

論文98-35S-12-14

유전 알고리즘에 기반한 퍼지 벌레 검색과 자율 적응 최소-최대 군집화를 이용한 영상 영역화

(Image segmentation using fuzzy worm searching and adaptive MIN-MAX clustering based on genetic algorithm)

河 星 旭 * , 姜 大 星 * , 金 大 鎭 * *

(Seong Wook Ha, Dae-Seong Kang, and Dai-Jin Kim)

요 약

본 논문에서는 퍼지 벌레 검색과 최소-최대 군집화 알고리즘에 기반한 영상 영역화 기법을 제안한다. 전체 영상에서 에지 정보는 픽셀들의 공간 관계를 포함하게 되며, 이를 위해 목적 함수들의 인자를 조정하여 퍼지 벌레의 행동을 정의하며, 에지 정보를 검사하는 방법으로 퍼지 벌레값과 최소-최대 노드를 이용한다. 에지 추출을 사용하는 현재의 영역화 방법들은 수학적 모델에 기반한 마스크 정보를 필요로 하며, 마스크 연산으로 인하여 수행 시간도 많이 걸리게 된다. 반면에, 제안하는 알고리즘은 퍼지 벌레의 검색에 따라 단일 연산을 수행하게 된다. 제안하는 알고리즘에서 필요한 범위의 크기를 스스로 결정하고 빠르고 강력한 계산을 수행하기 위해 최적해를 찾는 유전 알고리즘을 도입하고자 한다. 추가적으로, 영상의 그레이-히스토그램에서 퍼지 검색과 군집화를 수행하기 위해 유전 알고리즘을 사용하는 유전 퍼지 벌레 검색과 유전 최소-최대 군집화가 제안된다. 시뮬레이션 결과는 제안된 알고리즘이 히스토그램을 사용하여 적응적으로 양자화하며, 계산 시간과 메모리를 적게 요구하는 단일 검색 방법을 수행한다는 것을 보여준다.

Abstract

An image segmentation approach based on the fuzzy worm searching and MIN-MAX clustering algorithm is proposed in this paper. This algorithm deals with fuzzy worm value and min-max node at a gross scene level, which investigates the edge information including fuzzy worm action and spatial relationship of the pixels as the parameters of its objective function. But the conventional segmentation methods for edge extraction generally need the mask information for the algebraic model, and take long run times at mask operation, whereas the proposed algorithm has single operation according to active searching of fuzzy worms. In addition, we also propose both genetic fuzzy worm searching and genetic min-max clustering using genetic algorithm to complete clustering and fuzzy searching on grey-histogram of image for the optimum solution, which can automatically determine the size of ranges and has both strong robust and speedy calculation. The simulation results showed that the proposed algorithm adaptively divided the quantized images in histogram region and performed single searching methods, significantly alleviating the increase of the computational load and the memory requirements.

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

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

※ “이 논문은 1997년도 동아대학교 학술연구조성비
(공모과제)에 의하여 연구되었음”

接受日字: 1998年3月4日, 수정완료일: 1998年9月24日

I. 서론

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 서로 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]-[6]}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 기초를 두고 있다. 히스토그램 이치화 기법은 유사성을 이용하는 반면 에지 검출과 영역 추출 기법들은 불연속성을 이용한다. 기존의 영역화 기법들은 어느 정도 경험에 의존하여 영역의 개수와 파라미터를 설정하므로 인하여 다양한 종류의 영상들에는 적합하지 못하다. 그러므로, 입력되는 영상의 종류에 따라 자율적으로 적응하여 영역화하는 알고리즘의 개발이 필요하다. 영상 영역화하기 이전에 자율적으로 군집화를 수행하기 위해 본 논문에서는 ACA(adaptive clustering algorithm)를 사용하여 적응적으로 군집화를 수행한다. 이때, ACA는 군집화의 크기 제한과 초기 문턱치 설정에 대하여 영역화 결과가 판이하게 달라지는 경향이 있으므로 최소·최대 개념을 ACA에 도입하여 인자를 최대값과 최소값으로 확장하여 사용한다. 여기서 사용되는 최소·최대 군집화 기법은 ACA에서 사용되는 단일 인자로 인한 군집화의 집중 포화 문제점을 극복한다. 다시 말해, 단일 인자보다는 두 개의 유효 인자를 가지므로서, 군집화에 유연성을 부가할 수 있다. 최소·최대값을 사용하는데 있어서는 두 인자사이의 크기를 기준으로 군집화를 이루어 낼 수 있다. 이 방법은 영역 크기에 따라서

판이한 영역화 결과를 가져올 수 있다. 이러한 문제점을 해결하고, 최소·최대 인자를 최적으로 수행하기 위해서 유전 알고리즘을 도입한다. 영상 영역화를 위해 ACA에 유전 알고리즘을 도입한 논문^[7]에서는 최적의 문턱치를 구함으로써 최적의 영상 영역화를 수행하였음을 볼 수 있다. 여기서 사용되는 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 탐색 법이다. 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 선택하는가에 따라 좌우된다. GA를 이용하여 영상 영역화를 시도한 방법들^{[7]-[10]}이 여러 가지가 있다. 유전 알고리즘은 해당 문제에 따라서 적절한 적합도 함수를 선택하는 것이 가장 중요하다.

[7]에서는 유사성에 기초로 하여 ACA의 수행 후에 결과로 나타난 군집화에서 전체 에러를 적합도 함수로 선택하였던 방법이며, 적절한 군집단의 개수를 선택하기 위해서 개수와 군집단의 유사성을 바탕으로 적합도 함수에 대입하여 사용한 방법 등이 있다. 마찬가지로 본 논문에서도 적절한 함수를 선택하여 유전 알고리즘을 사용한다. 제안한 AMMCA와 FWSA에서는 몇 가지의 인자가 필요하다. 유전 알고리즘은 제안된 알고리즘의 인자를 최적화하기 위해 사용된다. 실험결과 인자의 값에 따라서 영역화 결과에 지대한 영향을 끼친다. 영역화의 방법에 있어서 기존의 에지 검출이나 영역 추출 등을 이용한 영상 영역화 방법은 3×3 , 5×5 , ..., $n \times n$ 의 마스크(mask)를 필요로 하며, 마스크를 사용한 영상 영역화의 수행 시간은 “영상크기 \times 마스크의 크기”로 나타낼 수 있다. 여기서, 마스크로 사용하는 영역화의 문제점은 수행시간이 많이 걸린다는 것이다. 본 논문에서 제안하는 방법은 수행시간의 단축을 위해 새로운 단일 검색 방법인 FWSA(fuzzy worm searching algorithm)이다. FWSA는 자연상의 벌레가 움직이는 동작 원리를 사용하며, 약자가 강자에게 먹히는 약육강식의 법칙을 퍼지 사상과

퍼지 집합 연산을 이용하여 검색에 적용한다. 이 방법은 기존의 검색법과는 달리 벌레가 현재의 픽셀에서 다음 픽셀로의 움직임 방향이 정해져 있지 않는 것이 특징으로, 자연상태의 벌레가 무작위로 움직이는 것을 모델로 사용한 것이다. 이러한 검색을 사용하면서 이전 위치와 현재 위치의 픽셀의 차이가 문턱치 이상이면 경계로 인식하는 것이다. FWSA의 검색법은 마스크나 다른 방법에서 사용하는 좌측상단에서 우측하단으로의 순차적 탐색법과는 달리 무작위로 생성된 퍼지 벌레가 자신의 무작위적 움직임으로 검색해 가는 방식으로 순차적인 검색법에서 필요한 이전 픽셀과의 보상이 필요없는 것이 장점이다. 여기서 사용되는 문턱치에 따라 영역화 결과가 ACA와 마찬가지로 판이하게 나타나게 된다. ACA에서 적용했던 방식과 마찬가지로 FWSA에서도 유전 알고리즘을 문턱치 최적화에 사용하게 된다. 이때, 유전 알고리즘에 필요한 적합도는 FWSA를 적용한 영역화 결과 영상과 이전 영상의 분산차를 이용하여 구한다. 본 연구에서 사용된 알고리즘들은 다른 영역화 알고리즘에 비해 검색의 단순성과 간단한 연산으로 인하여 빠른 수행속도와 메모리 낭비가 적다. 자율 적응 최소·최대 군집화 알고리즘과 퍼지 벌레 검색을 이용한 영역화에 있어서 알고리즘의 특성상 필요한 여러 제한 요소가 나타나게 된다. 이러한 제한 요소들은 알고리즘의 성능에 지대한 영향을 끼치게 되며, 여러 실험적 결과로서 제한 요소를 설정해야 하는 단점이 나타난다. 그러나, 여기에서는 이러한 제한 요소를 최적의 값을 유도해내는 유전 알고리즘을 사용하여 극복하였다. 본 논문의 구성은 영상의 연속성을 이용한 자율 적응 알고리즘에 최소·최대 개념을 도입한 자율 적응 최소·최대 군집화 알고리즘에 대해서 II에서 설명하고, 불연속적인 성질을 이용하는 퍼지 벌레 검색에 대한 방법을 III에서 제안하며, II와 III에서 필요한 제한 인자들을 결정하는 데 있어서 최적값을 유도하기 위한 유전 알고리즘에 대해서 IV에서 간단히 언급하겠으며, V에서는 III에서 언급한 퍼지 벌레 검색에 유전 알고리즘을 도입하는 방법과 퍼지 벌레 검색이전에 전처리 과정으로 자율 적응 최소·최대 알고리즘을 결합 알고리즘에 대해서 설명하겠으며, 알고리즘에 대한 실험 결과는 VI에서 평가하였으며, VII에서는 본 논문의 결론을 정리한다.

II. AMMCA(Adaptive MIN-MAX Clustering Algorithm)

AMMCA는 ACA(adaptive clustering algorithm)에 MIN-MAX 개념을 도입한 알고리즘이다. 연속적인 특성을 고려한 영역화 방법에서 자율적으로 군집화를 이루어내는 ACA의 경우는, 적절한 인자들의 선택으로 최적의 군집화를 이루어낼 수는 있으나 단일 인자를 사용하므로 각 군집단과 대표 인자사이에 많은 에러를 유발할 수 있다. 이러한 특성을 보완하기 위하여 기존의 ACA에서 사용된 단일 인자를 사용하여 거리 측정에 의한 군집화를 이루어 내는 방식을 이용하여 대표 인자를 최소·최대 군집영역으로 할당하므로서 ACA보다 유동적인 특성을 고려한 군집화 알고리즘을 설계할 수 있다. 구성은 아래와 같다.

- 1) AMMCA에 필요한 최대 유효 군집단의 수 (Cl_m), 군집단 영역의 크기(T_1)를 설정하고, 각 군집단의 최대, 최소 벡터를 0으로 초기화한다.
- 2) 입력 영상에 있어서 최최측 상단을 첫 번째 군집단의 최소, 최대 초기 벡터로 설정한다.
- 3) 2)가 아닌 경우 각 화소를 입력값으로 하여, 아래의 식으로 비교한다.

$$|Min_c - I_{ij}| + |Max_c - I_{ij}| = T_1 \quad (1)$$

여기서, I_{ij} 는 입력값이며, Min_c , Max_c 는 해당 군집단의 최소, 최대값을 나타낸다.

- 4) 식 (1)을 만족하면 현재의 입력을 해당 군집단에 포함시키고, 군집단의 최소, 최대 벡터를 아래의 식으로 갱신한다.

$$Min_c = \min(Min_c, I_{ij}), Max_c = \max(Max_c, I_{ij}) \quad (2)$$

- 5) 식 (1)을 만족하지 못하고 현재 군집단의 개수가 Cl_m 과 같은 경우, 현재의 초기 설정이 군집단을 너무 많이 만들어 영역화에 적합하지 못하므로, 1)의 과정에서 필요한 인자들을 새로운 값으로 설정하여 처음부터 다시 수행하도록 한다.
- 6) 식 (1)을 만족하지 못하고 현재 군집단의 개수가 Cl_m 의 값을 넘지 않는 경우, 새로운 군집단을 형성한다. 이때 새로이 설정되는 군집단의

최소, 최대 벡터를 현재의 입력값으로 설정한다.

7) 모든 입력에 대하여 검색이 완료될 때까지 2)에서부터 재수행한다.

AMMCA는 최소·최대 벡터를 이용하므로 인하여 ACA에 비해 단일 인자를 사용한 거리 측정보다 유연한 특성을 가지게 된다. 그림 1은 ACA와 AMMCA의 인자 범위에 따른 유연성에 대한 것을 나타내고 있다. 그림에서 ACA는 단일 중심인자로서 군집단의 범위를 수용하는 반면, AMMCA는 최소·최대 인자로서 군집단의 범위를 보다 유동적으로 수용할 수 있습니다. 영상의 1차원적인 grey 단계의 범위를 지정하는 것으로, 이해를 돕기 위해 ACA와 AMMCA의 두 인자를 y축과 같이 표기하였습니다.

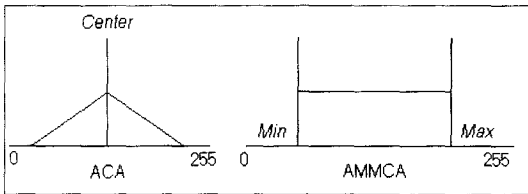


그림 1. ACA와 AMMCA의 영역 범위 비교
Fig. 1. Region range comparison of the ACA and the AMMCA.

III. FWSA(Fuzzy Worm Searching Algorithm)

II의 알고리즘에서 나온 결과에 불연속적인 특성을 이용하기 위해 FWSA를 사용할 것이다. 불연속적인 특성을 고려한 영역화 방법들은 마스크들을 사용하여 주위의 화소와의 차이로 에지를 검출하게 된다. 마스크 사용할 때에는 각 마스크 크기만큼의 수행 시간을 더 필요하게 된다. 불연속적인 특성에 대해서 단일의 수행으로 에지를 검출할 수 있는 새로운 탐색 방법을 고려할 수 있다. 이러한 탐색을 위해, FWSA는 자연상의 벌레의 특성을 컴퓨터의 모델로서 구성한 탐색 방법으로, 벌레 검색에 퍼지 개념을 도입한 단일 패스를 가지는 새로운 알고리즘이다. 에지를 검출하는 방법으로 FWSA에 사용되는 벌레의 생명치에 따라서 에지를 검출할 수가 있다. 제안하는 알고리즘은 퍼지 함수를 이용하여 화소값을 벌레 검색값으로 퍼지화시키게 된다. 또한, 알고리즘 특성상 벌레의 무작위적 검색으로 인하여 벌레들의 중복 할당이 일어나게 된다. 이러한 경우, 퍼지화된 검색값을 퍼지 집합 연산에 대

입하여 자연 현상인 약육강식의 법칙을 적용한다. 여타의 에지 검출 알고리즘은 3×3, 5×5, n×n등의 마스크에 대하여 검색을 수행하므로 인해 한 화소값은 마스크의 수만큼 계산이 수행되는 것에 비해, FWSA는 마스크가 필요없이 자연상의 벌레 진화와 번식의 방법을 이용하므로 인하여 빠른 검색 속도를 가지게 된다.

- 0) 모든 영상의 사상 벡터와 검색 벡터를 0으로 초기화하고 문턱치(threshold) T_2 를 초기화한다.
- 1) 영상 전체 크기의 10% 미만의 초기 벌레들을 무작위로 발생시키고 이들 벌레의 생명치(ξ)를 0.5로 초기화한다.
- 2) 모든 벌레들을 각각의 위치에서 8 방향으로 검색을 수행한다. 이때, 선택되지 않았던 영역에 대하여 검색 우선권(priority)을 주도록 한다. 그림 2는 현 위치(\cdot)에서 다음 위치를 8 방향으로 검색하는 방법을 나타내고 있다.

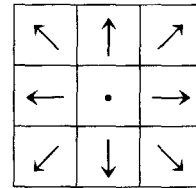


그림 2. 8 방향 검색 방법
Fig. 2. Octal directional searching method.

- 3) 검색 벡터가 0인 벡터 중 현재 좌표의 벌레값과 유클리디안 거리(Euclidean distance)가 최대인 승리 벡터를 구하고, 이때의 벌레 생명치를 일정한 비율로 증가시킨다. 8 방향에서 모든 검색 벡터가 1인 경우 무작위로 방향을 선택하여 승리 벡터를 획득하고, 이러한 경우 검색이 마무리되는 단계로 간주하여, 벌레의 생명치를 급격히 낮추어 주기 위해 이전 증가비율의 배수로 감소시킨다.
- 4) 승리 벡터와 현재 벡터의 거리가 문턱치(T_2) 이상이면 즉,

$$D(Cur_v, Win_v) \geq T_2 \tag{3}$$

여기서, $D(A, B) = \sqrt{A^2 + B^2}$, Cur_v 는 현재 좌표의 벡터, Win_v 는 승리 벡터를 나타낸다. 식(3)을 만족하면 승리 벡터의 사상 벡터는

1이 되며, 위의 식을 만족하지 못하는 승리 벡터의 사상 벡터는 0이 된다.

- 5) 퍼지를 이용하여 벌레의 생명치를 아래의 퍼지 함수를 이용하여 퍼지 사상(mapping)시킨다. 퍼지 함수 $f(x)$ 는 아래의 검색식이 적용되며, 각각의 검색 결과로 인한 벌레 생명치 조절의 결과이다.

$$f(x) = x + \xi$$

if $S_{win} = 0$ and $M_{win} = 1$, then $\xi = 0.2$

else if $S_{win} = 0$ and $M_{win} = 0$, then $\xi = 0.1$

else if $\xi = -0.2$

여기서, S_{win} 은 승리벡터의 검색벡터이며, M_{win} 은 승리벡터의 사상벡터이다.

- 6) 현재 벌레의 인덱스를 승리 벡터의 인덱스로 바꾸고, 바뀐 벌레의 검색 벡터는 1로 바꾸어 검색이 된 좌표에 대한 표시를 한다.
- 7) 생명치가 1이상이 되면 현재의 벌레가 충분한 우성인자를 가진다고 판단하고, 8방향으로 자식을 생성하고 각 자식의 생명치($\xi = 0.5$)를 초기화하게 된다. 이때, 만일 생명치가 0이 되는 벌레에 대해서는 열성인자를 가진 벌레로 판단하여 벌레를 소멸시킨다.
- 8) 서로 중복(overlapping)이 일어난 벌레들에 대하여서는 퍼지 집합 연산(fuzzy set operation)을 사용하여 합병한다. 중복 시에 두 벌레가 이동 전의 승리 벡터들과 검색 벡터가 모두 1인 경우는 둘 모두 빨리 소멸되는 것이 바람직하므로 퍼지 교집합 연산을 벌레의 생명치에 적용한다. 이와 반대로, 두 벌레가 이동 전 승리 벡터들의 검색 벡터 하나가 0인 경우는 좋은 벌레가 살아나도록 퍼지 합집합 연산을 사용한다. 중복의 경우 그림 3과 같고 알고리즘은 그림 4와 같다. 그림 3에서 *은 중복이 일어난 영역이며, w_i 과 w_j 는 i, j 번째 벌레의 생명치를 나타낸다.

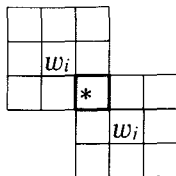


그림 3. 중복이 일어난 퍼지 벌레
Fig. 3. Overlapped fuzzy worms.

```

if Worm [i].SearchVector and Worm [j].SearchVector then
    Worm [i].Life = min(Worm [i].Life, Worm [j].Life)
    // Fuzzy intersection operation of Set A, B = min(A, B)
if not Worm [i].SearchVector or not Worm [j].SearchVector
then
    Worm [i].Life = max(Worm [i].Life, Worm [j].Life)
    // Fuzzy union operation of Set A, B = max(A, B)
    
```

그림 4. 중복된 퍼지 벌레에 대한 퍼지 연산 알고리즘
Fig. 4. Fuzzy operation algorithm for the overlapped fuzzy worms.

- 9) 모든 검색 벡터가 1이 되기 전에 모든 벌레가 소멸되면 무작위로 선택한 영역에 5% 미만의 벌레를 생성하고 각 벌레의 생명치는 0.5로 설정한다.
- 10) 모든 검색 벡터가 1이 되면 검색을 중단하고, 그렇지 않은 경우 단계 2)로부터 재수행한다.

IV. GA(Genetic Algorithm)

유전 알고리즘(GA)^{[11]-[13]}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)^{[14]-[16]}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 문제

에 대한 적합도를 구한다.

3. 재생산(reproduction) 단계

재생산 단계는 구해진 선택확률이나 개체의 인자 값으로 각 개체를 다음 세대 개체로 재생산하는 과정으로 개체들의 적합도나 인자를 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. 아래의 그림 5는 부모 개체로부터 자식 개체군을 생성하여 개체 저장소(population pool)에 넣는 재생산 과정을 설명한다.

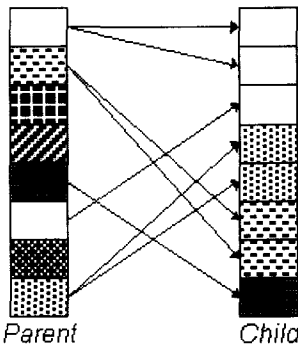


그림 5. 재생산
Fig. 5. Reproduction.

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 교

배법(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 복제(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. 그림 6은 여러 가지 교배법 중에 단일점 교배법에 대한 간단한 구조를 나타내고 있다. 그림 6에서 A와 B는 단일점을 중심으로 서로의 개체가 서로 교배됨을 볼 수 있다.

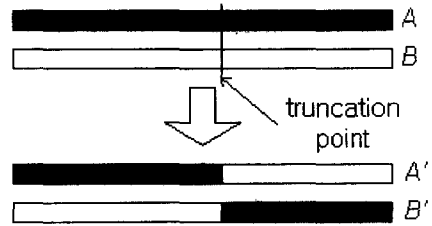


그림 6. 교배
Fig. 6. Crossover.

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. 그림 7은 돌연변이의 과정을 나타내는 것이다. 여기서, A의 개체 중에 무작위로 선택된 위치의 값이 1에서 0으로 바뀐 A'를 볼 수 있다.

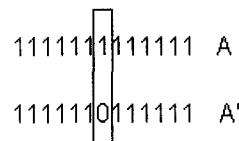


그림 7. 돌연변이
Fig. 7. Mutation.

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.

V. 전체 알고리즘

자율적으로 인자를 찾지 못하는 일반적인 문제의 경우, 반복적인 실험이나 경험에 의존하여 필요한 인자들을 설정하여야 한다. 본 논문에서는 어떠한 문제에 대하여 최적의 값을 유도하는 유전 알고리즘을 도입하였다. 제안하는 최종적인 알고리즘은 연속성을 기조로 영역화를 이루는 AMMCA의 결과에 불연속적인 특성을 고려한 FWSA를 이용하는 것으로, 이 두 알고리즘에서 반드시 필요한 인자 설정 문제를 유전 알고리즘을 사용하여 최적 인자를 유도하여 영역화하였다. 이를 위해 AMMCA에 대해서는 1차 유전 알고리즘을 사용하고 FWSA에 대해서는 2차 유전 알고리즘을 각각 사용한다. 전체 알고리즘의 구성은 다음과 같다. 2의 알고리즘은 1.2와 1.4에 포함된다. FWSA 이전의 전처리 과정으로 AMMCA를 사용하게 되나, 유전 알고리즘을 도입하는데 있어서 적합도를 계산하는 단계의 특성상 FWSA의 단계가 AMMCA에 포함되어 있다.

1. 1차 유전 알고리즘을 사용한 GAMMCA (genetic adaptive min-max clustering algorithm)

1) AMMCA의 영역 범위 T_1 을 위한 1차 유전 알고리즘의 초기 개체군(initial population)을 무작위로 생성한다.

- 2) 구성된 초기 개체군에 대하여 영역화 에러에 대한 적합도를 구한다. 이때 사용되는 적합도는 2)에서 구성된 FWSA로 영역화된 전체 에러로서 계산된다.
- 3) 적합도에 대하여 오름차순으로 정렬하고 교배와 돌연변이 연산을 수행하여 부모 개체군으로부터 자식 개체군을 생성한다.
- 4) 구해진 자식 개체군으로부터 1.2)에서와 같이 적합도를 계산한다.
- 5) 부모 개체군과 자식 개체군으로부터 적합도에 기인하여 가장 좋은 적합도를 가진 개체군으로만 구성되는 GBR(great best reinsertion)을 사용하여 새로운 부모 개체군을 생성한다.
- 6) 모든 개체군의 값이 수렴되거나 반복회수가 마지막일 경우 수행을 중단하고 그렇지 않은 경우 1.2)로부터 재수행한다.

2. 2차 유전 알고리즘을 사용한 GFWSA(genetic fuzzy worm searching algorithm)

- 1) FWSA의 문턱치 T_2 를 위한 2차 유전 알고리즘의 초기 개체군(initial population)을 무작위로 생성한다.
- 2) 구성된 초기 개체군에 대하여 FWSA를 수행한다. 이때 나누어진 영역화에 대하여 라벨링을 적용하여 영역화하고, 각 영역의 분산의 합을 전체 에러로 볼 수 있다. 전체 에러율을 전체 영상 크기로 나누어서 적합도 함수로 사용하고 이때의 값을 적합도로 사용한다. 초기 개체군의 적합도를 아래의 적합도 함수로 구한다.

$$f(x) = \frac{\sum_{i=0}^c \delta(i)}{\text{Image Size}}$$

$\delta(i)$: 해당 영역의 분산값, c : 군집단 개수

- 3) 적합도에 대하여 오름차순으로 정렬하고 교배와 돌연변이 연산을 수행하여 부모 개체군으로부터 자식 개체군을 생성한다.
- 4) 구해진 자식 개체군을 새로운 집단으로 구성하여 단계 2.2)로부터 재수행한다.
- 5) 부모 개체군과 자식 개체군으로부터 적합도에 기인하여 가장 좋은 적합도를 가진 개체군으로만 구성되는 GBR(great best reinsertion)을

사용하여 새로운 부모 개체군을 생성한다.

- 6) 모든 개체군의 값이 수렴되거나 반복회수가 마지막일 경우 수행을 중단하고 그렇지 않은 경우 2.2)로부터 재수행한다.
- 7) 2차 유전 알고리즘에서 FWSA의 문턱치 T_2 가 최적화된 경우 개체군 수렴의 전체 에러율이 1차 유전 알고리즘 개체의 적합도로 사용하게 된다. 1차 유전 알고리즘에서 자식 개체군의 생성된 값이 같은 경우 2차 유전 알고리즘을 재수행하여 적합도를 얻어야 한다. 본 연구에서는 이러한 중복 수행을 줄이기 위하여 1차 유전 알고리즘에서의 개체값으로서 사용된 적이 있는 데이터로 수행되는 경우 데이터베이스를 참조하여 수행 시간을 단축하도록 한다.

AMMCA와 FWSA에 GA를 적용한 전체 알고리즘의 구성은 그림 8과 같다.

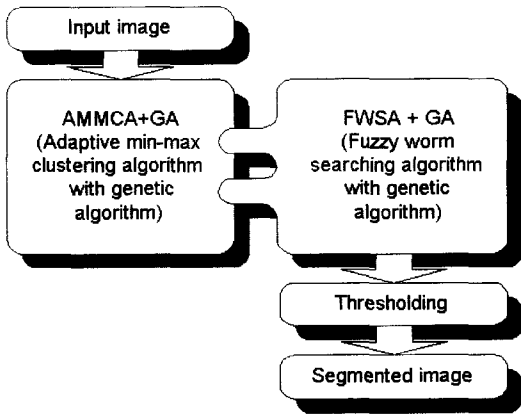


그림 8. 전체 블록도
Fig. 8. Block-diagram of the proposed algorithm.

VI. 실험

본 실험에 사용된 환경은 펜티엄 컴퓨터에서 Windows 95 운영체제를 사용하였으며, 사용된 시뮬레이션 프로그램은 Visual C++ 5.0으로 프로그래밍 하였다. 그림 9는 256 × 256 크기의 헤더를 가지지 않는 Peppers와 Lena raw 영상이 메모리에 읽혀진 시뮬레이션 프로그램을 나타내고 있다. 그림에서 좌측 영상이 Peppers 영상이며, 우측 영상이 Lena 영상이다. 그림 10은 Peppers와 Lena 두 영상에 대하여

AMMCA를 적용한 영상을 보여주고 있다. 그림 10의 영상은 자율 적응적으로 벡터 양자화가 이루어진 결과물이다. 일반적인 벡터 양자화의 경우, 분포가 전혀 없는 영역까지 포함하여 일정 간격으로 양자화를 이루어 적합하지 못한 결과를 나타낸다. 그러나, AMMCA의 경우 영상의 분포를 고려하여 각 화소들의 분포에 따라서 자율 적응적으로 최소·최대 개념에 기반하여 양자화를 이루어 내므로 비트의 손실이 적어진다. AMMCA에서 결과로 나온 영상에 대하여 FWSA를 적용한 결과물이 그림 11에 나타난다. FWSA를 적용한 후의 결과는 퍼지 벌레의 생명치와 관계가 있어 퍼지 벌레가 강하게 살아나게 될 조건의 경우 그레이 농도가 진하게 나타난다. 마찬가지로, 각 영상에 나타난 결과값은 퍼지화에 의하여 그레이 농도의 강약이 결정된 그림이다. 그림 12는 FWSA에서 검색을 퍼지 사상과 퍼지 연산에 의하여 검색되는 과정을 나타내는 그림이다.

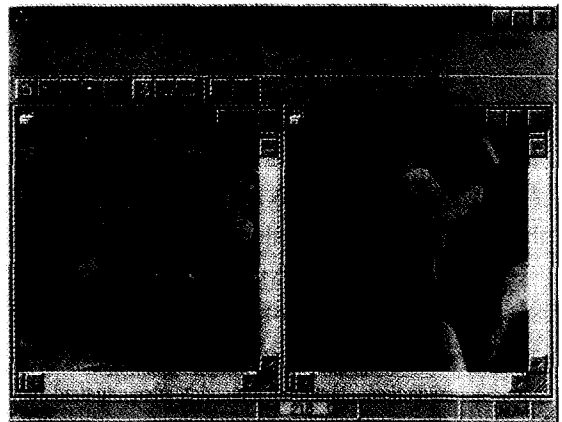


그림 9. Lena영상과 Peppers영상이 로딩된 시뮬레이션 프로그램
Fig. 9. The simulation program loaded both Lena image and Peppers image.

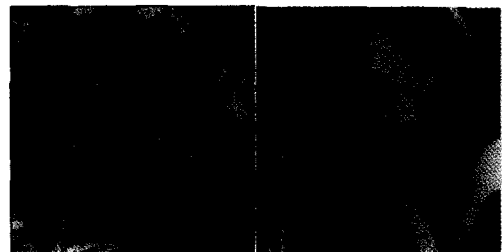


그림 10. GAMMCA를 적용한 영역화 결과
Fig. 10. The segmented images by using GAMMCA.



그림 11. GFWSA를 적용한 영역화 결과
Fig. 11. The segmented images by using GFWSA.

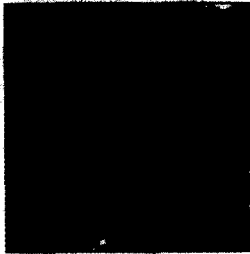


그림 12. FWSA의 검색 과정
Fig. 12. Searching process of FWSA.



그림 13. GAMMCA와 GFWSA를 적용한 후 이진화한 영상
Fig. 13. The segmented binary images after adapting GAMMCA and GFWSA.

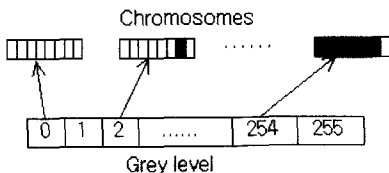


그림 14. 그레이 레벨의 개체 부호화
Fig. 14. Chromosome encoding for grey level.

두 알고리즘이 대입된 결과 영상에 대하여 더 확실한 에지 검출(edge detection)을 위해 적절한 문턱값으로 이진화시킨 영상이 그림 13에 나타난다. 그림 10, 11에서 사용된 AMMCA와 FWSA에서는 사용되는 영역 범위와 문턱치를 최적화하기 위해 유전 알고

리즘이 사용되었다. 여기서, 사용된 유전 알고리즘은 최적화하기 위해 몇 가지 인자 설정이 필요하게 된다. 실험에서 개체 부호화(encoding)는 0~255의 grey 레벨을 8비트의 이진데이터로 사용하였습니다(그림 14). 개체군의 개수는 20개로 사용하였으며, 돌연변이가 일어나는 시기는 반복회수가 10번이 되는 시점에 사용하였다. 유전 알고리즘은 적합도를 정하는 데에 따라서 요구되는 문제에 적절히 적용된다. AMMCA와 FWSA에 사용된 적합도 함수의 경우 전체의 에러를 사용하였으므로, 해당 군집단의 수에 따라서 에러 값이 크게 차이가 났다.

그림 15는 개체군 5, 10, 15, 20개 일 때 개체군의 값들이 수렴되는 그래프를 표현하고 있다. 여기서, 개체군의 수를 5~30개 사이에서 조정하였지만, 모든 경우에 있어서 각 범위에 따른 적합도의 차이가 현저하게 나타나므로 20회 이내에서 빠른 속도로 수렴되었다. 돌연변이는 조기 수렴을 해결하기 위한 것이므로 수렴되기 이전 10회마다 사용하였다. 돌연변이를 20회 내에서 자주 사용한 경우 값이 수렴되지 않고 발산하여 영역화를 제대로 이루지 못하여, 10회로 결정하였다. 또한, 교배는 다중점 교배를 사용하였고, 다중점은 2개의 점을 기준으로 교배하도록 하였다. 재삽입 단계에서는 GBR(great best reinsertion)법을 사용하였다. AMMCA와 FWSA에 사용되는 개체군의 부호화는 모두 정수값을 가지므로 이진화 코딩을 사용하였다. AMMCA의 최소·최대 인자 범위는 50으로 결정되었다. 이러한 값은 영상의 분포에 따라 적합도 함수의 결과로 생성되는 것이다. 최대·최소 인자의 범위를 변경시킨 결과를 그림 16에 표시하였다. 그림에서 최외측 상단을 기준으로 범위 10부터 오른쪽아래 가면서 10씩 증가하면서 실험하였다. 이때, 범위가 작을 때는 군집화 개수가 너무 많으며, 범위가 클 경우에는 군집단이 너무 적어서 영역화가 제대로 이루어지지 않음을 볼 수 있다. 또한, FWSA의 경계 결정에서 사용되는 문턱치는 GA에 의해 최적화된 경우, 26이라는 값으로 수렴되었다. 여기서도 마찬가지로 문턱치는 전체 영상의 에러 분포에 기초로 하는 적합도 함수에 의하여 최적으로 결정된 것으로 볼 수 있다. GA의 전역 검색으로 일반 국부 검색에서 일어나는 국부 수렴성이 극복된 것을 알 수 있다. 마지막으로, 제안하는 알고리즘의 속도와 수행능력을 비교하기 위해 에지 검출 방법 중에 Prewitt 필터를 사용한 결과는 그림 17에 나

타나 있다. 그림 18에서는 5개와 9개의 마스크 셀 (cell)을 사용한 Laplacian 필터를 두 영상에 적용한 영역화 결과이다. 상단의 영상은 5개의 마스크 셀을 가지는 Laplacian 필터를 적용한 것이며 하단부는 9개의 마스크 셀을 가지는 Laplacian 필터를 적용한 것이다. Laplacian 필터를 적용한 경우 Prewitt나 Sobel보다 작은 마스크를 사용하므로 수행 시간은 빠르고 에지 검출은 되나, 소영역이 많이 나타나는 경향이 있습니다. 여기서, 우리는 FWSA를 불연속성에 기초로 하는 마스크연산과 수행속도를 비교하였다. Prewitt과 Sobel 등의 필터를 사용하는 경우 마스크를 사용하게 되므로 인하여 “영상의 크기($O(n)$) × 마스크의 크기($O(m^2)$)” 만큼의 시간이 걸리게 된다. Laplacian 필터와 같이 좀더 작은 필터의 경우는 m^2 이 아닌 m 개의 마스크 요소를 사용하는데, 이전의 계산식에서 $O(m^2)$ 을 m 의 마스크 요소를 사용하므로서 계산할 수 있다. 연산 시간은 해당 시뮬레이션 컴퓨터의 CPU 연산 수행 속도에 비례한다. 상대적인 값으로 비교해보면, FWSA의 단일 연산은 “ $O(n + s) \times$ CPU 연산 시간”이 걸리고, 일반의 m 마스크 연산인 경우는 “ $O(n) \times O(m^2 \text{ or } m) \times$ CPU 연산 시간”으로 계산할 수 있다. 여기서 FWSA는 $O(n + s)$ 영상 크기만큼의 검색 시간에 중복 검색 시간(s)가 걸리나, 중복 검색시 퍼지 벌레가 빠르게 소멸되므로 $O(s)$ 는 아주 작은 상수값으로 볼 수 있으므로, $O(n + s)$ 는 $O(n)$ 으로 나타낼 수 있다. 제안한 알고리즘은 기존의 마스크를 사용한 에지 검출법에 비하여 “1 / 마스크의 크기” 수준으로 수행시간이 감소되었으며, 결과로 나타난 영상에서는 급격히 변하는 영역에 퍼지화를 사용하므로 일반적인 에지 추출의 에지가 아니면 일반 영역으로 나누는 기준에 퍼지의 모호성을 도입하여 퍼지화된 에지를 검출함을 볼 수 있다. FWSA는 불연속성에 기초로 하는 알고리즘이고 AMMCA는 연속성에 기반한 알고리즘으로 나눌 수 있다. 연속성에 기초로 하는 알고리즘에 대한 비교는 비슷한 수행시간을 가지는 연산이므로, 비교 대상에서 제외를 하였다. 전처리 과정에서 사용된 ACA에 최소·최대 개념을 도입한 방법에 대한 검증은 위해 그림 19에서는 상단에 ACA만을 사용한 영역화 결과를 나타내고 있으며, 하단부는 ACA에 유전 알고리즘을 도입한 영역화를 보여주고 있다. 상단부에서는, 이전의 화소값을 군집단 결정에 이용하므로서 군집단이 전파되어 울바르지 못

한 영역화를 보이고 있다. 하단부의 영역에서는 유사성에 기초로 하였으므로, 울바른 군집단 수를 선택하지 않을 경우 영역화가 제대로 이루어지지 않음을 볼 수 있다. 본 논문에서 제안한 AMMCA는 연속성에 기반한 알고리즘이며, FWSA는 불연속성으로 영역화 하는 알고리즘이다.

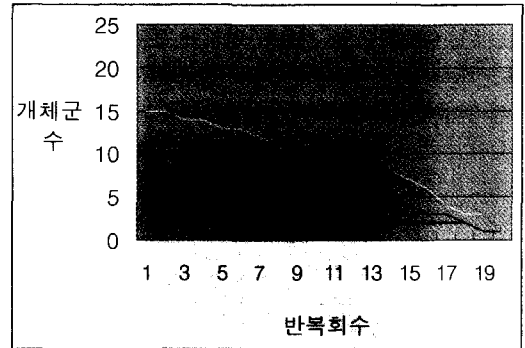


그림 15. 제안된 알고리즘에서의 GA 수렴성
Fig. 15. GA convergence for proposed algorithm.

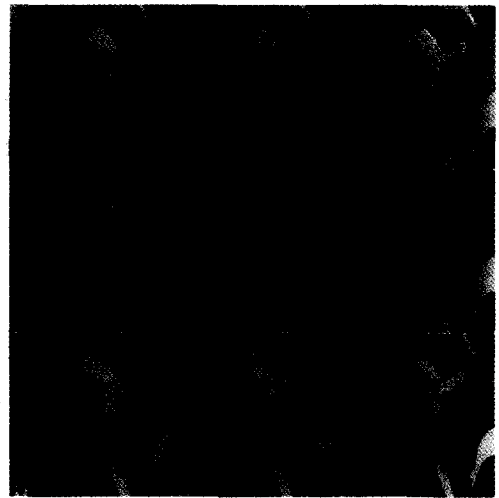


그림 16. AMMCA의 범위(10~90)에 따른 결과
Fig. 16. The results by ranges(10~90) of AMMCA.

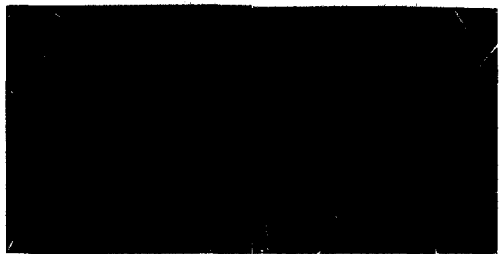


그림 17. Prewitt 필터를 적용한 영역화 결과
Fig. 17. The segmented images by using Prewitt filter.

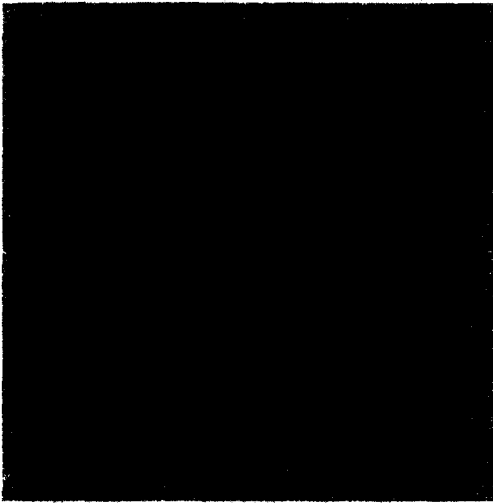


그림 18. Laplacian 필터를 적용한 영역화 결과
Fig. 18. The segmented images by using Laplacian filter.



그림 19. ACA와 GACA를 적용한 영역화 결과
Fig. 19. The segmented images by using ACA and GACA.

VII. 결 론

본 논문에서 제안하는 알고리즘은 영상 영역화에 있어서 새로운 검색 방법과 적용적으로 벡터 양자화 하는 알고리즘을 제안하며, 두 가지 알고리즘에서 필요한 인자의 최적화를 위하여 유전 알고리즘을 대입하여 사용하는 방법에 대하여 제안하였다. 여기서 사용되는 퍼지 벌레 검색이라는 방법은 자연상의 벌레의 자가 증식과 소멸에 대하여 퍼지 개념을 도입한 검색

법으로 기존의 에지 검출 등에 사용되는 마스크 등이 필요 없으므로, 검색 시간을 단축하였다. 제안된 알고리즘은 여러 가지 인자에 의하여 알고리즘의 수행능력이 좌우될 수 있다. 수행능력을 높이기 위해 유전 알고리즘을 사용하여 최적화된 인자로 수행할 수 있을 것이다. 여기서 제안된 퍼지 벌레 검색은 수행 시간이 짧으므로 영역화를 진처리로 사용하는 압축, 인식이나 색인 분야에서 이용 가능하다.

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] X. -P. Zhang and M. Desai, "Wavelet Based Automatic Thresholding for Image Segmentation," *Proceedings of ICIP-97*, Vol. 1, 1997, pp. 224-227.
- [7] 하성욱, 강대성, "적응 군집화 기법과 유전 알고리즘을 이용한 영상 영역화", *대한 전자공학회 논문지*, 34(8)S, pp. 92-103, 1997.
- [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] L. D. Davis, editor. *The Handbook of Genetic Algorithms*. Van Nostrand Reinhold, 1991.

[14] 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.

[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.

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

저 자 소 개

河 星 旭(正會員) 第 34卷 S編 第 8號 參照

姜 大 星(正會員) 第 34卷 S編 第 8號 參照

현재 동아대학교 전자공학과 조교수

金 大 鎮(正會員) 第 34卷 C編 第 7號 參照