

낮은 피사계 심도 영상에서 관심 물체의 효율적인 추출 방법

박정우*, 이재호**, 김창익***

요 약

본 논문은 낮은 피사계 심도 영상(low depth-of-field image)에 대해 사용자의 도움 없이 포커스 된 관심 영역을 고속으로 추출하는 효율적인 방법을 제안한다. 우리는 입력 영상에 존재하는 고주파 성분을 HOS(higher order statistics) 계산을 함으로써 영상의 포커스 된 영역을 찾아내는 중요한 지표로 활용한다. 본 논문에서 제안하는 방법은 크게 4가지 단계로 구분할 수 있다. 첫 번째 단계에서는 기존 연구[1] 방법과 동일하게 모든 화소에 관해 HOS 지도를 계산하고 블록화한다. 두 번째 단계에서는 블록화 된 HOS를 이용하여 포커스 된 물체가 존재하는 후보 관심 영역을 대략적으로 구한다. 이후 관심 영역 내부에 존재하는 구멍(hole)을 제거하기 위해 구멍(hole) 추적 및 제거 연산을 수행한다. 마지막으로 최종 관심 후보 영역에서 배경 부분의 화소만 제거하여 포커스 된 관심 물체만을 섬세하게 추출한다. 제안하는 방법은 기존 방법[1]에 비해 정지 영상에서 고속으로 관심 영역을 추출하므로 추후 알고리즘의 변형 없이 낮은 피사계 심도의 동영상에 확장 적용하여 관심 영역을 실시간으로 추출할 수 있다. 본 논문에서 제안하는 방법은 가상 현실(VR)이나 실감 방송, 비디오 인덱싱 시스템과 같은 여러 응용 분야에 효과적으로 적용될 수 있고, 이러한 유용성은 실험 결과를 통해 보였다.

An Efficient Object Extraction Scheme for Low Depth-of-Field Images

Jungwoo Park[†], Jaeho Lee^{**}, Changick Kim^{***}

ABSTRACT

This paper describes a novel and efficient algorithm, which extracts focused objects from still images with low depth-of-field (DOF). The algorithm unfolds into four modules. In the first module, a HOS map, in which the spatial distribution of the high-frequency components is represented, is obtained from an input low DOF image [1]. The second module finds OOI candidate by using characteristics of the HOS. Since it is possible to contain some holes in the region, the third module detects and fills them. In order to obtain an OOI, the last module gets rid of background pixels in the OOI candidate. The experimental results show that the proposed method is highly useful in various applications, such as image indexing for content-based retrieval from huge amounts of image database, image analysis for digital cameras, and video analysis for virtual reality, immersive video system, photo-realistic video scene generation and video indexing system.

Key words: Object of Interest(관심 물체), Low Depth of Field(낮은 피사계 심도), Image Segmentation (영상 분할), Immersive Video(실감 방송)

※ 교신저자(Corresponding Author): 김창익, 주소: 대전광역시 유성구 문지동 한국정보통신대학교 공학부(305-732), 전화: 042)866-6168, FAX: 042)866-6245, E-mail: ckim@icu.ac.kr
접수일: 2006년 1월 3일, 완료일: 2006년 6월 7일
* 삼성전자 기술 총괄팀

(E-mail: jwpark@icu.ac.kr)

** 한국정보통신대학교(ICU) 영상정보처리연구실 석사 과정
(E-mail: jaeho@icu.ac.kr)

*** 한국정보통신대학교(ICU) 조교수

※ 본 연구는 한국정보통신대학교 신입교원 기초연구과제의 연구비 지원에 의하여 수행되었음.

1. 서론

일반적으로 사용되는 영상 분할(image segmentation)에서는 영상의 화소 세기, 색상 또는 경계 정보와 같은 저 수준(low level)의 물리적 성질을 바탕으로 영상을 분할하기 때문에 영상의 주제가 되는 관심 영역만을 추출하기에는 한계가 존재한다[2,3]. 그리고, 사용자의 도움이 없이 자동으로 일반 영상에서 의미 있는 관심 영역을 찾기는 매우 어렵다[4]. 왜냐하면 영상에서 의미 있는 관심 영역의 판단은 사람마다 다른 매우 주관적인 요소이기 때문이다. 일반적으로 사용자는 영상에서 관심 영역을 추출하기 위하여 자신의 인지정보를 이용하여 관심 물체가 존재하는 영역을 수동으로 정한다[4,5,6]. 이는 배경이 존재하는 일반 영상의 관심 물체에 대한 인식이나 합성을 위한 영상 분할이 일반적으로 물리적인 특징을 이용하여 영상 내부를 균일한 영역으로 구분하는 것뿐만 아니라, 의미상으로 관심 물체 영역과 그렇지 않은 배경을 분리하는 기능도 필요하다는 것을 시사한다.

최근 사용자의 도움 없이 자동으로 영상의 주제가 되는 관심 영역을 추출하는 연구가 많이 진행되고 있다[1,7-11]. 기존 연구[1,7]는 그림 1과 같은 낮은 피사계 심도 영상에서 HOS 지도를 사용하여 자동으로 포커스 된 관심 물체(OOI: object-of-interest)를 추출하는 방법을 제안하였고, 프레임 사이의 분할된 객체의 연속된 움직임(motion) 정보를 추정하여 동영상의 경우로 응용하였다[8]. 이러한 연구[1,7,8]에서 공통으로 사용된 모폴로지 필터(morphological filters by reconstruction)는 관심 물체의 외곽선을 그대로 유지한 채 잡음(noise)과 물체 내부의 구멍(hole)을 효과적으로 제거함으로써 HOS 지도를 단순화한다. 하지만 필터의 사용을 위해 요구되는 많은 계산량으로 인해 고속으로 관심 영역을 추출하거나

동영상 정보를 다루기가 어렵다. 본 논문에서는 낮은 피사계 심도의 정지 영상을 다룰 때, 기존 연구[1,7,8]에서 사용했던 계산량이 높은 모폴로지 필터의 사용을 피하고, 영상을 블록화 하여 처리해 주는 방법을 제안함으로써 고속으로 포커스 된 관심 영역을 추출하는 것을 목표로 삼는다.

Cheng 등은 동영상에서 관심 영역을 설정할 때, 다양한 물리적인 특징(색상, 빛의 세기, 움직임 벡터 등)이 반영된 특징 지도(feature map)를 사용하였다. 지도에서 최대값이 관심 영역을 나타내는 두드러진 특징 값이라 정의하여 그 위치를 기준으로 효과적으로 관심 영역을 제한하였다[9]. 본 논문에서는 사진사가 포커스 정보를 활용하여 자신이 의도하는 관심 영역만을 선명하게 표현하는 낮은 피사계 심도 카메라 기법(low depth-of-field technique)을 이용한 그림 1과 같은 영상을 입력으로 한다. 기존 연구[1,7,8]에서는 관심 영역을 추출하기 위한 HOS 지도를 작성하기 위해 영상에 따라 다양한 범위를 가지는 HOS 값을 0 에서 255 사이로 다운 스케일링하여 사용하였다. 우리는 비록 이런 HOS가 영상마다 다양한 값을 가지지만 그것의 최대값의 위치는 기존 연구[9]와 유사하게 항상 포커스 된 관심 영역에 속한다는 것을 관찰과 실험을 통해 알 수 있다. 따라서 이러한 HOS 특징을 이용하여 포커스 된 관심 물체를 섬세하게 추출하기 전에 대략적인 관심 영역을 블록화하여 효율적으로 제한 가능하다.

본 논문에서 제안하는 기법은 포커스 된 객체를 기반으로 하는 다양한 영상 처리 응용분야에 효율적으로 적용할 수 있다. 예를 들어, 대용량 영상 저장 장치에서 내용 기반의 영상 검색을 위한 영상 인덱싱, 전자 현미경을 사용한 분자나 세포 수준의 3차원 영상 복원 및 분석, 디지털 카메라를 위한 영상 개선(image enhancement), 깊이 추정을 하기 위한 거리 영상 분할(range segmentation), 다양한 포커스 정보를 가진 여러 장의 영상 합성[15]과 같은 다양한 분야에 폭넓게 적용할 수 있다. 이와 같은 일련의 낮은 피사계 심도의 영상으로부터 포커스 된 관심 영역을 추출하여 3차원으로 재구성 가능한 객체 모델링에도 효율적으로 적용할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 광 기하를 통해 피사계 심도(depth of field)에 관한 개념을 설명하고, 낮은 피사계 심도의 영상을 수학적인 모델

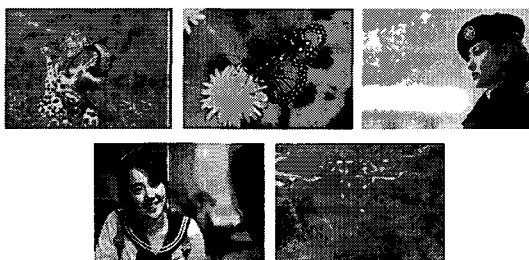


그림 1. 낮은 피사계 심도의 영상

링을 한다. 3장에서는 본 논문에서 제안하는 낮은 피사계 심도 영상에서 포커스 된 관심 영역을 블록화하여 추출하는 방법에 대해 자세히 알아보고, 4장에서는 제안한 방법과 기존 방법[1,8]을 실제 낮은 피사계 심도 영상에 적용하여 성능을 비교한 결과를 보여주며, 5장에서는 본 논문의 결론을 제시한다.

2. 광 기하(Optical geometry)와 낮은 피사계 심도(Low depth of field: Low DOF)

2.1 광 기하의 수학적 모델

카메라와 같은 광학 시스템은 렌즈와 렌즈를 통해 들어온 상이 맺히는 영상면으로 구성된다. 그림 2는 광학 시스템인 카메라의 광 기하를 보여준다. 렌즈의 초점 거리(focal length)가 f 이고, 피사체와 렌즈 사이의 거리가 s , 영상면으로부터 렌즈까지의 거리가 d 라고 하면, 포커스 된 영상에 관해서 다음과 같은 간단한 식이 성립한다.

$$\frac{1}{s} + \frac{1}{d} = \frac{1}{f} \tag{1}$$

물체와 렌즈 사이의 거리가 s 보다 크거나 작을 때 d_f 와 d_r 사이에 존재하는 물체는 영상면에 한 점으로 정확히 포커스 되지 않고 영상면 앞뒤에 초점이 형성되므로 실제 필름 상에는 원의 형태로 맺히게 된다. 광센서의 민감도에 따라 이렇게 원의 형태로 영상면에 맺힌 것과 포커스 되어 한 점으로 맺힌 물체를 구분하지 못하고 모두 포커스 되었다고 간주하는 최대 크기의 원을 착란원(circle of confusion, blur circle)이라 한다.

피사계 심도(DOF)는 카메라를 이용해서 사진을 찍을 때 광 축(optical axis)을 기준으로 물체가 선명하게 포커스 되는 거리($d_f + d_r$)를 의미한다. 즉, 피사

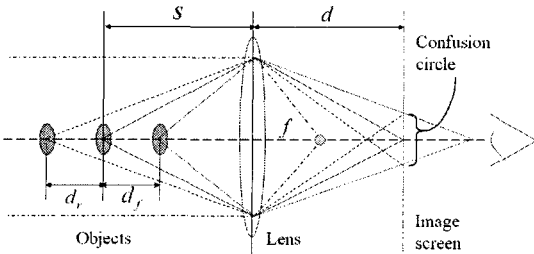


그림 2. 카메라의 광 기하

체에 초점을 맞추면 그 앞쪽과 뒤쪽의 일정한 거리 내에 초점이 맞는데, 이때 그 범위를 이르는 말을 피사계 심도라 한다. 사진을 찍을 때 피사계 심도를 짧게 하면 포커스 된 물체 부분만 선명하게 강조되고 피사체의 전, 후에 존재하는 배경 부분은 흐려지는 효과를 얻을 수 있다. 이러한 기법을 사용하면 그림 1과 같은 2차원 영상에서 3차원 영상과 같은 깊이감을 느낄 수 있는 낮은 피사계 심도 영상을 얻을 수 있다[10].

2.2 낮은 피사계 심도 영상의 수학적 모델

낮은 피사계 심도 영상은 영상의 주제가 되는 포커스 된 물체와, 초점이 맞지 않아 흐릿하게 보이는 물체나 배경 영역으로 구성된다. 포커스가 맞지 않아 흐릿하게 보이는 영상의 배경이나 물체의 모델은 다음과 같이 2차원 가우시안 함수를 이용하여 나타낼 수 있다[12].

$$G_o(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{2}$$

σ 는 영상의 흐릿한 (defocusing, blurring) 정도를 조절하는 필터 매개 변수로 사용된다. 그래서 (x, y) 화소 위치에서 포커스 되지 않아 흐릿한 영상 $I_d(x, y)$ 는 식 (3)처럼 원래 포커스 된 영상 $I_f(x, y)$ 와 가우시안 함수 $G_o(x, y)$ 를 선형 컨볼루션(linear convolution)을 한 형태로 모델링 할 수 있다.

$$I_d(x, y) = G_o(x, y) * I_f(x, y) \tag{3}$$

식 (3)과 같이 영상에 저역 통과 필터(low pass filter)를 적용하여 높은 주파수를 가진 부분을 제거하는 결과는 낮은 피사계 심도 영상에서 포커스 되지 않은 흐릿한 영역으로 나타낼 수 있다. 그래서 피사계 심도가 낮은 영상에서 확실하게 포커스 된 부분은 그렇지 않은 영역보다 높은 주파수 성분을 가진다는 가정은 주파수의 크기를 비교함으로써 포커스 된 선명한 영역과 그렇지 않은 영역을 구분하는 단서를 제공해준다.

3. 제안하는 알고리즘

낮은 피사계 심도의 입력 영상에서 포커스 된 관심 영역을 분할하기 위한 전체 시스템은 그림 3과

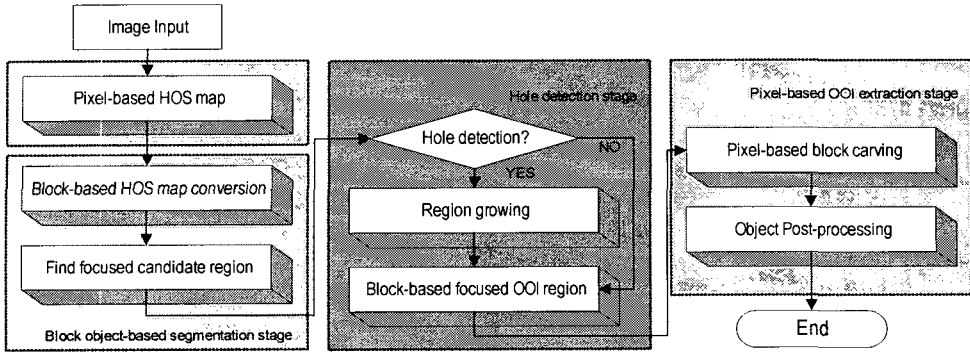


그림 3. 전체 시스템 블록 다이어그램

같이 크게 4개의 모듈로 나뉜다. 첫 번째 모듈에서는 주어진 입력 영상에 대하여 기존 연구[1]와 같이 각 화소에 대해서 색상에 기반한 4차 모멘트와 HOS 지도를 계산한다. 그리고 두 번째 모듈에서는 화소 별로 계산한 4차 모멘트의 결과를 블록화하여 HOS 지도를 재 작성한다. 이렇게 블록화 된 지도 내에서 포커스 된 관심 영역에 속한 블록들을 추출하는 과정을 수행한다. 세 번째 모듈에서는 추출한 관심 영역 내부에 잡음이 존재할 경우 그것을 추적하여 내부를 채우는 연산으로 이루어진다. 마지막으로 입력 영상의 포커스 된 물체를 화소별로 섬세하게 출력하기 위해 블록화 된 관심 영역 내에 존재하는 배경 부분과 주변 잡음을 제거하는 과정으로 구성된다. 다음 절에서는 HOS 지도를 바탕으로 관심 영역을 추출하는 알고리즘의 각 단계에 대해서 자세히 알아본다.

3.1 색상 정보에 기반한 HOS 지도 작성

낮은 피사계 심도 영상은 포커스 유무만 판단할 수 있다면 사용자의 도움 없이 의미 있는 관심영역을 추출할 수 있다. 영상의 포커스 정보를 바탕으로 관심 영역을 추출하기 위해서는 사물의 포커스 여부를 판단할 수 있는 특징 공간으로의 변환이 필수적이다. 앞에서 살펴본 바와 같이 입력 영상에서 포커스가 맞지 않아 흐릿한 배경 부분은 가우시안 잡음으로 구성된 모델[12]로, 포커스 되어 선명한 물체는 비가우시안 모델로 나타난다. 이와 같은 두 수학적 모델을 구분하기 위해 고차 통계 방법(HOS)인 영상의 4차 모멘트를 이용해 입력 영상을 특징 공간으로 변환한다[1,7].

색상 기반의 HOS 지도를 구하기 위해서는, 영상

크기가 $M \times N$ 인 RGB 입력 영상 $I(x,y)$ 에 대해서, 기존 연구[7]에서 제안된 바와 같이 각 채널에 대한 4차 모멘트를 각각 계산한다. 우선, 입력 영상의 red 채널에 관한 한 화소 위치 (x,y) 에서 4차 모멘트는 다음과 같이 나타낼 수 있다.

$$\hat{m}_r^{(4)}(x,y) = \frac{1}{N_\eta} \sum_{(s,t) \in \eta(x,y)} (I_{red}(s,t) - \hat{m}_r(x,y))^4,$$

$$\hat{m}_r(x,y) = \frac{1}{N_\eta} \sum_{(s,t) \in \eta(x,y)} I_{red}(s,t),$$

$$(0 \leq x < M, 0 \leq y < N) \tag{4}$$

여기서, $\eta(x,y)$ 은 (x,y) 를 중심으로 한 화소 그룹이고, $\hat{m}_r(x,y)$ 은 $\eta(x,y)$ 에서 영상 $I_{red}(x,y)$ 의 평균값이다. 또한, N_η 은 $\eta(x,y)$ 을 중심으로 한 화소 그룹의 크기를 나타낸다. 같은 방법으로 모든 RGB 채널에 대해 4차 모멘트 $\hat{m}_r^{(4)}(x,y)$, $\hat{m}_g^{(4)}(x,y)$, $\hat{m}_b^{(4)}(x,y)$ 를 구하고, 화소 기준으로 세 모멘트의 결과 중 가장 큰 최대값을 다음과 같이 취한다.

$$HOS_{original}(x,y) = \max(\hat{m}_r^{(4)}(x,y), \hat{m}_g^{(4)}(x,y), \hat{m}_b^{(4)}(x,y)) \tag{5}$$

이렇게 구한 4차 모멘트의 결과는 관심 영역의 위치를 추론할 수 있는 특징 좌표(saliency point)를 제공하고, 관심 영역을 구하기 위한 블록화 된 HOS 지도를 작성하는데 유용하게 사용된다. 식 (5)의 결과는 영상에 따라 그 결과값의 동적 범위(dynamic range)가 굉장히 다양하므로, 다음 식을 이용하여 HOS 지도가 0부터 255사이의 값을 갖도록 조절한다.

$$HOS(x,y) = \min(255, HOS_{original}(x,y) / DSF) \tag{6}$$

DSF는 다운 스케일 요소(down scale factor)로서 실험을 통하여 300이 적당함을 알았다[7]. 이렇게 색

상에 기반한 HOS 지도는 서로 뚜렷이 다른 색상 임에도 그 밝기 값이 비슷하여 흑백 영상 내에서 경계가 불명확한 경우에 효과적으로 포커스 된 영역을 구분할 수 있는 지표로 사용 가능하다.

그림 4-(a)는 입력 영상이고, 그림 4-(b~d)는 각 색상 채널에 대한 HOS 값을 나타낸다. 그림 4-(e)는 식 (6)를 적용하여 계산한 최종 색상 기반 HOS 지도이다.

3.2 블록에 기반한 관심 영역의 추출

기존 알고리즘[1]에서 적용된 모폴로지 필터 (morphological filters by reconstruction)는 관심 영역의 외관을 그대로 유지한 채 영역 내부의 잡음과 외부의 잡음을 효과적으로 제거하는 특징을 가지나, 반면 표 1과 같이 전체 알고리즘의 약 71.2%를 차지하는 많은 계산량을 가지므로 전체 시스템의 속도를 저하시키는 원인으로 작용한다. 실제로 이러한 약점은 고속으로 관심 영역을 추출하는 응용 프로그램이나 실시간으로 동영상을 다루는데 큰 제한 사항으로

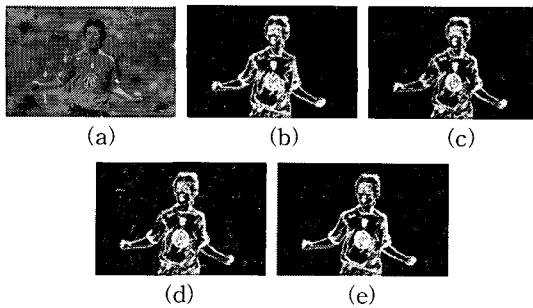


그림 4. 색상 기반 HOS 지도: (a) 낮은 피사계 심도의 입력 영상, (b) Red 채널 HOS, (c) Green 채널 HOS, (d) Blue 채널 HOS, (e) 최종 색상 기반 HOS 지도 (DSF = 300)

표 1. 낮은 피사계 심도의 영상에서 포커스 된 관심영역을 추출하는 기존 방법(1)의 단계별 평균 연산 시간

384×256 크기의 영상에 대한 기존 알고리즘[1]	연산시간 (ms)	(%)
HOS지도	47.33	1.0
모폴로지 필터를 사용한 HOS 단순화	3,537.00	71.2
영역 병합	1,377.67	27.7
임계치를 기준으로 관심영역 추출	6.67	0.1
합 계	4,968.67	100.0

작용한다[1,7,8]. 이 절에서는 낮은 피사계 심도의 정지 영상이나 동영상 데이터에서 고속으로 포커스 된 관심 영역을 추출하기 위해서 기존 연구[1]에서 소개했던 모폴로지 필터 대신 화소 기반의 HOS 지도를 블록화하여 처리하는 방법에 대해 설명한다.

3.2.1 영상의 HOS 특징

3.1절의 $HOS_{original}$ 값은 입력 영상(그림 5-(a))에 따라 그 값의 범위가 다양하므로(그림 5-(b)), 식 (6)을 이용하여 다운 스케일링을 한다. 이렇게 계산한 HOS 지도는 기존 연구[1]에서처럼 모든 입력 영상에 대해 포커스 여부를 가려내는 수단으로 사용된다(그림 5-(c)). 그러나, 식(6)을 통해 $HOS_{original}$ 지도가 [0, 255] 범위를 갖도록 스케일링을 할 경우, 그림 5-(b)와 같이 원래 $HOS_{original}$ 결과의 상세함을 반영하지 못한다. 우리는 이처럼 낮은 피사계 심도 영상의 정규화 되지 않은 HOS를 관찰한 결과, HOS의 최대값인 피크(peak)는 영상 내의 포커스 된 관심 영역의 일부뿐인 것을 확인하였다. 즉, 영상 내에서 포커스 된 물체는 높은 HOS 값을 가지고, 그 물체 내에서 가장 높은 HOS 값이 존재한다. 그래서 다음 절에서 정규화 되지 않은 $HOS_{original}$ 의 속성을 이용하여 블록 기반의 HOS 지도를 만들고, 이를 바탕으로 영상에서 포커스 된 관심 영역을 검출하는 방법에 대해서 알아본다.

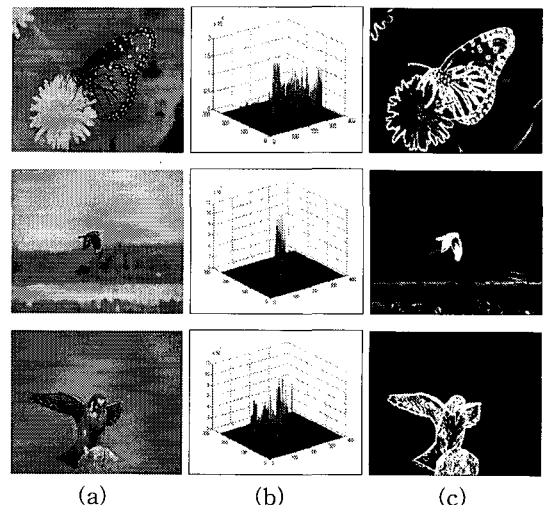


그림 5. HOS의 특징: (a) 낮은 피사계 심도의 영상, (b) HOS 계산 결과, (c) HOS 지도

3.2.2 블록 기반 HOS 지도 작성

낮은 피사계 심도의 입력 영상에 대해서 고속으로 포커스 된 영역의 추출이 가능하다면, 정지 영상뿐만 아니라 동영상을 다루는 여러 가지 멀티미디어 시스템에도 쉽게 확장하여 적용할 수 있다. 이 절에서는 영상에서 고속으로 관심 영역을 검출하기 위해 기존 연구[1,8]에 사용한 모폴로지 필터(morphological filters by reconstruction)를 사용하지 않고, HOS 지도를 블록화하여 이를 바탕으로 영상의 포커스가 존재하는 후보 영역을 고속으로 찾는 방법에 대해 설명한다.

블록화한 HOS 지도를 작성하기 위해서 식 (5)를 이용한 화소에 기반한 정규화 된 HOS 지도를 이용한다. 그림 5-(b)와 같은 M×N 화소 기반의 정규화 된 HOS 지도를 m×n 블록 크기로 나눈 후, 그 블록 내에 존재하는 최대값을 블록화 된 HOS_{block} 지도의 각 블록을 대표하는 값으로 정한다.

$$HOS_{block}(u, v) = \max_{\substack{mu \leq x < m(u+1), \\ nv \leq y < n(v+1)}} [HOS(x, y)] \quad (7)$$

여기서, 블록화한 HOS_{block}의 (u,v) 범위는 $0 \leq u < \frac{M}{m}, 0 \leq v < \frac{N}{n}$ 이다. 블록화 된 HOS 지도를 계산하는 과정은 그림 6과 같다. 그림 6-(a)와 같은 입력 영상에 대하여 3.1절에서 설명한 정규화된 HOS 지도는 그림 6-(b)와 같고, 이 결과를 바탕으로 그림 6-(c)와 같은 m×n의 블록화 된 HOS_{block} 지도를 얻을 수 있다.

그림 6-(b)는 피사계 심도가 낮은 입력 영상의

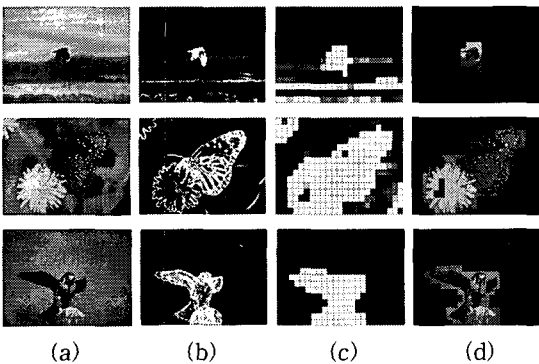


그림 6. 블록화 된 HOS_{block}: (a) 입력 영상, (b) 스케일링 된 HOS 지도, (c) HOS_{block} 결과, (d) (b)의 최대 값을 기준으로 하여 (c)의 HOS_{block} 지도에서 포커스 된 후보 영역을 추출한 결과

HOS_{original}을 식 (6)을 이용해[0, 255]로 스케일을 조절 한 HOS 지도이다. 이러한 HOS 지도에서 관심 영역만을 추출하기 위해서는 관심 영역 주변의 잡음을 제거하는 과정이 필요하다. 기존 연구[1]에서는 관심 물체의 윤곽선을 유지한 채로 잡음 성분만을 효과적으로 제거하기 위해 모폴로지 필터(morphological filters by reconstruction)를 사용했지만, 본 연구에서는 HOS_{original}의 최대값을 특징점(salient point, seed point)으로 삼아 그림 6-(c)의 블록화 된 HOS_{block} 지도에서 대략적인 관심 영역을 고속으로 찾도록 한다.

3.2.3 블록화 된 관심 영역 추출

이 절에서는 기존 연구[1,7,8] 방법과 다르게 잡음을 제거하지 않고 포커스 정보가 포함된 블록화 된 관심 영역을 찾는 방법에 대해서 알아본다. 우리는 3.2.1 절의 HOS 특징을 통해 낮은 피사계 심도 영상 내의 포커스 된 물체는 높은 HOS 값을 가지고, 가장 높은 HOS 값은 포커스 된 관심 영역에 포함되는 것을 관찰하였다. 이를 바탕으로 그림 6-(b)와 같은 HOS_{original}의 최대값 좌표는 영상 내의 최대 HOS 값을 갖는 포커스 된 물체 영역에 포함되고, 포커스 된 물체를 검출하기 위한 기준인 시작점(seed point)으로 정한다.

$$(\hat{x}, \hat{y}) = \arg \max_{(x, y)} [HOS_{original}(x, y)] \quad (8)$$

$$(seed_u, seed_v) = \left(\left(\text{int} \right) \frac{\hat{x}}{m}, \left(\text{int} \right) \frac{\hat{y}}{n} \right) \quad (9)$$

포커스 된 물체 후보 영역을 구하기 위해 영상의 처음 (0,0)부터 순차적으로 검사하여 관심 후보 영역을 찾는 방법(exhaustive search) 대신, HOS_{original} 최대값의 블록 좌표인 시작점 (seed_u, seed_v)을 기준으로 하여 그림 6-(c)와 같은 HOS_{block}(u,v)에서 후보 관심 영역을 구하기 위해 객체 기반 찾기(object-based search) 방법인 4방향(동,서,남,북) 깊이 우선 탐색(DFS: depth first search)을 사용한다. 그래서 고속으로 식 (9)의 결과인 시작점(seed point)에 연결된 흰색 블록 부분을 후보 관심 영역으로 얻을 수 있다. 이와 같이 후보 관심 영역을 찾기 위한 시작점을 제공함으로써 잡음이 존재하는 블록화 된 HOS_{block} 지도에서 포커스 된 물체의 후보 영역만을 효과적으로 검출할 수 있다. 그림 6-(d)는 낮은 피사계 심도 영상에서 포커스 된 물체가 존재할 블록화 된 후보

영역을 추출한 결과이다. 첫 번째와 세 번째에 존재하는 포커스 된 새를 위한 후보 영역은 내부에 잡음인 구멍(hole)을 가지지 않으므로 다음 모듈에서 바로 포커스 된 물체인 새를 추출할 수 있다. 그러나 나비 검출을 위한 후보 영역처럼 포커스 된 물체의 내부질감이 균일한 경우에는 후보 영역 내부에 구멍을 포함할 수도 있으므로 물체를 섬세하게 분할하기 전에 반드시 내부 잡음을 제거하는 과정이 선행되어야 한다.

3.3 관심 후보 영역 내의 잡음(hole) 검사

영상에서 관심 물체가 존재할 후보 영역을 검출할 때, 그림 6-(d)의 나비 그림처럼 구멍이 포함될 경우 포커스 된 물체를 제대로 추출할 수 없다. 그래서 앞 절의 블록단위의 후보 관심 영역을 추출한 후에 그 영역 내에 구멍의 존재 여부를 파악한다. 그래서 만약 이 영역 내에 구멍이 존재한다면 그것을 추적하여 제거하는 과정을 수행한다. 본 연구에서 16×16 블록 단위로 후보 관심 영역을 추출하여 내부 잡음(hole)의 존재 여부를 파악하는 것은 기존 연구 [1]에서 영역 병합을 하기 위해 수립한 경계 조건과 일치하는 것을 알 수 있다. 즉, 그림 7의 흰색 사각형

영역과 같이 블록 단위의 구멍의 경계(boundary) 영역이 추출된 후보 관심 영역에 50% 이상 속하면 그 구멍이 관심 영역에 속한다고 판단해서 구멍을 채우는 연산을 하고, 그렇지 않으면 채우지 않는다[1,7].

이는 블록 단위로 연산하므로 잡음 검사 후 내부의 구멍(hole)을 채우는 연산은 기존 연구[1] 보다 빠르고 간단하다. 그림 8은 블록화 된 HOS 지도를 이용하여 최종 포커스 된 관심 영역을 섬세하게 검출하는 과정을 보여준다. 그림 8-(b)는 블록화 된 HOS_{block} 지도를 나타내고, 그림 8-(c)는 이를 바탕으로 관심 영역 후보를 검출한 결과를 보여준다. 그림 8-(d)는 영역 내에 잡음이 존재할 경우 이를 제거하는 과정을 보여준다. 그림 8-(e)는 최종 관심 영역을 추출한 결과이다.

3.4 포커스 된 관심 영역 추출

낮은 피사계 심도 영상에서 관심 물체가 존재하는 그림 8-(d)와 같은 블록화 된 후보 영역은 주변 잡음에 관계없이 영상의 관심 영역만을 효율적으로 제한하는 마스크 역할을 한다. 그림 9-(a),(c)와 같은 HOS 지도에 해당하는 마스크를 적용하여 관심 영역을 제한한 후, 임계값을 기반으로 그림 9-(b),(d)처럼 주변 잡음은 제거하고 포커스 된 관심 영역만을 나타내는 HOS 지도를 수정할 수 있다. 관심 물체만을 나타내는 HOS 지도를 작성하기 위해 임계값은 100으로 정하였다. 이렇게 포커스 된 관심 영역을 나타내는 HOS 지도는 실제 영상에서 관심 영역을 검출하기 위한 기준으로 사용된다.

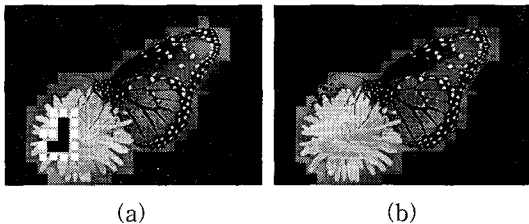


그림 7. 블록화 된 관심 영역 추출: (a) 후보 관심 영역, (b) 구멍 채움 연산 결과

3.4.1 관심 영역 분할

낮은 피사계 심도 영상에서 포커스 된 관심 물체

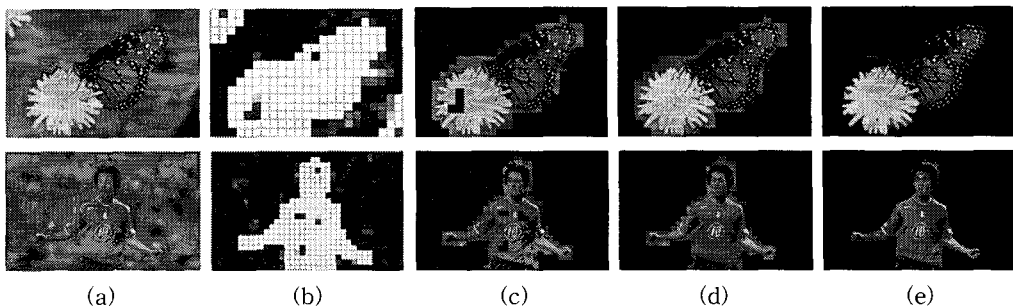


그림 8. HOS_{block}에서 관심 영역을 검출하는 과정: (a) 낮은 피사계 심도 영상, (b) HOS_{block} 결과, (c) 후보 영역 검출, (d) 내부 잡음 제거, (e) 최종 관심 영역

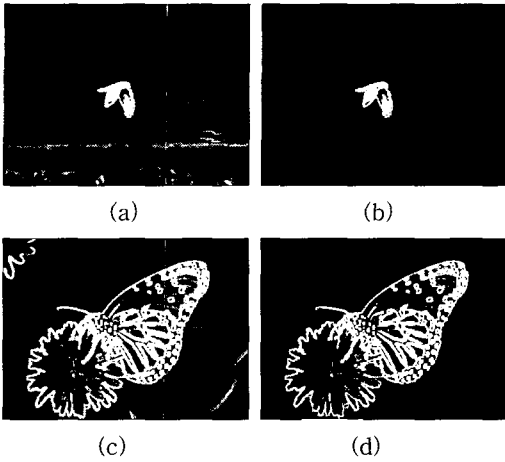


그림 9. HOS 지도에서 관심 영역의 제한: (a),(c) 피사계 심도가 낮은 입력 영상에 식 (6)를 이용하여 계산한 HOS 지도, (b),(d) 그림 7의 결과에 기반하여 기존 HOS 지도에서 관심 영역만 제한한 결과

를 고속으로 추출하기 위해 ‘영역 채우기 기법’(filling-in technique)[14]을 사용하였다. 그림 10은 이 기법을 사용하여 입력 영상에서 관심 영역을 추출하는 방법을 보여준다. 그림 10-(b)는 블록 기반의 관심 영역을 바탕으로 하여 포커스 된 영역만을 나타내는 HOS 지도이다. 이것을 기준으로 하여 입력 영상에 대해 수평방향으로 배경 성분을 제거한 것이 그림 10-(c)이고, 수직방향으로 배경 성분을 제거한 것이 그림 10-(d)이다. 이러한 두 가지 결과를 화소 별로 AND 연산을 수행하면, 그림 10-(e)와 같이 영상에서 포커스 된 최종 관심 영역을 얻을 수 있다.

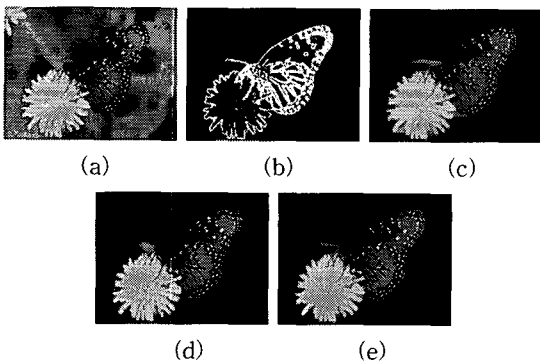


그림 10. 관심 영역의 검출: (a) 낮은 피사계 심도 영상, (b) 관심 물체의 HOS 지도, (c) (b)를 바탕으로 수평 방향의 배경 영역을 제거한 영역, (d) (b)를 바탕으로 수직방향의 배경 영역을 제거한 영역, (e) (c)와 (d)의 결과를 화소 별로 AND 연산하여 얻은 최종 관심 영역

3.4.2 후처리 과정

본 논문에서 제안한 방법을 사용하여 추출한 포커스 된 관심 영역은 잡음도 함께 추출될 수도 있으므로, 이번 단계에서는 에러 출력들을 제거하거나 최소화하고 추출한 관심 영역의 경계 영역을 부드럽게 만드는 연산을 한다. 이를 위해 두 종류의 열림, 닫힘 연산을 하는 단순한 모폴로지 필터를 사용한다. 이러한 필터를 적용하면 관심 물체 분할을 위한 최종 이진 마스크를 얻을 수 있고, 이를 바탕으로 포커스 된 관심 영역에 대한 영상 분할이 가능하다.

4. 실험 결과

본 논문에서 제안한 알고리즘은 384×256 크기의 낮은 피사계 심도의 코렐 데이터베이스 영상에 적용하여 실험하였다. 낮은 피사계 심도의 입력 영상에서 효과적으로 포커스 된 관심 영역을 추출하기 위해 입력 영상의 RGB 색상 채널을 모두 고려한 색상 기반의 HOS 지도를 계산하였다. 식 (6)에서 색상 기반 HOS 지도를 계산하기 위해 n 는 기존 연구[1]와 동일한 3×3 크기의 주변 화소를 고려하였다. 잡음에 강인하게 포커스 된 관심 영역만 추출하기 위하여 화소 단위로 잡음을 제거하는 기존 방법[1]과 다르게, HOS 특징을 이용하여 후보 관심 영역만을 16×16 크기의 블록단위로 찾는다. 이렇게 찾은 후보 영역은 관심 물체의 특징으로 인하여 그 내부에 잡음이 포함될 수 있으므로, 구멍(hole)을 제거하는 연산을 수행하여 관심 물체가 존재할 영역을 찾는다. 블록화 된 관심 영역은 포커스 된 물체와 배경을 포함하게 되는데, 여기에 속한 배경을 제거함으로써 포커스 된 관심 물체만 추출한다. 이러한 영상 분할을 위하여 HOS_{block} 를 이용하였다. 후처리 과정으로 추출한 관심 영역의 잡음을 제거하기 위하여 3×3 크기의 열림, 닫힘 모폴로지 연산을 이용하였다. 그림 11은 본 논문에서 제안한 알고리즘을 낮은 피사계 심도의 영상에 적용하여 포커스 된 관심 영역을 구하는 방법을 단계별로 보여준다. 그림 11-(a)는 낮은 피사계 심도 영상이고, 11-(b)는 HOS 지도를 블록화 된 HOS_{block} 으로 변환한 결과를 보여준다. 그림 11-(c)는 영상의 포커스 영역을 대표하는 특징점(saliency point)을 제공하는 $HOS_{original}$ 을 이용하여 4방향 깊이 기반 탐색(DFS) 알고리즘을 HOS_{block} 에 적용하여 관심 영역

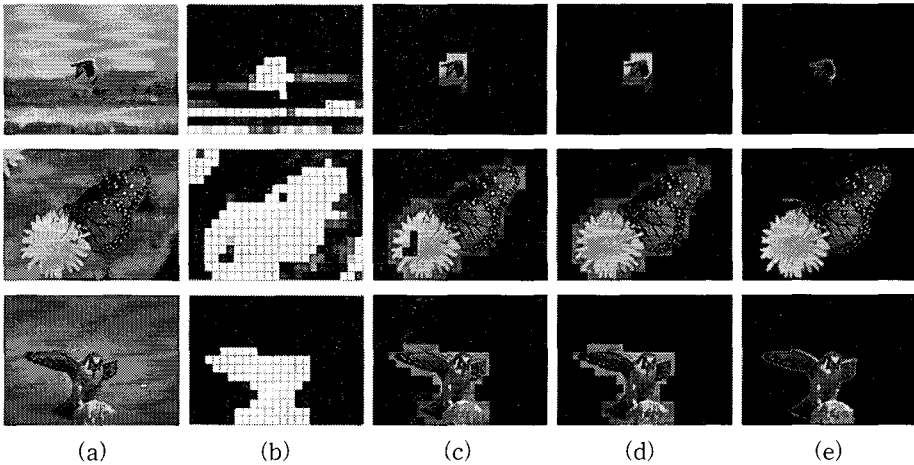


그림 11. 제안된 관심영역을 추출하는 알고리즘: (a) 낮은 피사계 심도 영상, (b) HOS_{block} 지도, (c) 후보 관심 영역 검출, (d) 관심 영역 내부의 잡음 추적 및 제거, (e) 최종 관심 영역 추출

을 구한 결과를 보여준다. 이렇게 구한 관심 후보 영역은 그 영역 내부에 잡음인 구멍(hole)을 포함할 수 있으므로, 이를 추적해서 제거하는 연산을 한다. 그림 11-(d)는 영역 내부의 잡음을 제거하여 관심 영역을 구한 결과이다. 그림 11-(e)는 블록 기반의 관심 영역 내부에 존재하는 배경 영역을 ‘영역 채우기 기법’(filling-in technique)[14]을 사용하여 섬세하게 포커스 된 물체를 검출한 결과를 보여준다.

영상의 잡음에 강인하게 포커스 된 물체만을 추출하기 위하여, 정규화 된 HOS 지도에서 관심 영역을 주변의 잡음을 제거하는 기존의 방법[1] 대신, 정규화 되지 않은 HOS_{original} 특징을 이용하여 영상의 포커스 된 물체가 존재하는 후보 영역을 제한함으로써 주변 잡음에 강인하게 관심 영역의 제한이 가능하다. 또한, HOS 지도를 블록화하여 계산하므로 효율적인 관심 물체의 분할이 가능하다. 실험은 인텔 펜티엄-IV 3.4GHz PC에서 수행되었고, 384×256 크기의 코렐 영상에 대해 제안한 알고리즘의 평균 수행 시간은 약 0.17 초가 소요되는 것을 알 수 있었다. 제안한 알고리즘을 구성하는 각 단계에 해당하는 평균 수행 시간을 표 2에 정리하였다. 표 2를 표 1과 비교하면 본 논문에서 제안한 블록에 기반한 관심 영역의 추출 방법이 약 30배 정도 빠른 것을 알 수 있다. 이는 고속으로 포커스 된 객체의 영상 분할이 가능하기 때문에 피사계 심도 낮은 동영상에 적용 가능함을 보인다. 또한, 포커스 된 객체의 분할 결과의 성능 평가를 위해 기존 연구[1,14]

표 2. 제안한 방법을 이용하여 포커스 된 관심 영역을 추출하는 단계별 평균 연산 시간

384 × 256 크기의 영상에 대해 제안한 방법의 적용 결과	연산시간(ms)	(%)
블록 기반 HOS 지도	110.25	64.6
블록 기반의 후보 관심 영역 추출	2.00	1.1
구멍 추적	1.00	0.6
최종 관심 영역 추출	57.50	33.7
총 계	170.75	100.0

에서 사용한 화소 기반의 측정 방법을 사용하였다. 그림 12는 포커스 된 관심 영역을 추출한 결과를 보여준다. 그림 12-(b)는 입력 영상에서 포커스 된 부분을 수동으로 직접 찾아준 것이고, 그림 12-(c), (d)는 기존 연구[1] 방법과 본 논문에서 제안한 방법을 이용하여 추출한 관심 영역이다. 그림 12-(b)와 같은 관심 영역 모델을 바탕으로 기존 연구 방법과 본 논문에서 제안하는 방법을 통해 추출한 포커스 된 관심 영역 사이의 왜곡 계산을 하면 표 3과 같다. 이를 통해, 제안한 방법의 관심 영역 분할 성능이 기존 연구 방법[1,7]의 성능과 비슷하고 그 속도는 약 30배 정도 향상된 것을 알 수 있다.

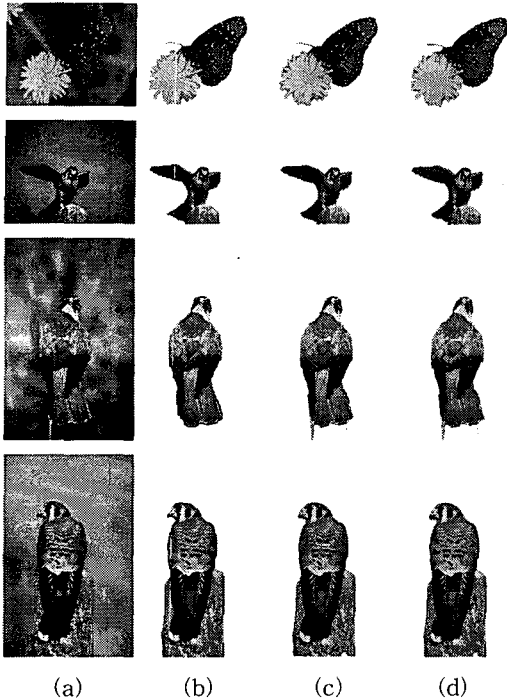


그림 12. 영상 분할 성능 비교: (a) 낮은 피사계 심도의 영상, (b) 수동으로 포커스 된 영역을 제한한 결과, (c) 기존 연구(7)를 이용한 포커스 된 물체 검출 결과, (d) 제안한 방법을 사용하여 관심 영역을 추출한 결과

표 3. 그림 12의 영상에서 추출한 포커스 된 관심 영역의 정확도

영상	기존[7] 알고리즘(%)	제안한 알고리즘(%)
(a)	98.98	98.95
(b)	95.09	95.22
(c)	96.27	96.22
(d)	98.13	98.22

5. 결 론

본 논문은 낮은 피사계 심도 영상에서 블록화 된 HOS_{block} 지도를 사용하여 고속으로 포커스 된 관심 영역만을 추출하는 알고리즘을 제안하였다. 본 논문에서 제안한 알고리즘은 HOS 지도에서 관심 영역을 제외한 주변 잡음을 제거하는 기존방법[1,7] 대신, $HOS_{original}$ 특징을 이용하여 영상의 관심 물체가 존재하는 후보영역만을 고려함으로써 주변 잡음에 강인하고, 영상을 블록 단위로 처리하기 때문에 고속으로 관심 영역 분할이 가능하다. 관심 물체를 추출하는 속도는 영상의 포커스 된 물체의 크기에 영향을 받는

다. 이 방법은 고속으로 동작하므로 낮은 피사계 심도 동영상에 바로 적용 가능성을 보인다.

향후 개선사항으로는 본 논문에서 제안한 영상 객체 분할 시스템을 가상 현실(VR)이나 실감 방송과 같은 분야에 적용하는 것이다. 본 논문에서는 관심 영역을 추출하기 위해 기존 연구[14]에서 제안한 ‘영역 채우기 기법’(filling-in technique)을 사용하였다. 이 알고리즘은 관심 영역의 경계가 모호할 경우 배경을 효과적으로 제거할 수 없는 한계를 지니므로, 현재 배경이 포함된 블록화 된 관심 영역에서 배경을 효율적으로 제거하여 관심 물체만을 추출하는 알고리즘 개발을 진행 중이다.

참 고 문 헌

[1] C. Kim, "Segmenting a Low Depth-of-Field Image Using Morphological Filters and Region Merging," *IEEE Trans. on Image Processing*, vol. 14, issue. 10, pp. 1503-1511, Oct. 2005.

[2] L.M. Lifshitz and S.M. Pizer, "A Multiresolution Hierarchical Approach to Image Segmentation based on Intensity Extrema," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, pp. 529-540, June 1990.

[3] D. Comaniciu and P. Meer, "Robust Analysis of Feature Spaces: Color Image Segmentation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR'97)*, San Juan, Puerto Rico, 750-755, 1997.

[4] J. Pan, S. Li, and Y. Zhang, "Automatic Extraction of Moving Object Using Multiple Features and Multiple Frames," *IEEE International Symposium on Circuits and Systems*, vol. 1, pp. 36-39, May 2000.

[5] C. Gu and M.C. Lee, "Semiautomatic Segmentation and Tracking of Semantic Video Object," *IEEE Trans. Circuits Syst. Video Technology*, vol. 8, no. 5, Sep. 1998.

[6] M. Kass, A. Witkin, and D. Terzopoulos, "Snake: Active Contour Model," *Proceedings of First International Conference on Computer Vision*, pp. 259-269, 1987.

[7] J. Park and C. Kim, "Performance Improvement of Object-of-Interest Extraction from the Low Depth-of-Field Image Using Color-based HOS (high-order statistics)," *신호처리학술대회 (KSPC)*, 2005.

[8] J. Park and C. Kim, "Extracting Focused Object from Low Depth-of-Field Image Sequences," in *Proc. SPIE Visual Communications and Image Processing*, 2006.

[9] W. Cheng, W. Chu, J. Kuo, and J. Wu, "Automatic Video Region-of-Interest Determination Based on User Attention Model," *IEICE Trans. Information and Systems*, E88-D(7), pp. 1578-1586, 2005.

[10] Z. Ye, C. Lu. "Unsupervised Multiscale Focused Objects Detection Using Hidden Markov Tree," in *Proc. International Conference on Computer Vision, Pattern Recognition & Image Processing, 2002 (CVPRIP'2002)*, pp. 812-815, Durham, North Carolina, USA, Mar. 2002.

[11] J.Z. Wang, J. Li, R.M. Gray, and G. Wiederhold, "Unsupervised Multiresolution Segmentation for Images with Low Depth of Field," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 23, No. 1, pp. 85-90, Jan. 2001.

[12] Yim and A.C. Bovik, "Multiresolution 3-D Range Segmentation using Focused Cues," *IEEE Trans. Image Processing*, Vol. 7, No. 9, pp. 1283-1299, Sep. 1998.

[13] G. Gelle, M. Colas, and G. Delaunay, "Higher Order Statistics for Detection and Classification of Faulty Fanbelts Using Acoustical Analysis," in *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics (SPW-HOS '97)*, pp. 43-46, Banff, Canada, July 1997.

[14] C. Kim and J.-N. Hwang, "Video Object Extraction for Object-Oriented Applications," *Journal of VLSI Signal Processing - Systems for Signal, Image, and Video Technology, Special Issue on Multimedia Signal Processing*, Vol. 29, No. 1/2, pp. 7-21, Aug. 2001.

[15] K. Aizawa, A. Kubota, K. Kodama, "Implicit 3D Approach to Image Generation: Object-Based Visual Effects by Linear Processing of Multiple Differently Focused Images," in *Proc. 10th International Workshop on Theoretical Foundations of Computer Vision*, Vol. 2032, pp. 226-237, Germany, Mar. 2000.



박 정 우

2003년 2월 성균관대학교 정보통신공학부 (학사)
 2006년 2월 한국정보통신대학교 (ICU) 공학부 (석사)
 2006년 2월 ~ 현재 삼성전자 기술총괄 생산기술연구소 연구원

관심분야 : HCI, 컴퓨터비전, 기계학습, 유비쿼터스 컴퓨팅



이 재 호

2006년 2월 부산대학교 전자전기통신공학부 (학사)
 2006년 2월 ~ 현재 한국정보통신대학교(ICU) 공학부 석사과정

관심분야 : Medical Image Processing, Multi-view coding.



김 창 익

1989년 2월 연세대학교 전기공학과 (학사)
 1991년 2월 포항공과대학교(POSTECH) 전기전자공학과 (석사)
 1991년 1월 ~ 1997년 7월 SKC Ltd. R&D 센터 선임 연구원

2000년 12월 위싱턴주립대학교 전기공학과 (박사)
 2000년 12월 ~ 2005년 1월 Senior Member of Technical Staff, Epson Palo Alto Laboratory, Epson R&D Inc.
 2005년 2월 ~ 현재 한국정보통신대학교(ICU) 공학부 조교수

관심분야 : 지능형 비디오, 3D 비디오, Next Generation Video Communication Systems, Multimedia Signal Processing, Digital TV Broadcasting, Video Analysis.