 해결하는 방법을 제시한다. GBR은 자식 저장소에 있는 개체의 적합도가 부모 개체의 적합도보다 좋은 개체를 선택하며, 조기 포화 문제는 부모 집단에서 돌연변이를 발생시키는 방법으로 해결한다. 또한, 영역화 결과에서 발생할 수 있는 소영역을 제거하기 위한 새로운 기법을 제시한다. 실험 결과는 가우시안 잡음 환경하에서 제안된 알고리즘이 ACA보다 더 우수한 영역화 결과를 보였다.

Abstract

This paper proposes a new gray-level image segmentation method using GA(genetic algorithm) and an ACA(adaptive clustering algorithm). The solution in the general GA can be moving because of stochastic reinsertion, and suffer from the premature convergence problem owing to deficiency of individuals before finding the optimal solution. To cope with these problems and to reduce processing time, we propose the new GBR algorithm and the technique that resolves the premature convergence problem. GBR selects the individual in the child pool that has the fitness value superior to that of the individual in the parents pool. We resolve the premature convergence problem with producing the mutation in the parents population, and propose the new method that removes the small regions in the segmented results. The experimental results show that the proposed segmentation algorithm gives better performance than the ACA ones in Gaussian noise environments.

I. 서 론

유전 알고리즘(GA)은 1970년대에 정립된 이론으로 자연의 선택(natural selection)과 자연 유전학(natural genetics)의 원리인 재생산(reproduction), 교배(recombination or crossover), 돌연변이(mutation)에 근거하는 최적해를 찾는 탐색법이다. 기존의 알고리즘은 국부 탐색(local search)을 이용하는데 반해 GA는 여러 집단을 이용하여 해를 찾는 전역 탐색(global search)을 이용하므로서 최적해를 찾을 확률이 기존의 탐색 알고리즘보다 높다. GA의 성능은 해의 부호화 기법(encoding method), 재생산, 교배, 돌연변이, 재구성법(reinsertion method)에 해당하는 각각 알고리즘 중 어느 것을 선택하는 가에 따라 좌우된

* 正會員, 東亞大學校 電子工學科
(Dept. of Elec. Engineering Dong-A Univ.)

接受日字: 1997年3月3日, 수정완료일: 1997年7月23日

다. 본 논문에서는 GA의 최적해를 찾는 능력을 영상 영역화에 응용하여 새로운 기법을 유도하고자 한다. 영상 영역화(image segmentation)는 영상처리 시스템에서 영상 정보 추출의 중요한 첫 번째 단계이며, 하나의 영상을 비슷한 특징을 갖는 여러 개의 영역들로 나누는 것이다. 영역화된 영역은 명암도(gray level)와 같은 특징에 대하여 동질적(homogeneous)이며, 균일하여야 하고, 그 이웃하는 영역들은 전혀 다른 값들을 가져야 한다. 영상 압축이나 패턴 인식 시스템의 성능 향상을 위하여 이들 시스템의 전 처리 단계인 영역화가 성공적으로 수행되어야 하는 선행 조건이 있으며, 영역화의 성패가 전체 시스템의 효율을 결정짓는다고 말할 수 있다. 많은 영역화 기법들이 이미 발표되었고, 히스토그램 이치화(histogram thresholding), 에지 검출(edge detection), 영역 추출(region extraction)[1]-[5] 과 같은 목적들을 달성하기 위해 사용되어 진다. 이들을 두 가지 형태로 분류하면, 첫 번째는 유사성(similarity) 개념에 바탕을 두고 있으며, 또 다른 형태는 불연속성(discontinuity) 개념에 그 기조를 두고 있다. 히스토그램 이치화 기법은 유사성을 이용하는 반면 에지 검출과 영역 추출 기법들은 불연속성을 이용한다. 기존의 영역화 기법들은 어느 정도 경험에 의존하여 영역의 개수와 파라메터를 설정하므로 인하여 다양한 종류의 영상들에는 적합하지 못하다. 그러므로, 입력되는 영상의 종류에 따라 자율적으로 적응하여 영역화하는 알고리즘의 개발이 필요하다. 영상 영역화 방법중 본 논문의 기본이 된 ACA의 학습 방법에서는 어떤 입력 벡터가 이미 존재하는 군집단(cluster)의 참조 모델에 충분히 유사성을 갖는다면 그 군집단에 포함시킨 후 연결강도를 수정하는 알고리즘으로 구성이 간단한 장점이 있다. 그리고, ACA는 영상에 관계없이 강도를 자율적으로 적응하는 알고리즘으로 본 연구에서는 자율적응의 문제인 경계값의 최적치를 찾기 위한 알고리즘을 제안하고자 한다. 입력 벡터가 이미 존재하는 군집단의 참조 모델에 대하여 충분히 유사성(similarity)을 갖지 못하면 새로운 군집단을 생성하고 그 군집단에 포함시킨다. 모든 입력 벡터에 대하여 위의 과정을 반복하면 군집단의 개수는 무수히 증가하게 된다. 이는 경계값에 의하여 결정되며, 학습 초기부터 군집단의 개수를 예상할 수 없는 단점을 갖고 있으며, 만일 매우 많은 입력 벡터들을 취급하는 경우에는 시스템 자원의 고갈을 야기할 수 있다. 두 번째, 경계값의 문제점은 충분한 개수의 입력 벡터들이 입력되기 전의 벡터들은 잘못 분류될 수 있으며 이는 벡터들이 입력됨에 따른 연결강도의 이동에 기인한 것이다. 세 번째 문제점은 벡터들의 입력 순서에 의한 영향으로서 순서를 달리할 때 그 부류의 결과는 다르게 나타난다. 이러한 특성으로 잡음이 첨가된 영상의 영역화 결과에서 군집단의 개수가 기하급수적으로 늘어나며, 강도 벡터의 수정이 일정 횟수 이전에는 영상 영역화가 좋지 않은 문제점이 있다. 본 연구에서는 GA를 사용하여 최적화된 경계값으로 간단한 ACA를 사용하여 영상을 영역화하였다. ACA에 내재하는 문제중 첫 번째인 경계값 설정 문제는 GA를 사용하여 최적화된 경계값을 구하여 해결될 수 있으며, 두 번째 문제로 강도 벡터가 입력에 대하여 계속 수정되는 단점은 GA의 반복과정에서 나온 최적 경계값에 의해 고정된 강도 벡터를 사용하여 해결하였다. 잡음 영상의 영역화 이후에 잘못 분류된 소영역이 나타날 수 있다. 여러 영상 영역화 알고리즘에서는 잡음에 대한 문제를 해결하기 위해 잡음제거용 필터를 사용하여 전처리 과정으로 필터링을 수행한다. 그러나, 필터를 사용할 경우, 잡음에 의해 원영상에 포함되었던 특징들이 소멸될 수 있다. 본 연구에서는 필터링 대신에 영역화 이후에 잡음에 의해 나타날 수 있는 소영역들을 소영역 제거법으로 제거하였다. 원영상에서의 소영역은 영역화에서 불필요한 정보를 가지므로, 영역화의 효율을 높이기 위해 제거된다. 소영역을 제거하기 위해서도 영역화의 결과가 최적이 되는 것이 우선이다. ACA는 정밀하지 못한 경계값으로 불필요한 영역들이 나타나 소영역 제거후에도 영역화가 잘 이루어지지 못하나, 제안한 알고리즘은 정밀한 최적 경계값으로 불필요한 영역이 나타나지 않으므로 소영역 제거 효과가 높다. 제안된 알고리즘에 사용된 GA는, 기존 알고리즘에서 재구성 방법을 달리 하였으며, 이때에 발생하는 조기 포화 문제(premature convergence)를 해결하는 방법을 제안하고자 한다. 제안한 알고리즘의 최적 경계값을 구하는 방법에 있어서 개체군의 개수와 유사성 중 어느 인자에 중점을 두고 비교하느냐에 따라서 영역화 결과가 달리 나오며, 인자에 따라 여러 시스템에 응용될 수 있다. 본 논문의 구성은 기본적인 유전 알고리즘을 Ⅱ절에서 설명하고, 간단한 영역화 기법인 ACA에 대해서는 Ⅲ에서 설명하며, 본 논문에서 제안하는 알고리즘은 Ⅳ에 설명한다. Ⅴ절에서는

영역화에 사용되는 필터링 대신에 소영역을 제거하는 방법이 설명되며, 알고리즘에 대한 실험은 VI절에 설명한다.

Ⅱ. 유전 알고리즘(Genetic Algorithm)

유전 알고리즘(GA)[7]-[16]은 자연 생태계의 유전학과 자연 선택 원리에 의존하는 확률적 탐색(stochastic search) 방법으로 J. Holland에 의해 개발된 기법이다. GA는 최근 여러 분야에서 최적해를 구하는데 많이 사용되고 있다. GA는 재생산(reproduction), 교배(crossover), 그리고 돌연변이(mutation) 등의 유전 연산자(genetic operators)와 적합도 평가 함수(fitness evaluation function)로 구성이 되며, 그 과정은 다음과 같다

1) 부호화(encoding) 및 초기화(initial) 단계

주어진 문제의 해를 유전자의 염색체(chromosome) 구조와 같이, 유전 연산자에 사용할 수 있도록 부호화한다. 부호화는 문제에서 요구하는 해의 특성에 따라 이진(binary) 값과 실수 값을 가질 수 있으며, 이들 값에 따라 유전 연산자가 달라지게 된다. 이때, 초기 개체군(parent population or initial population)은 무작위(random)로 구성한다.

2) 적합도 평가(fitness evaluation) 단계

구성된 개체군이 주어진 문제에 얼마나 적합한지를 평가하고, 다음 세대에 살아남을 선택확률을 정하기 위해, 초기 개체군을 적합도 함수의 인자로 하여 문제에 대한 적합도를 구한다.

3) 재생산(reproduction) 단계

재생산 단계는 구해진 선택확률이나 개체의 인자 값으로 각 개체를 다음 세대 개체로 재생산하는 과정으로, 개체들의 적합도나 인자값을 기준으로 하여 새로운 집단을 구성하고, 구성된 집단을 교배 저장소(mating pool)에 저장한다. 재생산의 방법에는 정렬(sorting)된 개체를 순위대로 선택하는 순위적 선택법(rank-based fitness assignment), 무작위로 선택된 수를 개체의 적합도 비율로 사상(mapping)시키는 룰렛 휠 선택법(roulette wheel selection), 일정 간격의 무작위 수로 개체를 선택하는 일반 확률적 추출(stochastic universal sampling), 교배될 개체의 이웃(neighborhood)을 선택하는 국부적 선택법(local selection), 일정 개수 이상만 선택하는 절단적 선택법(truncation selection), 토너먼트 경기 방식으로 무작위로 구성된 개체군에서 좋은 적합도의 개체를 선택하는 토너먼트 선택법(tournament selection)등의 여러 가지 방법이 있다.

4) 교배(recombination) 단계

교배 저장소에서 각각 교배할 두개의 개체를 선택한다. 이진 값의 교배 방법으로는, 주로 교배되는 하나의 위치를 무작위로 선택하여 교배하는 단일 점 교배법(single-point crossover)이 많이 쓰이며, 이외에 여러 위치를 무작위로 선택하여 교배하는 다중 점 교배법(multi-point crossover), 개체와 같은 수의 비트를 무작위로 발생시켜 0비트 부분은 다른 부모(other parent)의 비트로 바꾸는 일정 배율 교배법(uniform crossover), 각 비트마다 다른 부모의 비트를 무작위로 섞는 뒤섞임 교배법(shuffle crossover)등이 있다. 실수 값의 교배 방법에는 개체의 인자를 무작위 선택한 이산 값에 따라 부모의 인자를 복재(reproduction)하는 이산 교배법(discrete recombination), 부모의 개체에서 일정 계수 값만큼 증가·감소하는 중재 교배법(intermediate recombination), 선형적인 값으로 최적해에 접근해 가는 선형 교배법(line recombination), 선형 교배법을 확장시킨 확장 선형 교배법(extended line recombination)등의 방법이 있다. 교배된 개체를 새로운 개체군으로 구성하여 자식 저장소(child pool)에 저장한다.

5) 돌연변이(mutation) 단계

구성된 개체군을 일정한 돌연변이의 확률로 자식 저장소에서 무작위로 선택한 후, 개체가 이진 값으로 구성된 경우는 임의의 비트를 역변환(0은1, 1은0으로 만드는 과정)하고, 실수 값인 경우는 돌연변이 함수로 개체를 변환시킨다.

6) 재구성(reinsertion) 단계

초기 개체군을 재구성하는 방법은 유전 연산자로 생성된 자식 저장소의 모든 개체를 다시 초기 개체군으로 구성하는 순수 재구성법(pure reinsertion) 외에, 부모보다 적은 자식을 생성하고 무작위로 부모와 교체시키는 일정 배율 재구성법(uniform reinsertion), 마찬가지로 부모보다 적은 자식을 생성시키고 적합도가 나쁜 부모와 교체하는 엘리트 우선 재구성법(elitist reinsertion), 재구성에 필요한 개체보다 많은 자식을 생성하고 적합도가 가장 좋은 개체만을 선택하는 적합도 의존 재구성법(fitness-based reinsertion)등이 있다.

7) 반복 단계

요구하는 식의 최적해를 구할 때까지 2) - 6)을 반복한다.

Ⅲ. ACA의 학습 방법

ACA는 연속 값을 갖는 입력 벡터를 처리하기 위해 제안되었으며, 승리 대표 벡터를 선정하는 기준은 최소 거리 측도(minimum distance measure)에 의존한다. 자율 적응 분류기의 기본적인 학습 방법[6]은 다음과 같다.

1) 새로운 입력 벡터 ($\overline{x_i}$)가 제공되면, 연결강도 벡터 ($\overline{w_j}$)와 입력 벡터 사이에 대하여 왜곡 $d(\overline{x_i}, \overline{w_j}) = \|\overline{x_i} - \overline{w_j}\|$이 최소가 되는 승리 대표 벡터 j^*를 구한다.

2) 다음의 수식이 만족되면 부류 j^*는 경계값 검사를 통과하게 된다:

$$\|\overline{x_i} - \overline{w}_{j^*}\| < \rho$$

여기서, 경계값 ρ는 어떤 군집단의 반경을 나타낸다.

3) 승리 대표 벡터가 경계값 검사에서 실패되면, 새로운 군집단 k는 $\overline{w}_k = \overline{x_i}$로 생성된다.

4) 승리 대표 벡터가 경계값 검사에서 통과되면, $\overline{w}_{j^*}$를 다음과 같이 수정한다:

$$\overline{w}_{j^*}^{(+)} = \frac{\overline{x_i} + \overline{w}_{j^*}^{(-)} \| Nclst_{j^*}^{(-)} \|}{\| Nclst_{j^*}^{(-)} \| + 1}$$

여기서, $\|Nclst_i\|$는 $Nclst_i$에 소속된 벡터의 개수를 나타낸다.

적응 군집화의 학습 알고리즘은 완벽한 입력 벡터에 대해서는 잘 동작하지만 약간의 잡음이 첨가되어도 문제가 발생되는 단점을 갖고 있다. 즉, 잡음이 있는 입력 벡터를 계속 학습시키면, 인식을 잘못하여 발생되는 불필요한 군집단의 수가 급격히 증가한다.

Ⅳ. 제안된 영역화 알고리즘

영상 영역화는 영상을 여러 개의 영역으로 나누는 경계값을 필요로 한다. 본 연구에서 사용된 GA는 실수 경계값을 개체의 값으로 이용하여 유전 연산자를 수행한다. 일반적인 GA는 확률적인 재구성으로 인하여 해가 유동적일 수 있다. 이런 문제점를 해결하고 수행 시간 단축을 위하여 새로운 재구성 방법(GBR: great best reinsertion)을 제안하고, 이때 일어날 수 있는 조기 포화 문제를 해결하는 방법을 제시한다. 최상 선택 재구성 방법(GBR)은 자식 저장소의 개체가 부모의 개체보다 적합도가 좋은 개체를 선택하며, 이때에 일어나는 조기 포화 문제는 초기 집단의 돌연변이를 발생시키는 방법으로 해결한다. 제안된 알고리즘의 구성은 다음과 같다.

1) 초기 개체군의 해(경계값)를 무작위로 선택한다.

2) 각 개체군의 적합도를 구한다. 이때에 적합도 함수는 초기 개체군의 경계값으로 영역화하여 군집단(cluster)의 개수와 유사도(similarity)를 인자로 하여 다음과 같이 표현된다.

$$f(x) = \frac{\sum_{i=1}^{N} \| \overline{w}_{j^*} - \overline{x_i} \|}{N} + \| C^* - GC^* \|$$

$\overline{w}_{j^*}$: 승리 벡터, $\overline{x_i}$: 현재 화소의 명암도,

C^*: 군집단의 수, GC^*: 목표 군집단의 수

3) 주어진 군집단의 수와 현재의 군집단의 수가 같으면서, 주어진 유사도보다 크거나 같은 개체가 있으면 수행을 마친다.

4) 주어진 군집단의 수와 현재의 군집단의 수가 다르면서, 개체들의 유사도가 같은 값으로 수렴하는 조기 포화가 발생하면, 초기 개체군에서 1차 돌연변이[그림 1 참조]를 발생시킨다.

5) 순위적 선택법(rank-based selection)[그림 2 참조]으로 교배 저장소에 초기 개체군 모두를 아래의 위치 값에 따라 새로이 구성하고, 부모 개체군을 적합도 순으로 정렬하여 저장한다. 여러 선택법 중 순위적 선택법은 부모 개체 모두를 교배에 사용하게 된다. 우성인자(1111 | 000)와 우성인자(1101 | 100)의 교배(1111 | 100)보다, 열성인자(0000 | 111)와의 교배(1111 | 111)가 더 좋은 우성인자를 만들 수 있다. 순위적 선택은 선형적으로 선택될 수 있고, 비선형적으로 선택될 수 있다[그림 3 참조]. 아래의 수식은 [그림 3]에 따라 선형적으로 사상될 수 있게 유도된 식이다.

$$MP_i = 2 - SP + \frac{2 \cdot (SP-1) \cdot (L_i - 1)}{(NP-1)}$$

MP_i: i번째 개체의 위치 값, NP: 개체의 수,

L_i: i번 위치, SP: 평균 선택 확률에 대한 적합도가 가장 좋은 개체의 선택 확률 비(ratio)

6) 실수 값의 교배를 위해 중재 교배법(intermediate recombination)[그림 4, 5 참조]을 쓰도록 한다.

offspring = p1 + α * (p1 - p2)

α값은 0~1의 실수를 무작위로 선택한다. p1과 p2는 교배 저장소에서 자식 개체(child individual; offspring)를 생산할 부모 개체(parent individual)로서 선택된다. 새로이 만들어진 개체는 자식 개체군에 저장한다.

7) 2차 돌연 변이[그림 5 참조]

$MV = C_i \pm RV * \delta$; (+와 -는 같은 비율)

$RV = 0.5 * DV$

$\delta = \sum \frac{a(i)}{2^i}$, $a(i)$ 는 돌연변이가 일어날 확률(Pm)일 때 1의 값을 갖고, 그 외는 0.

MV: 돌연변이된 자식 개체, C_i: 무작위로 선택된 i 번째 자식, DV: 개체가 가질 수 있는 해의 범위

8) 새로이 생성된 자식 개체들의 적합도를 구하여, 3)을 수행한다.

9) GBR[그림 6 참조]을 사용하여 부모 개체군의 개체보다 좋은 적합도를 가지는 자식 개체군의 개체로 교환한다.

10) 구성된 새로운 개체군을 2)로 부터 다시 수행한다.

대부분의 선택법은 낮은 적합도의 개체가 다음 세대로 복제될 확률이 낮다. 그러나, 순위적 선택법은 적합도에 관계없이 각 개체의 위치에 따른 복제과정을 가지므로, 영상 영역화에서 경계값 설정 문제와 같이 해가 유동적일 경우에 다른 선택법들보다 적합하다. 구성 7)의 2차 돌연변이는 100세대(generation)마다 수행한다. 제안된 알고리즘에서 얻은 경계값으로 영상의 명암도를 세분화하여 영역화(segmentation)한다. 공간 영역에서 구해지는 거리에 근거하여, 각 화소를 군집단(cluster)에 포함시킨다. 영상을 영역화하는 과정은 ACA의 후단에 GA를 사용하여 최적 경계값을 구하는 부분을 첨가한 것과 같다. GA에서 GBR을 사용한 최적화 과정이 ACA의 경우 보다 빠른 수행 속도를 보이며, GA를 사용한 경계값 조정이 기존의 알고리즘보다 미세하여 영상 구분이 뚜렷해졌다. 일반적인 GA를 사용하는 경우, 확률적인 재구성으로 인하여 적합도가 낮은 자식 개체가 적합도가 높은 부모 개체와 바뀌며, 이때 개체군의 해가 최적해와는 멀어질 수 있다. 그 결과, 구성된 개체군이 다음 세대에서는 낮은 선택 확률을 가지므로 인해, 여러 개체가 이러한 해로 구성이 이루어지는 조기 포화가 발생된다. 이러한 단점을 보완하기 위해 제안한 GBR은 부모보다 적합도가 높은 자식 개체만 선택하도록 하여 개체들의 평균값이 낮아지지 않도록 하였다. 조기 포화 현상은 최적해가 아니면서 개체의 95%이상이 같은 값으로 수렴이 되고, 자식도 부모와 같은 값들로 구성되어 개체의 값들이 바뀌지 않는 현상을 일컫는다. 조기 포화 현상을 두 가지로 나누어 볼 때, 낮은 값으로 수렴되는 저 조기 포화(low premature convergence)와 높은 값으로 구성되는 고 조기 포화(high premature convergence)로 나눌 수 있다. GBR은 저 조기 포화는 방지하지만, 높은 적합도만으로 다음 세대가 구성이 되는 특성으로 인하여 발생하는 고 조기 포화의 경우, 무작위로 하나의 부모 개체를 선택하여 부모 개체군이 가지는 해의 범위 내의 임의의 수를 대입하는 1차 돌연변이를 사용하여 해결한다.

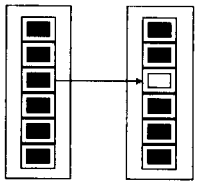


그림 1. 1차 돌연변이
Fig. 1. The first mutation.

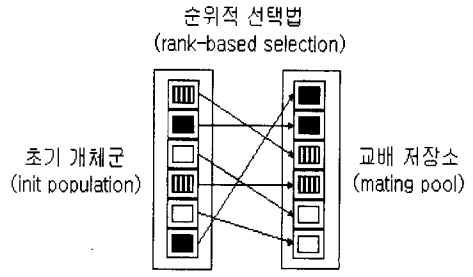


그림 2. 순위적 선택법
Fig. 2. Rank-based selection.

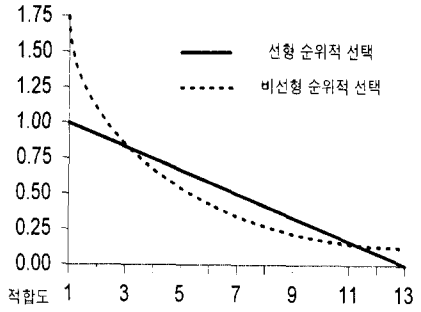


그림 3. 순위적 선택을 도식화한 선형 그래프
Fig. 3. Pictured linear graph for rank-based selection.

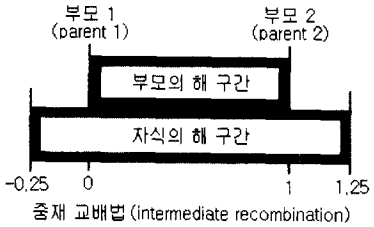


그림 4. 중재 교배법
Fig. 4. Intermediate recombination.

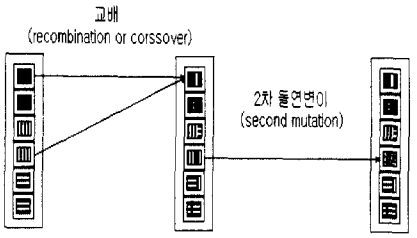


그림 5. 교배와 2차 돌연변이
Fig. 5. Recombination and the second mutation.

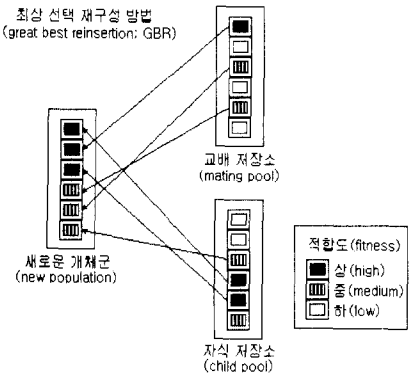


그림 6. 최상 선택 재구성 방법
Fig. 6. Great best reinsertion method(GBR).

V. 소영역 제거 방법

소영역 제거 방법은 영상 영역화에서 잡음 제거에 필요한 과정이고, 특정 물체를 인식하는 데에 기여를 한다. 소영역 제거를 위한 전처리 과정으로 소영역 검출(detection)을 수행 해야하며, 수행 방법은 순차적 검색으로 유크리디안 거리가 멀고 그레이 레벨이 같은 영역을 다른 영역으로 세분화하는 것이다. 다른 영역으로 세분화되는 4 가지의 경우를 아래와 같이 고려하여 각 화소를 소영역으로 검출하였다.

0) 첫 번째 군집단을 초기화하고, 첫 화소를 군집단의 대표 값으로 한다.
1) 이전 화소와 현재 화소가 다른 경우, 새로운 군집단을 만들어 현재 화소를 포함시킨다[그림 7 참조].
2) 1)의 경우이면서 상위 화소와 현재 화소가 같은 경우, 현재 화소를 상위의 군집단으로 포함시킨다[그림 8 참조].
3) 현재 화소가 상위 화소와 같은 값이면서, 다른 군집단으로 영역화된 경우는, 현재 화소와 같은 줄의 이전 화소 모두를 상위 군집단에 포함시킨다[그림 9 참조].
4) 상위의 다른 군집단(1, 2차 집단)이 현재 화소 군집단으로 연결되는 경우, 상위의 2차 집단을 1차 집단에 포함시킨다[그림 10 참조].
5) 모든 화소를 검색할 때까지 1)~4)를 반복한다.
6) 검색을 마친 화소의 군집단의 수가 원하는 값(limit cluster)보다 작은 경우, 바로 옆의 군집단에 포함시킨다[그림 11 참조].

제안된 소영역 제거 방법은 ACA와 제안된 알고리즘에 모두 적용되었으나, 최적의 경계값을 가지는 제안된 알고리즘의 경우에 소영역 제거의 효율이 높아진다. ACA의 경우는 정밀하지 못한 경계값으로 영역화된 결과에 소영역 제거가 이루어지면, 정밀하지 못한 특성이 계속 보존된다.

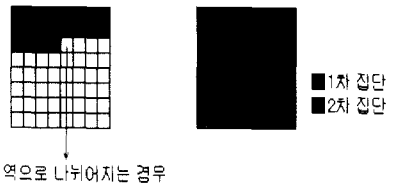


그림 7. 소영역 검출의 첫 번째 경우
Fig. 7. The first case of small region detection.

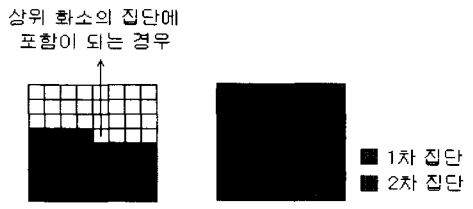


그림 8. 소영역 검출의 두 번째 경우
Fig. 8. The second case of small region detection.

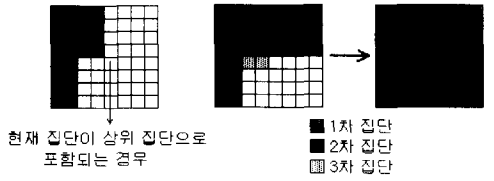


그림 9. 소영역 검출의 세 번째 경우
Fig. 9. The third case of small region detection.

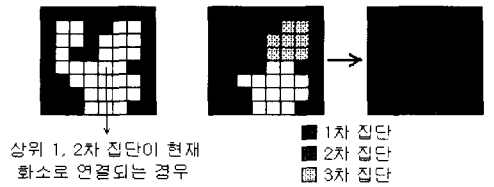


그림 10. 소영역 검출의 네 번째 경우
Fig. 10. The fourth case of small region detection.

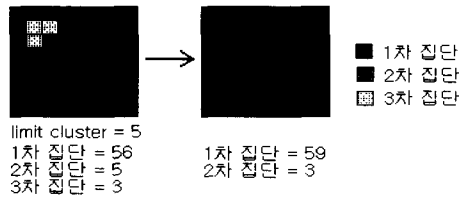


그림 11. 소영역 제거법
Fig. 11. Small region removing method.

VI. 실험 및 고찰

제안된 알고리즘의 성능을 평가하기 위해 실험은 태극 영상, 막대 영상, lena 영상과 girl 영상에 대해서 ACA와 제안한 알고리즘으로 비교해 보았다. 입력으로 사용된 영상 크기는 256*256 화소이며, 각 화소의 명암도는 0~255의 분포를 갖는다. 입력 영상 데이터는 헤더가 없고 바로 영상 데이터 값을 가지는 raw 영상 파일을 사용하였다. BYTE형 메모리를 동적으로 할당하여 파일로부터 읽어 들여 실험에 사용하였으며, 실험에 사용된 환경은 윈도즈 95에서 Visual C++ 5.0으로 시뮬레이션 되었다. 잡음 영상에 대한 영역화를 검사하기 위해 원 영상에 아래의 식과 같은 잡음을 첨가하여 실험하였다.

$$\overline{G} = (1 + l\, G_{rnd}(\cdot)) * \overline{G_0}$$

여기서, $G_{rnd}(\cdot) = \sqrt{-2\ln x}\cos(2\phi)$, x는 [0, 1] 사이의 값이며, ϕ는 $[-\pi, \pi]$ 사이의 값이다. $\overline{G_0}$는 원 영상의 명암도이며, $\overline{G}$는 $\overline{G_0}$에 잡음이 첨가된 영상을 나타낸다. 잡음 변위 l 은 [0, 1]의 범위를 가지는 인자이며, $G_{rnd}(\cdot)$는 정규화된 가우시안 (Gaussian) 난수 발생 함수이다.

그림 12(a)와 (c)는 실험에 사용된 태극과 막대의 원 영상이고, 그림 12(b)와 (d)는 그림 12(a)와 (c)에 가우시안 잡음을 10% 첨가한 영상이다. GA의 초기 집단은 영상에 관계없이 무작위로 추출된 값을 경계값으로 사용하였으며, 영상에 따라 원하는 군집단의 수를 달리하여 수행하였다. 반복횟수가 증가함에 따라서 해의 변동이 적도록 GA의 교배과정에서 α값의 크기를 감소시켰다. 그림 13(a)는 태극 영상에 대하여 ACA를 사용한 영역화 결과이며, 영역화된 영역은 4개로 분류되었다. 제안된 알고리즘에 의한 영역의 개수는 3이며 결과는 그림 13(b)에 나타내었다. ACA는 최대, 최소의 군집단 수로 결과 영상을 만들어 내므로, 영역의 개수가 유동적이다. 제안된 알고리즘과 비교를 위하여 ACA의 목표 군집단의 수에서 ±1의 값으로 실험하였다. 잡음이 첨가된 영상의 영역화 유사도를 측정하기 위해 ACA와 제안된 알고리즘에 대한 결과를 각각 그림 13(c)와 (d)에 나타내었다. 막대 영상에 가우시안 잡음을 10% 첨가한 경우에 대하여, ACA와 제안된 알고리즘의 실험 결과를 각각 그림 14에서 보여준다. 그림 13과 14의 결과로부터, ACA에서는 군집단의 개수가 잡음에 민감한 반응을 보이며, 제안된 알고리즘의 경우는 GA에서 구한 최적 경계값으로 잡음에 강한 결과를 나타내며, 이때 생겨난 불필요한 영역에 대해 영상에 관계없이 영역의 크기가 10이하인 소영역을 제안한 방식으로 검출하여 제거하였다. 군집단의 개수, 초기 경계값, 그리고 유사성을 달리하여 실험한 결과를 표 1과 표 2에 정리하였다.

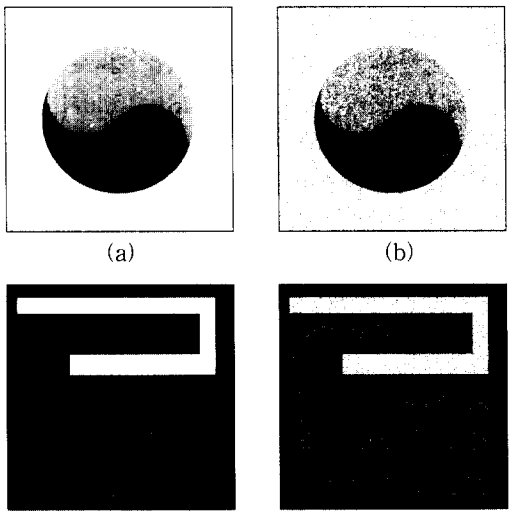


그림 12. (a) 태극 영상 (b) 가우시안 잡음(10%) 첨가된 태극 영상 (c) 막대 영상 (d) 가우시안 잡음(10%) 첨가된 막대 영상

Fig. 12. (a) taeguk image (b) taeguk image with Gaussian noise(10%) (c) bar image (d) bar image with Gaussian noise(10%)

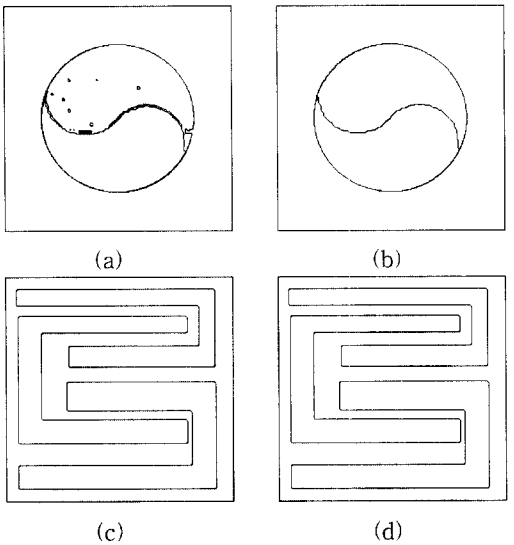


그림 13. (a) ACA를 적용한 태극 영상 영역화 결과 (b) 제안된 알고리즘을 적용한 태극 영상 영역화 결과 (c) ACA를 적용한 막대 영상 영역화 결과 (d) 제안된 알고리즘을 적용한 막대 영상 영역화 결과

Fig. 13. (a) taeguk image segmentation by the ACA (b) taeguk image segmentation by the proposed algorithm (c) bar image segmentation by the ACA (d) bar image segmentation by the proposed algorithm

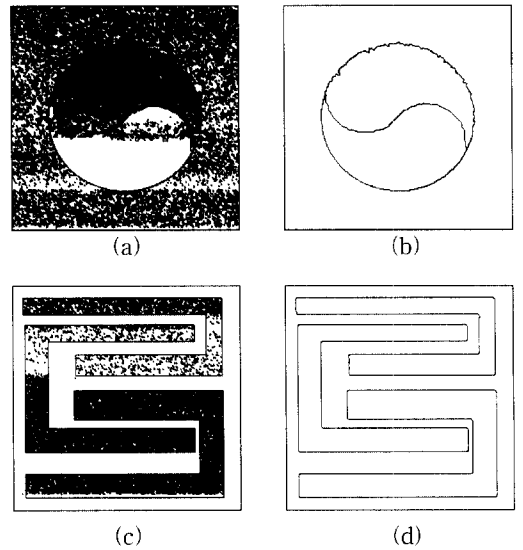


그림 14. 가우시안 잡음(10%) 첨가된 영상에 대한 ACA와 제안된 알고리즘의 실험 결과
(a)(c) ACA에 의한 영역화 결과 (b)(d) 제안된 알고리즘에 의한 영역화 결과

Fig. 14. Simulation results by the ACA and the proposed algorithm for the image with Gaussian noise(10%).
(a)(c) Segmentation results by the ACA (b)(d) Segmentation results by the proposed algorithm

표 1. 태극 영상 영역화 결과

Table 1. Taeguk image segmentation results.

	결과 군집단		결과 경계 값		유사성	
초기 인자	ACA	GA	ACA	GA	ACA	GA
0.1	4	3	0.7750	0.6391	0.9934	0.9911
0.2	4	3	0.8000	0.6175	0.9938	0.9904
0.3	3	3	0.6500	0.6071	0.9921	0.9909
0.4	3	3	0.6750	0.6468	0.9871	0.9911
0.5	4	3	0.7500	0.6024	0.9933	0.9909
0.6	3	3	0.6000	0.6612	0.9824	0.9906
0.7	3	3	0.7000	0.6494	0.9921	0.9911
0.8	4	3	0.8000	0.6514	0.9938	0.9911
0.9	5	3	0.4500	0.6401	0.9944	0.9911
1.0	5	3	0.2500	0.6885	0.9944	0.9912

표 1은 잡음이 없는 태극 영상에 대한 실험 결과이며, 표 2는 가우시안 잡음 10% 갖는 막대 영상에 대한 실험 결과이다. 군집단의 개수는 일정하게 고정시키고 경계값을 0.1씩 변화시키면서 ACA와 GA로 실험한 결과, 유사성은 둘 다 0.99이상으로 높게 나왔으나, ACA를 사용한 경우는 원하는 군집단 개수와는 달리

나오는 경우가 많았으며, GA를 이용한 영상의 경우 목적 군집단의 개수를 명시한 경우 최적의 경계로 나누어 졌다. 제안된 알고리즘을 사용한 경우 영역화를 수행하는 시간이 단축되었으며, 원하는 군집단의 개수를 정확히 분류해 내었다.

표 2. 가우시안 잡음 10% 첨가된 막대 영상 영역화 결과

Table 2. Gaussian noise(10%) added bar image segmentation results.

목표 유사도	결과 군집단		결과 경계 값		유사성	
	ACA	GA	ACA	GA	ACA	GA
0.1	5	3	0.7750	0.7750	0.9911	0.9799
0.2	3	3	0.6000	0.6607	0.9904	0.9818
0.3	3	3	0.6500	0.7210	0.9909	0.9819
0.4	3	3	0.7000	0.6819	0.9911	0.9818
0.5	4	3	0.7500	0.6152	0.9909	0.9818
0.6	3	3	0.6000	0.7027	0.9906	0.9819
0.7	3	3	0.7000	0.6305	0.9911	0.9818
0.8	5	3	0.8000	0.7171	0.9911	0.9819
0.9	3	3	0.7000	0.7179	0.9911	0.9818
0.99	5	3	0.2500	0.7604	0.9761	0.9819

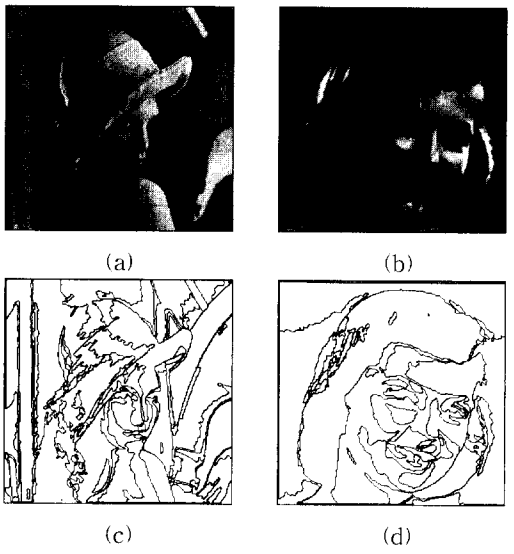


그림 15. 제안된 알고리즘에 의한 영역화 결과
(a) lena (b) girl (c) lena 영상 영역화 결과 (d) girl 영상 영역화 결과

Fig. 15. Segmentation results by the proposed algorithm.
(a) lena (b) girl (c) segmentation result of the lena image (d) segmentation result of the girl image

표 2의 GA를 쓴 경우는 0.9이상의 목표 유사도를 주고 목표 군집단의 수를 3으로 주었을 경우는 유사도 0.991246으로 수렴이 되었다. 그림 15(a)와 (b)는 lena와 girl의 원 영상이고, 그림 15(c)와 (d)는 GA에 의한 영역화 결과이며, 그림 16은 lena와 girl얼굴 영상에 대한 유사도 측정 결과이다. 그림 15(a)와 (b)의 원 영상의 얼굴부분은 명암도의 변화가 적다. 이럴 경우, 미세한 경계값 조정에 따라 전혀 다른 결과를 나타낸다. 제안된 알고리즘의 교배 과정에서 α값의 미세 조정으로, 결과영상이 얼굴부분의 명암도에 따라 최적으로 영역화되었다.

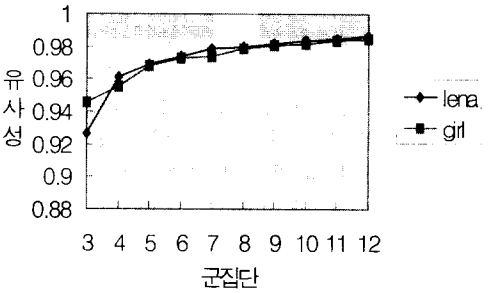


그림 16. 가우시안 잡음(10%) 첨가된 막대 영상에 대한 영역화 알고리즘의 결과 비교

Fig. 16. Performance comparison of segmentation algorithms for the bar image with Gaussian noise(10%).

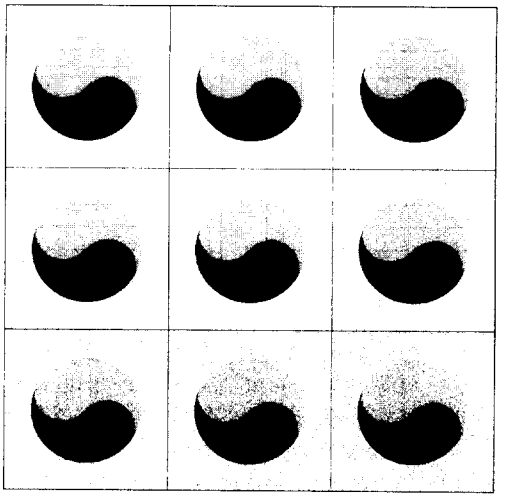


그림 17. 잡음(0~20%)이 첨가된 태극 영상

Fig. 17. Taeguk images with noise(0~20%).

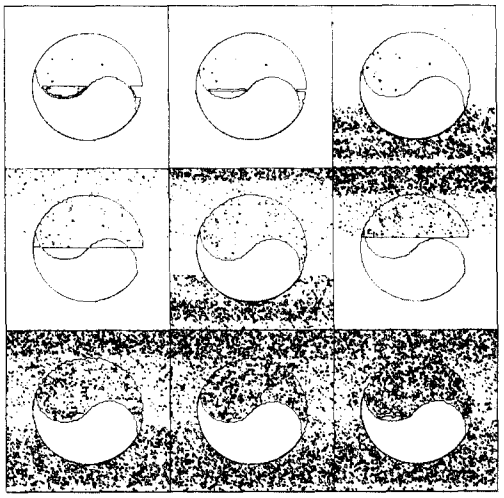


그림 18. 그림 17에 ACA를 적용한 결과
Fig. 18. Results of Fig. 17 by the ACA.

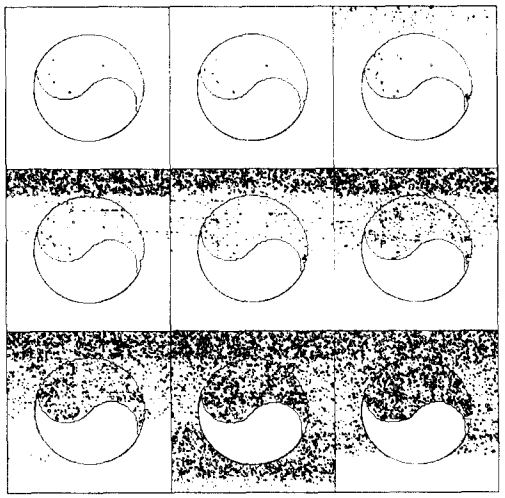


그림 19. 그림 17에 제안한 알고리즘을 적용한 결과
Fig. 19. Results of Fig. 17 by the proposed algorithm.

ACA와 제안한 알고리즘의 잡음에 대한 더욱 미세한 결과를 측정하기 위해, 가우시안 잡음을 0~20% 추가한 영상에 대하여 실험해 보았다. 그림 17은 태극 영상에 잡음을 점차적으로 발생시킨 영상이다. 그림 18과 그림 19는 그림 17의 영상들에 대해 ACA와 GA를 적용한 결과이다. 그림 18을 보면 잘못된 경계값 설정으로 영상중의 몇 개는 잘못된 영역화 결과를 갖게 된다. 그림 19는 제안된 알고리즘의 결과로서 최적화된 경계값으로 잡음의 영향을 최소화하는 결과를 보인다.

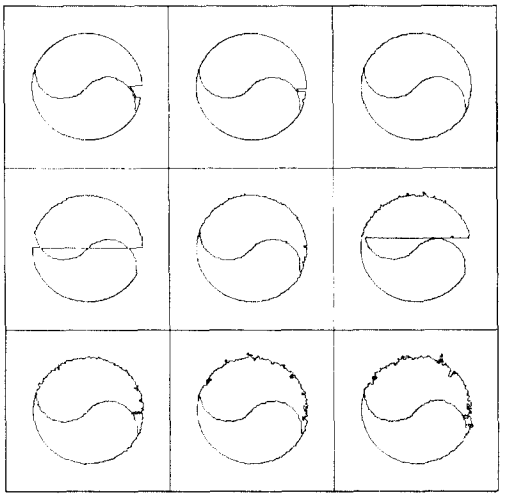


그림 20. ACA 적용후 소영역 제거 결과
Fig. 20. Results by the small region removing after using ACA.

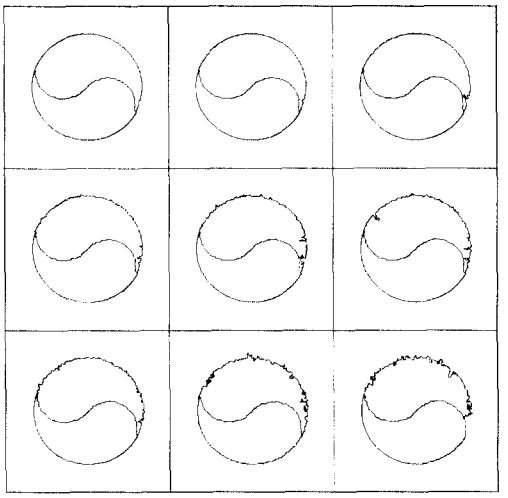


그림 21. 제안한 알고리즘 적용후 소영역 제거 결과
Fig. 21. Results by the small region removing after using proposed algorithm.

그림 20과 21은 ACA와 제안된 알고리즘을 적용후 소영역을 제거한 결과를 나타낸다. 표 3은 잡음에 대한 영역화 결과로서, 제안한 알고리즘의 경우는 영상의 특성에 따라서 최적의 경계가 구해졌으나, ACA는 영상의 특성과는 상관없이 알고리즘의 조건에 맞는 영역화 결과로서 경계가 구해진다. 두 결과에 대해 인식이나 분할에 필요없는 소영역 제거를 수행하였을 때, ACA는 영역화가 잘못 이루어진 영상은 소영역 제거 후에 원하는 영상과는 다른 결과를 나타내었다. 그러

나, 제안한 알고리즘의 소영역 제거는 최적 경계값으로 인해 잡음이 없는 영상과 거의 흡사한 영역화 결과를 보인다. 표 4는 두 알고리즘의 차이를 수치적으로 비교하기 위한 표이다. 표 4에 나타난 결과치는 아래의 식을 사용한다.

$$f(x) = \frac{\sum_{i=0}^{N}\sum_{j=0}^{N}[S_{i,j}(0) - R_{i,j}(x)]}{N*N}$$

N*N: 영상의 크기, $S_{i,j}(0)$: 잡음 0% 영상의 소영역 제거후의 최적화된 영역화의 i행 j열의 화소값, $R_{i,j}(x)$: x%잡음 영상에서의 소영역 제거후 영역화된 영상의 i행 j열의 화소값

표 3. 잡음영상에 대한 경계값 비교
Table 3. Vigilance parameter comparison by noisy images.

잡음 정도 (%)	ACA	제안한 알고리즘	잡음 정도 (%)	ACA	제안한 알고리즘
0	0.62500	0.67718	11	0.50000	0.67903
1	0.62500	0.61003	12	0.50000	0.69874
2	0.62500	0.68257	13	0.62500	0.68441
3	0.62500	0.70348	14	0.53125	0.65183
4	0.62500	0.63399	15	0.62500	0.67716
5	0.56250	0.69013	16	0.50000	0.63373
6	0.56250	0.68728	17	0.50000	0.64852
7	0.62500	0.62884	18	0.62500	0.63216
8	0.50000	0.68531	19	0.62500	0.61234
9	0.62500	0.67035	20	0.56250	0.63267
10	0.62500	0.71258	-	-	-

표 4. 잡음영상의 ACA와 제안한 알고리즘 적용후 소영역 제거결과 비교
Table 4. Comparison of small region removing results after using the ACA and the proposed algorithm.

잡음 정도 (%)	제안한 알고리즘	ACA	잡음 정도 (%)	제안한 알고리즘	ACA
0	0.00000	0.00552	11	0.00445	0.21669
1	0.00087	0.00180	12	0.00521	0.21762
2	0.00053	0.00480	13	0.00549	0.00691
3	0.00102	0.00213	14	0.00547	0.00874
4	0.00141	0.00111	15	0.00552	0.00740
5	0.00161	0.00369	16	0.00808	0.19255
6	0.00256	0.00326	17	0.00941	0.22401
7	0.00337	0.00370	18	0.00965	0.00915
8	0.00312	0.17909	19	0.01046	0.01211
9	0.00346	0.00419	20	0.01370	0.01226
10	0.00444	0.01596	-	-	-

Ⅶ. 결 론

ACA는 주어진 경계값에 의존하여 영역화하며, 잡음 영상 입력에 불필요한 군집단을 분류해 내므로 잡음에 약한 특성을 보인다. ACA에서 의존하는 경계값을 GA를 이용하여 최적의 값으로 찾는 알고리즘을 사용하여 잡음 영상에 대한 특성을 강화시켰다. 본 논문에서 제안한 알고리즘은 ACA보다 수행 시간 단축과 정확한 경계값 조정으로 영상 영역화 결과가 우수하였으며, 잡음에 대해서도 강한 특성을 보였다. 제안된 알고리즘의 응용 분야는 패턴 인식 분야나 공장 자동화 분야 등에서 특정 개수의 군집단을 빠르게 찾는 시스템에서 적용될 수 있으며, 잡음이 많은 전송 매체에서의 물체 인식(object recognition)에도 사용될 수 있다. 인식하는 시스템의 특성에 따라 유사성과 분할될 영역의 수를 조정함으로써 여러 시스템에 적용 가능하다. 앞으로의 연구 과제는 영상 영역화를 하는데 있어서 SOFM(Self Organizing Feature Maps)을 사용한 후, SOFM의 수행 시간 단축을 위해 각 강도 벡터(weight vector)를 개체군으로 구성하여 GA를 이용한 영상 분할법에 관한 것이다.

참 고 문 헌

[1] E. Riseman and M. Arbib, "Segmentation of static scenes," *Computer Graphics and Image Processing*, Vol. 6, 1977, pp. 221-276.

[2] S. W. Zucker, "Region growing: childhood and adolescence," *Computer Graphics and Image Processing*, Vol. 5, 1976, pp. 382-399.

[3] K. S. Fu and J. K. Mui, "A survey on image segmentation," *Pattern Recognition*, Vol. 13, 1981, pp. 3-16.

[4] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer Vision, Graphics, and Image Processing*, Vol. 29, 1985, pp. 100-132.

[5] T. Peli and D. Malah, "A study of edge detection algorithms," *Computer Graphics and Image Processing*, Vol. 20, 1982, pp.

1-21.

[6] S. Y. Kung, *Digital Neural Networks*, PTR Prentice Hall, Englewood Cliffs, New Jersey, pp. 80-85.

[7] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, Reading, Massachusettes(1989).

[8] B. Bhanu and S. Lee, *Genetic Learning for Adaptive Image Segmentation*, Kluwer Academic Publishers, Boston(1994).

[9] G. P. Babu and M. N. Murty, "Clustering with Evolution strategies", *Pattern Recognition*, Vol. 27, 1994, pp. 321-329.

[10] P. Andrey and P. Tarroux, "Unsupervised image segmentation using a distributed genetic algorithm", *Pattern Recognition*, Vol 27, 1994, pp. 659-673.

[11] T. Bäck, F. Hoffmeister, and H. -P. Schwefel, A survey of evolution strategies. In R. K. Belew and L. B. Booker, editors, *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 2-9, San Mateo, CA, 1991. Morgan Kaufmann.

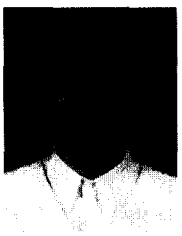
[12] H. Bersini and F. J. Varela. The immune recruitment mechanism: A selective evolutionary strategy. In R. K. Belew and L. B. Booker, editors, *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 520-526, San Mateo, CA, 1991. Morgan Kaufmann.

[13] R. J. Collins and D. R. Jefferson. Selection in massively parallel genetic algorithms. In R. K. Belew and L. B. Booker, editors, *Proceedings of the Forth International Conference on Genetic Algorithms*, pages 249-256, San Mateo, CA, 1991. Morgan Kaufmann.

[14] L. D. Davis, editor. *The Handbook of Genetic Algorithms*. Van Nostrand Reinhold, 1991.

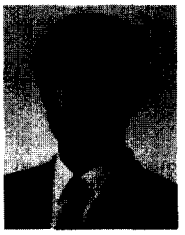
[15] D. B. Fogel and J. W. Atmar. Comparing genetic operators with Gaussian mutations in simulated evolutionary processes using linear search. *Biological Cybernetics*, 63:111-114, 1990

저 자 소 개



河 星 旭(正會員)

1997년 2월 동서대학교 컴퓨터공학과 졸업(공학사). 1997년 3월 ~ 동아대학교 전자공학과 석사과정 재학중. 주관심분야는 영상처리, 패턴인식, MFC Programming



姜 大 星(正會員)

1984년 2월 경북대학교 전자공학과 졸업(공학사). 1991년 8월 Texas A&M Univ. Electrical Eng.(M.S.). 1994년 8월 Texas A&M Univ. Electrical Eng.(Ph.D.). 1984년 3월 ~ 1989년 6월 국방과학연구소 연구원. 1994년 7월 ~ 1995년 8월 한국전자통신연구소 선임연구원. 1995년 ~ 현재 동아대학교 전자공학과 전임강사. 주관심분야는 영상처리, 패턴인식, 영상코딩, 통신시스템 등임