

전역 임계치 벡터의 유전적 진화에 기반한 적응형 배경차분화

임 양 미[†]

요 약

주어진 배경 이미지로부터 전경 객체를 분리하는 것을 목표로 하는 배경 차분화 기법에 관한 많은 연구가 있어 왔다. 최근에 발표된 몇 가지 통계 기반 배경 차분화 기법들은 동적인 환경에서 동작할 수 있을 정도로 안정된 성능을 보이는 것으로 보고되고 있다. 그러나 이들 기법은 일반적으로 매우 많은 계산 자원을 요구하며, 객체의 명확한 윤곽을 획득하는데 있어서는 아직 어려움이 있다. 본 논문에서는 점진적으로 변화하는 배경을 모델링하기 위해 복잡한 통계 기법을 적용하는 대신 간단한 이동-평균 기법을 사용한다. 또한 픽셀별로 할당되는 다중의 임계치 대신 유전자 학습에 의해 최적화되는 하나의 전역적 임계치를 사용한다. 유전자 학습을 위해 새로운 적합도 함수를 정의하여 학습하고 이를 이용하여 이미지의 분할 결과들을 평가한다. 본 논문의 시스템은 웹 카메라가 장착된 개인용 컴퓨터에서 구현하였으며, 실사 이미지들에 대한 실험 결과에 의하면 기존의 가우시안 믹스처 방식보다 우수한 성능을 보이는 것으로 나타났다.

Adaptive Background Subtraction Based on Genetic Evolution of the Global Threshold Vector

Yang Mi Lim[†]

ABSTRACT

There has been a lot of interest in an effective method for background subtraction in an effort to separate foreground objects from a predefined background image. Promising results on background subtraction using statistical methods have recently been reported are robust enough to operate in dynamic environments, but generally require very large computational resources and still have difficulty in obtaining clear segmentation of objects. We use a simple running-average method to model a gradually changing background, instead of using a complicated statistical technique. We employ a single global threshold vector, optimized by a genetic algorithm, instead of pixel-by-pixel thresholds. A new fitness function is defined and trained to evaluate segmentation result. The system has been implemented on a PC with a webcam, and experimental results on real images show that the new method outperforms an existing method based on a mixture of Gaussian.

Key words: background subtraction(배경차분화), image segmentation(이미지분할), genetic algorithm(유전자알고리즘), fitness(적합도)

1. 서 론

배경 차분화(background subtraction)는 주어진

배경 이미지로부터 관심의 초점이 되는 전경 객체들을 분리해내는 작업으로[1,2], 가장 단순한 방법은 고정된 배경과 다른 모든 변화된 픽셀들을 전경 객체로

※ 교신저자(Corresponding Author) : 임양미, 주소 : 서울시 노원구 공릉동 172번지(139-743), 전화 : (02)970-6425, E-mail : yosimi@chol.com

접수일 : 2009년 3월 2일, 수정일 : 2009년 5월 18일
완료일 : 2009년 6월 27일

[†] 정희원, 서울산업대학교 매체공학과 외래강사

간주하는 것이다[3]. 그러나 배경 또한 조명과 지형의 변화를 포함하는 것이 일반적이므로, 가우시안 믹스처(mixture of Gaussians, MOG)와 같은 통계적 방식들을 채택할 필요가 있다. MOG 접근법에서는 하나 혹은 그 이상의 가중치가 부여된 가우시안들의 평균과 편차를 매 프레임마다 갱신하는 방식에 의해 각 배경 픽셀의 변화를 모델링한다. 그러나, 이 경우 가우시안의 개수가 작으면 배경을 정확히 식별하지 못하게 되는 문제가 있다[4,5].

커널-기반 밀도 추정(kernel-based density estimation)을 사용하는 비인자적(non-parametric) 방식들은 MOG를 일반화한 방식이다. n 개의 최근 픽셀 값들의 히스토그램에 의거하여 배경의 확률 밀도 함수를 근사한 후 커널 함수(주로 가우시안)와 결합하여 사용한다[6,7] 이 방식들은 변화에 빠르게 적응할 수 있는 반면 계산 비용이 매우 많이 들어 실시간 구현에 적합하지 않다[8].

MOG와 커널 밀도 추정은 감시용 프로그램들에서 움직임을 포착하는 경우와 같이 움직이는 객체의 식별에 주로 사용되는데, 객체가 움직이지 않는 경우에는 전경 객체를 정확히 식별할 수 없게 된다[9]. 그 밖의 실시간 비디오 분할 기술들은 단순화된 배경 모델을 사용한다. Hong과 Woo는 복잡한 색상 모델들을 대신하여 널리 알려진 RGB 색상 공간과 표준화된(normalized) RGB 색상 공간을 사용하였다[2]. 픽셀의 값들이 시간에 따라 가우시안 분포를 따른다는 가정 하에, 이들의 기법에서는 학습 기간 동안 기록된 평균과 표준 편차의 함수로 각 배경 픽셀에 적용될 임계치를 결정하였다. Baoxin과 Sezan 또한 단순화된 색상 공간인 이동 평균(running average)을 사용하여 배경 픽셀들의 평균과 편차를 유지하였고, 지역적인 공간 정보의 문맥을 활용하면서 각 픽셀을 전경과 배경으로 분류하는 과정에 확률 함수를 사용하였다[10].

이전의 몇몇 연구 [1,10,11] 에서와 같이 본 논문에서는 각 픽셀의 이동 평균으로 배경을 모델링함으로써 계산 비용을 절감하는 대신, RGB 색상 공간[12] 상의 최적 임계치 벡터를 찾기 위해 유전자 알고리즘(genetic algorithm, GA)을 사용한다. 이렇게 구한 최적 임계치 벡터를 한 프레임에 속한 모든 픽셀들의 분류에 동일하게 적용함으로써, 픽셀별로 임계치를 계산하는 방식에 비해 공간적, 시간적 비용을 절감할

수 있다. 연결된 전경 영역들을 찾고 후보 임계치 벡터들에 대한 적합도(fitness)를 계산하기 위해 라벨링(labeling) 알고리즘을 사용한다. 본 논문에서 제안하는 방식은 배경 모델이 단순하기 때문에 빠른 분할이 가능하며, 주어진 장소의 조명 및 배경색 분포에 가장 적합한 최적 임계치 벡터를 적응적으로 계산하여 사용함으로써 기존 방식들에 비해 정확한 분할이 가능하다.

논문의 구성은 다음과 같다. 다음 절에서는 본 논문에서 제안하고 있는 방식의 전체적인 구조를 제시하고, 전경 분할을 위한 적합도 함수에 대해 소개하며, 유전 학습 과정에 대하여 기술한다. 3절에서는 실험 결과를 제시하고, 4절에서는 결론과 함께 제안된 방식의 한계에 대하여 서술한다.

2. 적응형 배경차분화 시스템

유전자 알고리즘은 현재 컴퓨터 비전을 포함한 많은 분야에서 최적화 혹은 분류 문제에 대한 해를 탐색하기 위해 널리 사용되고 있다. 유전자 알고리즘을 컴퓨터 비전 프로그램에 적용하는 가장 직접적인 방법은 유전자 알고리즘이 프로그램의 현재 출력들을 참조하여 제어 인자들을 동적으로 최적화하는 것이다. 이렇게 배경 차분화 문제에 유전자 알고리즘을 적용할 경우, 색상들의 분포에 대한 가설들은 더 이상 필요하지 않게 된다[13].

그림 1은 적응형 배경 차분화를 위해 본 논문에서 제안하는 방식의 세 단계를 설명하고 있다.

1. 배경 모델링: 일정 시간 동안 배경 이미지의 각 픽셀에 대한 이동 평균을 계산하여 배경 픽셀들의

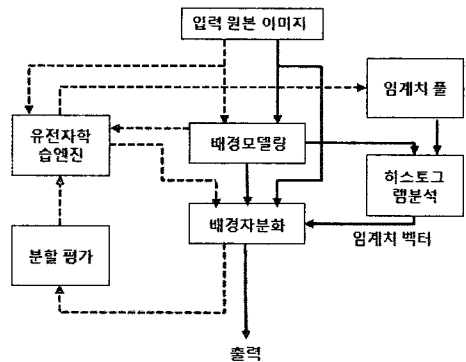


그림 1. 적응형 배경 차분화를 위한 시스템 구조도

색상 분포를 모델링한다.

2. 유전 학습: 반복적인 유전 피드백 과정에 의해 각 이미지에 대한 배경 차분화 결과를 개선해 나간다(그림 1의 점선). 새로운 이미지와 배경 모델이 주어지면, 배경 차분화 모듈(module)은 무작위로 생성된 초기 임계치 벡터를 사용하여 이미지의 각 픽셀을 전경 픽셀과 배경 픽셀로 분류하고, 분할 평가 모듈은 분류의 결과에 대한 적합도를 계산하여 유전자 학습 엔진에 전달한다. 유전자 학습 엔진은 유전 연산에 의해 새로운 임계치 벡터를 선택하고 이를 배경 차분화 모듈에 전달하여 이전의 과정을 반복한다. 입력 이미지에 대해 최적의 임계치 벡터가 결정되었다고 판단되면 이 임계치와 입력 이미지의 히스토그램을 반영하여 임계치 풀을 갱신한다. 임계치 풀이 완전히 갱신될 때까지 2단계를 반복하다가 갱신이 완료되면 학습을 종료한다.

3. 실시간 배경 치환: 학습 결과로 갱신된 임계치 풀을 사용하여 실시간으로 입력 이미지에 대해 배경 차분화를 수행한다(그림 1의 실선). 현재 배경의 색상 분포와 각 임계치 벡터와 함께 저장되어 있는 색상 분포들을 비교하여 하나의 임계치 벡터를 선택하며, 이를 사용하여 배경에 대한 차분화를 수행한다.

2.1 배경 모델링

본 연구에서는 각각의 색상 채널에 8비트를 할당하는 24비트 표준 RGB 색상 공간을 사용한다. 제안하고 있는 유전자 알고리즘은 배경 화면과 전경 객체들을 명확히 분리하기 위해 [0,0,0]에서부터 [255,255,255]까지의 전체 탐색 공간으로부터 최적의 임계치 벡터를 탐색해 내는 것을 목표로 한다. 픽셀 i 의 RGB 색상값은 전경의 경우 $F_i=(F_R(i),F_G(i),F_B(i))$ 로 표기하고 배경의 경우에는 $B_i=(B_R(i),B_G(i),B_B(i))$ 로 표기한다.

실용적인 수준의 배경 차분화를 수행하기 위해서는 서서히 변화하는 조명과 카메라의 움직임 등으로 인해 발생하는 점진적인 변화에 대해 적절히 대처할 수 있어야 한다. 본 논문에서는 픽셀별 임계치들 대신 하나의 임계치 벡터만을 사용하고, 배경 모델은 각 픽셀의 색상 값들에 대한 이동 평균만을 포함한다. 초기 평균은 미리 준비된 배경 이미지로부터 계산하며, 이후 아래 공식에 따라 입력 영상의 매 프레임마다 평균을 갱신한다.

$$B_{i,t+1} = \alpha F_{i,t} + (1-\alpha)B_{i,t} \quad (1)$$

여기서 $\alpha(0 \leq \alpha \leq 1)$ 는 학습률(learning rate), $B_{i,t}$ 는 픽셀 i 의 현재 평균 색상, $F_{i,t}$ 는 새 프레임에서의 해당 픽셀의 색상, 그리고 $B_{i,t+1}$ 은 $t+1$ 시간에 사용할 갱신된 이동 평균이다. 각 색상은 R, G, B 성분을 가지고 있다. α 가 크면 최근의 색상 값들이 더 많이 반영되고, α 가 작으면 적용의 속도가 느려지게 된다.

배경 모델은 전경 객체가 없는 상태의 배경 장면을 처리하여 얻는다. 하지만 전경 객체가 영상에 나타나게 되면, 배경 차분화에 의해 배경으로 분류된 픽셀에 한하여 갱신을 수행한다. 따라서 배경 중 전경에 의해 가려진 부분은 오랜 시간 동안 배경 모델이 변화하지 않을 수 있다.

2.2 배경 차분화

MOG 혹은 이와 관련된 기술들은 변화하는 이미지들도 안정적으로 분할할 수 있지만, 픽셀별 임계치들을 사용해야만 픽셀을 분류할 수 있기 때문에 계산 비용이 많이 든다. 또한, 이들 기술은 정확한 윤곽선의 구분 보다는 커다란 영역(region)들의 위치에 대한 탐지를 위해 설계되었기 때문에 윤곽선이 뚜렷하지 않다는 문제가 있다[10].

회색도(gray level) 히스토그램을 사용하는 다른 몇 가지 접근법에서는 전체 이미지에 대해 하나의 단일한 임계치를 선택하여 사용한다[14,15]. 임계치 이상의 회색 레벨에 해당하는 영역들은 전경 객체로 분류하고, 이미지의 나머지 부분은 배경으로 분류한다. 이러한 전역적 임계치 방식들은 구현하기가 쉽고 빠르게 동작하지만, 전경과 배경이 확연히 구분되는 회색도 분포를 보이는 단순한 화면들만을 다룰 수 있고, 색상에 대하여는 고려하지 않는다. 그러나 그림 2에서 알 수 있는 바와 같이 각 색상 채널은 매우 다른 히스토그램을 가질 수 있다. 따라서 본 논문에서는 빨강, 녹색, 파랑에 대해 서로 다른 임계치 값을 적용하며, 이 세 임계치를 결합하여 배경 차분화(BS, background subtraction)에 사용할 분리 임계치(separating threshold)를 $T_{BS}=(T_R, T_G, T_B)$ 로 표기하기로 한다.

본 논문에서의 배경 차분화는 현재의 배경 모델을 필요로 한다. 먼저, 배경의 RGB 히스토그램[16,17]을 임계치 풀에 있는 학습된 배경들의 k -막대 RGB

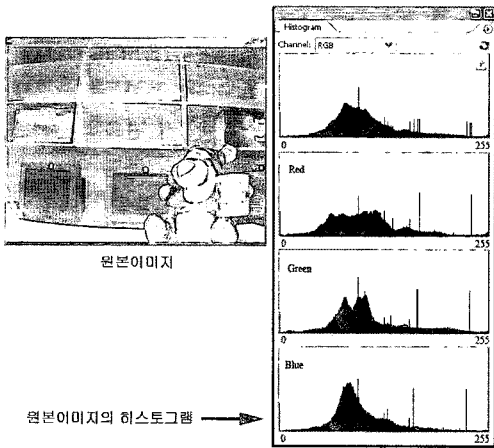


그림 2. 샘플 컬러이미지의 RGB 히스토그램

히스토그램과 막대 단위로 비교한다. 여기서 각각의 학습된 배경은 해당 화면에 대해 유전자 알고리즘이 찾아 낸 최적 벡터와 쌍을 이루고 있다. 두 개의 이미지 X와 Y간의 거리는 식 (3)과 같이 표현할 수 있다.

$$Dist_{XY} = \sum_{i=0}^k \sqrt{(x_i^R - y_i^R)^2 + (x_i^G - y_i^G)^2 + (x_i^B - y_i^B)^2} \quad (2)$$

$$d_i = \begin{cases} 1 & \text{if } |F_R(i) - B_R(i)| > T_R \text{ and } |F_G(i) - B_G(i)| > T_G \text{ and } |F_B(i) - B_B(i)| > T_B \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

여기서 x_i^R, x_i^G, x_i^B 는 이미지 X의 빨강, 녹색, 파랑 히스토그램의 i번째 막대들의 높이이고, y_i^R, y_i^G, y_i^B 는 이미지 Y의 히스토그램 막대들의 높이이다. 임계치 풀의 각 쌍을 Dist에 따라 오름차순으로 정렬하여 최상위 쌍의 임계치 벡터를 배경 차분화에 사용한다.

임계치 벡터가 선택되면, 전경 객체를 배경 화면으로부터 분리하여 이진 맵 di로 저장할 수 있다. 입력 이미지 $F_i=(F_R(i), F_G(i), F_B(i))$ 와 배경 모델 $B_i=(B_R(i), B_G(i), B_B(i))$ 간의 거리를 임계치 벡터 $T_{BS}=(T_R, T_G, T_B)$ 와 비교하여 식 (3)과 같이 구할 수 있다.

2.3 분할에 대한 평가

배경 차분화는 일반적인 이미지 분할(segmentation)에 비해 단순하다고 할 수 있다. 동일한 전경 객체에 해당하여 서로 구분할 필요가 없는 이미지의 각 부분들을 하나로 묶어내기만 하면 되기 때문이다. 따라서 본 논문에서는 분할의 질에 대한 기존의 척도(measure)들을 사용하는 대신, 목표 함수(objective

function)에 해당하는 새로운 척도를 정의하고 이를 이용하여 유전 진화 과정을 제어한다[18].

다른 두 채널에 대한 임계치를 고정한 채, 녹색 채널에 대한 임계치 T_G 가 5에서 90까지 변화하는 동안 잡음 수준이 변화하고 있는 그림 3의 분할 결과들을 살펴보자. 높은 임계치들을 적용한 경우(예를 들어 $T_G=90$), 배경 픽셀의 녹색 값과 매우 다른 녹색 값을 가진 픽셀들만이 전경으로 분류되어, 전경이 여러 개의 작은 잡음 영역들로 분리되거나, 반대로, 낮은 임계치를 적용한 경우, 실제 전경 영역의 바깥 부분에 잡음 영역들이 형성된 두 가지 경우, 모두 잘못된 분류된 픽셀들이 전경 영역의 총 개수가 많아지게 하는 결과를 낳은 셈이다.

이는 분리된 영역들의 평균 크기가 분할의 질을 평가할 수 있는 좋은 척도가 될 수 있음을 의미한다. 그러나 실제 전경이 잡음 영역의 크기보다 작은 객체를 포함하는 여러 객체로 이루어져 있다면 이 척도만으로는 분할의 질을 적절히 평가하기 힘들다. 이러한 문제를 완전히 해결하는 것은 매우 힘든 일이다. 하만, 일반적으로 잡음이 많은 분할 결과는 많은 수의 매우 작은 잡음 영역들을 포함하고 있다고 가정한다면, 이러한 작은 잡음 영역들에 관련된 척도는 분할의 질에 대한 또 하나의 효과적인 척도가 될 수 있다.

앞에서 언급한 척도들을 결합하여 아래와 같은 목표 함수를 분할 결과의 질에 대한 척도로 사용하기로 한다.

$$fitness = ms_b^\alpha - ts_s^\beta \quad (4)$$

여기서 ms_b 는 미리 정해진 임계치 T보다 큰 모든 영역들의 평균 크기(mean size)이고 ts_s 는 T보다 작은 모든 영역의 전체 크기이다. 스케일링 요소 α 와 β 는 두 항 사이의 변위(variance) 편차를 조절하기 위한 것이다. 작은 영역들의 평균 크기 대신 총 크기를 사용하는 것은 잘못된 분류된 픽셀의 총 수를 최종 척도에 반영하기 위한 것이다. 회색도-이미지(gray image)는 잔디 불꽃(Grassfire) 라벨링 알고리즘을

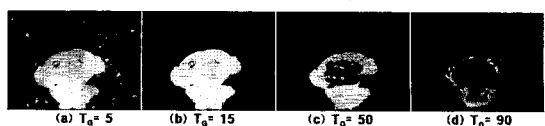


그림 3. 녹색 채널에 각각 다른 임계치 값을 적용하여 얻은 노이즈 분할 결과

사용하여 전경 영역들을 탐색하고 각 영역들의 크기를 구한다[12,19,20]. 재귀 호출의 오버헤드를 피하기 위해 본 논문의 실험에서는 명시적 스택을 사용하여 구현한 버전을 사용하였다.

2.4 유전 학습 시스템

유전자 알고리즘은 각 반복(iteration)마다, 현재 개체군(population)에 포함된 모든 개체(individual)들을 적합도(fitness) 함수에 의거하여 평가한다. 적합도가 높은 일부 개체는 다음 세대에 계속 포함되고, 두 부모의 부분(portion)들을 결합하는 교배(crossover)와 무작위 돌연변이(mutation)에 의해 새로운 개체들을 생성한다.

개체들은 유전자 연산들로 다루기 쉬운 비트열(bit sequence)의 형태로 표현되는 것이 일반적이다. 본 논문에서는 그림 4에 나타낸 바와 같이 2.2절에서 기술한 세 개의 임계치 T_R, T_G, T_B 각각에 8-비트열을 할당하고 이들을 연결한 24-비트열로 하나의 개체를 표현한다. 그림 5는 유전자 알고리즘의 학습 과정에 대한 블록 다이어그램이다. 전역적(global) 적

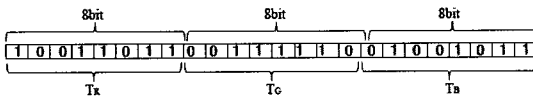


그림 4. 제안 GA에서 개체의 비트열 표현

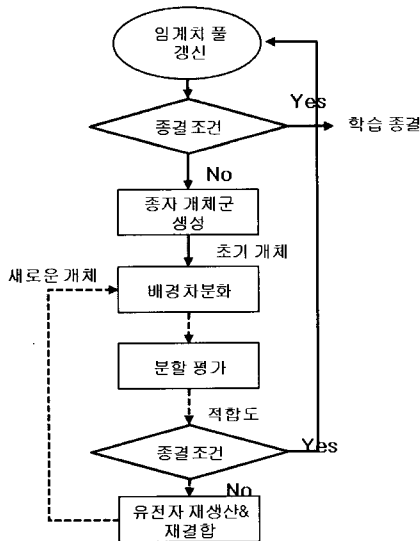


그림 5. 전역적과 지역적 유전자 학습 사이클

용 사이클은 입력 이미지와 현재의 배경을 가지고 시작하며 무작위로 선택된 임계치 벡터들로 구성된 초기 종자 개체군(seed population)에서부터 입력 이미지와 현재의 배경에 적합한 임계치 벡터를 찾아 나간다.

종자 개체군의 각 임계치 벡터를 사용하여 현재 이미지를 처리한다. 만약 가장 좋은 분할 결과의 질이 미리 정해 놓은 임계치를 초과하였으면 지역적(local) 적응 사이클이 종결된다. 임계치 풀의 크기를 N으로 유지하면서, 가장 높은 적합도를 얻은 임계치 벡터와 현재 배경의 히스토그램의 쌍(pair)으로 임계치 풀의 가장 오래된 멤버를 교체한다. 시스템이 아직 만족할만한 분할의 질에 도달하지 못한 경우에는 개체군의 일정 비율을 선택하여 다음 세대에 전달하기 위해, 각 개체 s의 선택 확률이 아래와 같은 선형 순위(linear ranking) 방식의 선택을 수행한다[21].

$$P_s(s) = \frac{1}{N} [\eta_{max} - (\eta_{max} - \eta_{min}) \frac{j-1}{N-1}] \tag{5}$$

여기서 N은 개체군의 크기, $\eta_{mir} = 2 - \eta_{max}$ ($1 < \eta_{max} < 2$) ≥ 0 , j는 s의 순위, η_{max}/N 는 최상 순위 개체가 선택될 확률, 그리고, η_{min}/N 는 최하 순위 개체가 선택될 확률이다.

새 개체군에는 교배(crossover)와 돌연변이(mutation)에 의해 생성된 새로운 개체들이 포함되어 배경 차별화 모듈로 다시 전달된다. 지역적 적응 사이클(local adaption cycle)은 한 세대의 마지막에서 달성된 최대 적합도가 임계치를 초과하거나 그 밖의 종결 조건을 만족하면 종결된다. 시스템은 갱신된 임계치 풀을 가지고 새 입력 이미지와 새롭게 갱신된 배경을 처리하기 위한 전역적 적응 사이클(global adaption cycle)을 다시 시작한다. 임계치 풀이 완전히 새로운 멤버들로 교체될 때까지 이 사이클을 반복한다. 이러한 적응 사이클을 통해 주어진 환경에 적합한 최적의 임계치 벡터를 구할 수 있으며, 이 벡터를 분리 임계치로 사용하여 배경 차별화를 수행한다.

3. 실험 및 평가(Experiment and Evaluation)

3.1 시스템 사양(System Configuration)

테스트를 위한 플랫폼으로는 1-Gbyte 램과

64-Mbyte 비디오 카드가 장착된 펜티엄4(3.0GHz)를 사용하였고, 640×480 해상도의 웹캠을 사용하여 이미지들을 캡처하였다. 크고 작은 영역을 구분하기 위한 임계치로는 100을 사용하였고, 교배와 돌연변이 비율은 각기 50%와 25%로 하였다. 선형 순위 선택을 위한 인자 η_{max} 와 η_{min} 는 각기 1.6과 0.4로 하였다. 각 지역적 유전 학습 사이클마다 30개의 임계치 벡터를 생성하여 초기 개체군으로 사용하였다. 종결 조건은 (1) 세대수가 10에 도달하였거나 (2) 가장 높은 적합도를 가진 개체가 마지막 10 세대 동안 개선되지 않았을 때로 하였다. 중간값 필터를 위한 임계치 값은 4로 하였다.

적합도 계산에 필요한 스케일링 요소 α 와 β 값의 다양한 조합을 테스트하여 선택하였다(그림 6). $\alpha=0.5$ 와 $\beta=1.5$ 로 하였을 때 가장 좋은 결과를 얻을 수 있었다. α 가 너무 작으면 ms_b 가 적합도에 미치는 영향이 너무 작아서 ms_b^a 항에 비해 ts_s^β 가 지배적인 역할을 하게 되었고, 그 결과로 어떠한 전경 영역도 인식되지 않는 방향, 즉 ts_s^β 를 최소화하는 방향으로 시스템이 진화하였다. 역으로, α 가 크고 β 가 작았을 경우에는 ts_s^β 가 적합도에 미치는 영향이 미미하였고, 동시에 ms_b^a 를 최대화하려는 성향으로 인해 전체 이미지를 하나의 전경 객체로 인식해 버리는 문제가 발생하였다(그림 3 참조). β 를 3.0으로 하였을 경우에는 ts_s^β 를 최소화하기 위해 전경 객체의 경계 부근에 위치한 픽셀들을 배경으로 잘못 분류하였고 이로 인해 인식된 객체의 크기가 원래 객체의 크기에 비해 작아지는 현상을 관찰할 수 있었다.

3.2 성능 평가(Performance Evaluation)

본 논문에서 제안하고 있는 방식이 비교적 정적인(static) 배경 화면을 가정하고 있기 때문에, 갑작

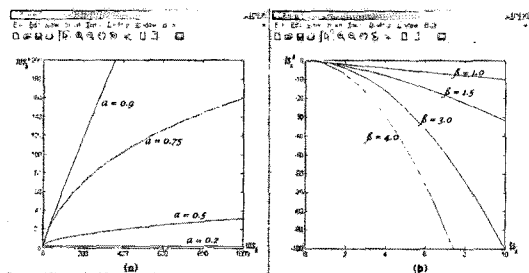


그림 6. 적합도에 관한 α 와 β 의 상호 영향

스런 조명 혹은 배경 지형의 변화가 발생하지 않는 한 학습된 임계치 벡터는 만족스러운 출력을 생성할 수 있다. 유전 학습을 수행한 후, 천정에 6개의 조명이 달려 있는 사무실 환경에서 실험을 수행하였다. 아래는 본 논문에서 제안한 배경차분화 방법과 기존의 MOG 방법의 성능을 비교한 결과이다. MOG의 학습률과 본 논문의 배경 모델링 과정에 적용할 학습률을 모두 0.2로 하였고, 두 방식 모두에 대해 동일한 형태학적 연산(morphological operation)¹⁾을 적용하였다.

그림 7(a)의 경우, 본 논문의 방식은 창틀 부근의 색상 유사(camouflage)²⁾ 영역으로 인해 약간의 잡음 영역들이 머리 내부 영역에 남기는 하였지만, 그림자를 성공적으로 제거하였다. 또한 사람의 다리 사이에 있는 작은 배경 공간도 정확히 배경으로 인식하였다. MOG의 경우에는 결과 이미지의 바다과 책상 밑에 진한 그림자들이 나타났다. 그림 7(b)는 양모(woolen) 재질의 인형을 실험한 것으로 불규칙하게 빛을 반사하는 표면에 대해 전경과 배경을 만족스럽게 분리한 것을 확인할 수 있다. 반면, MOG는 몇몇 잡음 영역을 올바르게 분류하지 못했는데 특히 인형의 어두운 부분 중 카무플라주 현상이 발생한 부분들에 대하여 잘못된 분류를 하였다. 그림 7(c)는 다중의 객체가 너무 근접 배치되어 상호간에 그림자를 드리우게 되어 정확한 전경 식별이 어려운 상황의 예이다. 본 논

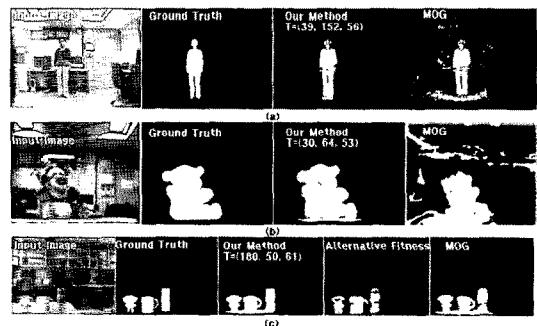


그림 7. 사무실에서 (a)사람, (b)털인형, (c)여러 재질의 객체들을 전경으로 하였을 때의 결과

- 1) 영상의 화질 개선을 위해 영상 처리 프로그램에서 널리 사용되는 기반 연산으로, 임펄스 잡음 제거와 침식(erosion) 및 팽창(dilation) 연산 등을 들 수 있다. 본 논문의 고유한 연산이 아니고 전체 성능에 미치는 영향도 미미하므로 부가적인 설명은 생략하기로 한다.
- 2) 배경의 색상이 우연히 전경의 색상과 일치하는 현상

문의 방식은 세 개의 객체를 결합된 하나의 객체로 잘못 인식하였다. 같은 입력 이미지에 대하여 객체의 총 개수가 많아지면 적합도가 높아지도록 적합도 함수에 새로운 항을 추가하여 실험을 수행하였으며 MOG의 결과 바로 앞에 있는 결과가 그 결과이다. 객체들이 원래 모양보다 축소되었지만 대신 그림자는 완전히 제거할 수 있었다.

표 1은 하나 혹은 그 이상의 전경 객체를 포함하고 있는 다양한 종류의 실내 화면으로부터 얻은 240개의 이미지들에 대한 정량적 성능 평가이다. 이 이미지들에는 그림 7에 보인 이미지들도 포함된다. 표 1은 데이터 집합들의 명칭을 설명하고 있다. 각각은 30개의 이미지로 구성되며, 사무실과 가정집 방 각각에 대하여 한 명의 사람, 하나의 객체, 다중의 객체의 각 조합이 하나의 데이터 집합에 해당한다.

본 논문의 방식을 사용하여 얻은 결과를 실제 객체의 모양(ground truth) 및 MOG를 사용한 결과와 비교한다. 표 2는 Toyama 등이 수행한 방식과 동일한 방식으로 수행한 성능 비교 결과이다[22]. 여기서

표 1. 샘플 이미지

	F	Single person	Single object	Multiple objects
E				
사무실(office)		OSP	OSO	OMO
가정집 방(domestic room)		DSP	DSO	DMO

F: Foreground E: Environment

표 2. 잘못 인식된 픽셀들을 비교한 정량적 실험 결과

알고리즘	오류 유형	OSP	OSO	OMO	DSP	DSO	DMO
제안한 방법	False neg.	1227	309	532	1835	721	1390
	False pos.	1125	478	1998	2989	1039	2779
	합계	2352	787	2530	4824	1760	4169
MOG	False neg.	306	667	2714	2944	1839	4939
	False pos.	8633	4306	10254	11131	6844	7920
	합계	8939	4973	12968	14075	8713	12859

F: Foreground E: Environment

표 3. 정량적 실험 비교: S(A,B)

Algorithm	OSP	OSO	OMO	DSP	DSO	DMO	Average
Proposed	0.823	0.912	0.742	0.678	0.851	0.598	0.767
MOG	0.626	0.708	0.606	0.552	0.697	0.532	0.620

false-negative 값은 주어진 데이터 집합에서 전경으로 분류하지 못한 전경 픽셀 수의 평균이며, false-positive 값은 배경을 전경으로 분류한 픽셀 수의 평균이다. OSP의 false-negative의 경우를 제외한 모든 유형의 오류에 대해 본 논문의 방식이 우수한 성능을 보였다.

잘못 분류한 픽셀의 수는 서로 다른 크기의 객체들을 포함하고 있는 분할 결과들을 비교하는 경우에는 그 자체로 이상적인 척도가 될 수 없다. 따라서 본 논문에서는 아래와 같이 정의되는 추가적인 유사성 척도 S(A,B)를 사용한다.

$$S(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{6}$$

여기서 A는 탐지한 전경 영역이고 B는 실제 전경 영역이다. 표 3은 이 척도를 사용하여 실험 결과를 비교하고 있다. 표 2와 3 모두 서로 다른 유형의 모든 테스트 화면에 대해 본 논문의 방식이 MOG에 비해 우수한 성능을 보임을 알 수 있다. 특히, false-positive의 수를 현저히 감소시킬 수 있었다. 전경 바깥의 잡음 영역에 대해 패널티를 부가하는 적합도 함수가 이러한 결과에 크게 기여한 것으로 판단된다.

3.3 복잡도 비교(Complexity Comparison)

형태학적 필터링에 소요되는 경비를 고려하지 않는다면, m을 가우시안의 개수라고 할 때 MOG는

$O(m)$ 의 복잡도를 가진다. 각 프레임의 중심에 위치한 가우시안 커널들을 계산하는데 사용된 최근 프레임의 수를 n 이라 할 때, 커널 밀도 추정 모델은 $O(n)$ 의 복잡도를 가진다. 마찬가지로 형태학적 필터링의 경비를 고려하지 않을 때, 본 논문의 방식은 각 픽셀을 분류하기 위해 RGB 각 채널에 대해 단 한 번의 색상 차이를 구하기 위한 비교를 수행하고, 배경 모델 또한 각 픽셀에 대해 세 개의 인자만을 갱신한다. 따라서 복잡도는 $O(1)$ 이다. 또한, 본 논문의 방식에서는 오직 하나의 전역적인 임계치 벡터를 유지하는 과정에 이동 평균을 사용하므로, 배경 모델 자체 이외에는 추가적인 픽셀 당 메모리를 필요로 하지 않는다.

4. 결 론

본 논문에서는 유전자 알고리즘을 사용한 새로운 배경 차분화 방식을 제안하였다. 복잡한 통계 모델 대신 이동 평균을 사용하여 배경을 모델링하였고, 배경 차분화 과정에서 모든 픽셀들에 대해 동일한 하나의 전역 임계치 벡터를 사용함으로써, MOG와 비인자적 밀도 추정 방식들에서 사용되는 픽셀별 임계치들로 인한 복잡도 문제를 개선하였다. 이 벡터를 최적화하는 과정에, 전경 객체를 정확히 분할하는 것을 목표로 하여 설계된 새로운 적합도 함수를 사용하는 유전자 알고리즘을 사용하였다.

실내 환경에서 수집한 실제 이미지 데이터를 사용한 테스트 결과, 제안된 방식은 비슷한 인자를 사용한 MOG에 비해 우수한 성능을 보였다. 이러한 유전적 접근법은 전경 객체의 명확한 분할이 필요한 여타 실시간 이미지 프로세싱 응용들에도 효과적으로 적용될 수 있을 것으로 판단된다. 그러나, 현재의 시스템은 빠르게 움직이는 전경 객체나 급속한 배경의 변화에 대처하는 데에는 한계가 있으며, 이러한 문제를 해결하는 것이 향후 연구의 주된 내용이 될 것이다.

참 고 문 헌

- [1] R.J. Qian and M.I. Sezan, "Video background replacement without a blue screen," *International Conference on Image Processing (ICIP 99)*, Vol. 4, pp. 143-146, 1999.
- [2] D. Hong and W. Woo, "A background subtraction for a vision-based user interface," *Fourth Pacific Rim Conference on Multimedia (IEEE PCM)*, Vol. 1 pp. 263-267, 2003.
- [3] M. Piccardi, "Background subtraction techniques: a review," *IEEE SMC International Conference on Systems, Man, and Cybernetics*, Vol. 4, pp. 3099-3104, 2004.
- [4] L. Li, W. Huang, I.Y. Gu and Q. Tian, "Statistical modeling of complex backgrounds for foreground object detection," *IEEE Trans. on Image Processing*, Vol. 13, Issue 11, pp. 1459-1472, 2004.
- [5] A.E. Igammal, D. Harwood and L. Davis, "Non-parametric model for background subtraction," *6th European Conference on Computer Vision*, Part II, pp. 751-767, 2000.
- [6] A. Mittal and N. Paragiosm, "Motion-based background subtraction using adaptive kernel density estimation," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, Vol. 2, pp. 302-309, 2004.
- [7] K. Kim, T. Chalidabhongse, D. Harwood and L. Davis, "Real-time foreground-background segmentation using codebook model," *Real Time Imaging*, Vol. 11, Issue 3, pp. 172-185, 2005.
- [8] M. Harville, "A framework for high-level feedback to adaptive, per-pixel, mixture-of-Gaussian background models," *7th European Conference on Computer Vision*, Vol. 3, pp. 543-60, 2002.
- [9] L. Greengard and J. Strain, "The fast Gauss transform," *SIAM Journal on Scientific Computing*, Vol. 2, Issue 1, pp. 79-94, 1991.
- [10] L. Baoxin and M.I. Sezan, "Adaptive video background replacement," *Proc. IEEE International Conference on Multimedia and Expo*, pp. 269-272, 2001.
- [11] Y.S. Raja, J. Mckenna and S. Gong,

“Segmentation and tracking using colour mixture models,” *Asian Conference on Computer Vision*, Vol. 1, pp. 607-614, 1998.

[12] 권혁중, 배상근, 김병국, “스테레오 CCD 카메라를 이용한 이동체의 실시간 3차원 위치추적” 한국GIS학회지, 13권, 2호, pp.129-138, 2005.

[13] B. Bhanu, L. Sungkee and J. Ming, “Adaptive image segmentation using a genetic algorithm,” *IEEE Trans. on Systems, Man and Cybernetics*, Vol. 25, No. 12, pp. 1543-1567, 1995.

[14] L. García-Pérez, M.C. García-Alegre, J. Marchant and T. Hague, “Dynamic threshold selection for image segmentation of natural structures based upon a performance criterion,” *3rd European Conference on Precision Agriculture (3ECPA)*, pp. 193-198, 2001.

[15] P.K. Sahoo, A.A Farag and Y.P. Yeap, “Threshold selection based on histogram modeling,” *IEEE International Conference on Man and Cybernetics*, Vol. 1, pp. 351-356, 1992.

[16] 이창수, 지정규, “멀티미디어 서비스를 위한 동영상 이미지의 특징정보 분석 시스템에 관한 연구,” 한국데이터베이스학회, 9권, 3호, pp. 1-12, 2002. (RGB)

[17] D.E. Goldberg, “Applications of genetic-based machine learning,” Chap. 7 in *Genetic Algorithms in Search, Optimization & Machine Learning*, Addison Wesley Publishing, pp. 261-307, 1989.

[18] B. Bhanu, “Automatic target recognition: state of the art survey,” *IEEE. Trans. on*

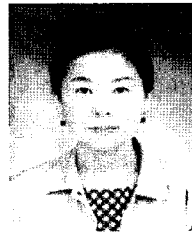
Aerospace and Electronic Systems, Vol. AES-22, pp. 364-379, 1986.

[19] K. Ma and H. Wang, “Region-based non-parametric optical flow segmentation with pre-clustering and post-clustering,” *IEEE International Conference on Multimedia and Expo (ICME)*, Vol. 2, pp. 201-204, 2002.

[20] 강동중, 하종은, “Visual C++ 을 이용한 디지털 영상처리,” *사이텍미디어*, pp. 257-259, 2003.

[21] J. E. Baker, “Adaptive selection methods for genetic algorithm,” *First International Conference on Genetic Algorithms and Applications*, Their Applicat., pp. 101-111, 1985.

[22] K. Toyama, J. Krumm, B. Brumitt and B. Meyers, “Wallflower: principles and practice of background maintenance,” *7th International Conference on Computer Vision*, pp. 255-261, 1999.



임 양 미

1993년 서울산업대학교 매체공학과 학사

1998년 큐슈예술공과대학(현 큐슈대학교) 정보전달 전공 예공석사

2009년 중앙대학교 첨단영상대학원 영상공학과 박사

2004년 9월 ~ 2006년 8월 성신여자대학교 미디어정보학부 조빙교수

2000년 3월 ~ 현재 서울산업대학교 및 성신여자대학교 외래강사

관심분야 : 영상처리, 인터랙티브 커뮤니케이션