

사진의 자동 구도 보정 제시 기법

최용섭*, 박대현*, 김윤*

A Method of Auto Photography Composition Suggestion

Yong-Sub Choi*, Dae-Hyun Park*, Yoon Kim*

요약

본 논문에서는 일반 사용자가 카메라로 사진을 촬영할 경우에 시선을 집중시키면서 안정적인 구도의 영상을 얻을 수 있는 자동 구도 제시기법을 제안한다. 일반 사용자는 대부분 사진의 구도에 대한 배경지식이 없이 사진을 촬영하고, 촬영된 사진은 피사체의 위치가 적절하지 않아 전문가들이 촬영한 안정적인 구도의 사진과 대조된다. 따라서 비전문가 사용자들에게 촬영 후 영상을 처리하는 방법이 아닌 촬영 시 안정적인 구도를 자동으로 제시해주는 방법을 제공한다. 제안하는 방법은 Saliency Map, Image Segmentation, 윤곽선 검출 등을 통해 피사체를 분석하고 피사체를 안정적인 구도가 구성될 수 있는 위치에 황금분할 가이드라인과 함께 출력한다. 실험결과를 통해 피사체를 분석하고 윤곽선을 검출하여 사용자에게 자동으로 구도가 제시되는 것을 알 수 있다.

▶ Keywords : 사진 구도, 관심영역 지도, 이미지 세그멘테이션, 얼굴영역 검출

Abstract

In this paper, we propose the auto correction technique of photography composition by which the eye line is concentrated and the stable image of the structure can be obtained in case the general user takes a picture. Because the general user photographs in most case without background knowledge about the composition of the photo, the subject location is not appropriate and the unstable composition is contrasted with the stable composition of pictures which the experts take. Therefore, we provide not the method processing the image after photographing, but the method presenting automatically the stable composition when the general users take a photograph. The proposed method analyze the subject through Saliency Map, Image Segmentation,

•제1저자 : 최용섭 •교신저자 : 김윤

•투고일 : 2013. 09. 27, 심사일 : 2013. 10. 12, 게재확정일 : 2013. 12. 24.

* 강원대학교 컴퓨터정보통신공학과(Dept. of Computer and Communications Engineering, Kangwon National University)

※ "2013년도 강원대학교 학술연구구조성비로 연구하였음(과제번호-C1009829-01-01)"

Edge Detection, etc. and outputs the subject at the location where the stable composition can be comprised along with the guideline of the Rule of Thirds. The experimental result shows that the good composition was presented to the user automatically.

▶ Keywords : Photography Composition, Saliency Map, Image Segmentation, Face Detection

I. 서론

모바일 기기는 반도체의 나노공정, 저전력 고성능 프로세서 등의 발전과 함께 PC에 버금가는 성능을 제공하면서 많은 사람들이 널리 사용하는 일반적인 기기가 되고 있다. 최근에는 개인이 휴대폰, 태블릿PC 등 다수의 모바일 기기를 사용하는 추세이며, SNS(Social Networking Service)가 보편화되면서 모바일 기기에 탑재되어있는 카메라는 필수기능으로 자리매김하였다.

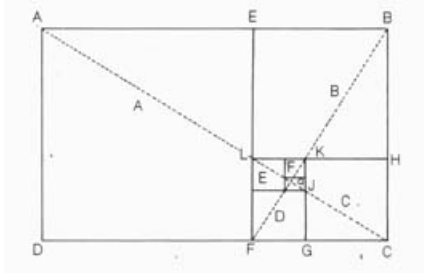


그림 1. 황금분할의 예제.
Fig. 1. Example of golden section.

해상도가 점차 증가함과 동시에 모바일 기기의 카메라 및 디지털 카메라에는 촬영할 때에 얼굴인식, 파노라마 등과 같은 다양한 촬영 지원 기능과 보조 기능들이 제공되고 있다. 이와 같은 기능 중 오래전부터 3등분할 기법(Rule of Thirds)이 제공되고 있지만 대부분의 이용자들은 3등분할 기법의 가이드라인만을 알고 있을 뿐 기능과 역할에 대해서는 모르고 있다. 3등분할 기법은 황금분할(Golden Section)의 개념을 사용하여 카메라로 촬영되는 피사체가 사진 내에서 시선을 이끌고 안정적으로 배치될 수 있도록 보조하는 역할을 한다[1, 2]. 황금분할의 구조는 인간이 만든 특정 조형물이나 자연의 주어진 정적인 상태에 대한 심미적 분석에 유용하나

동적인 상태의 분석에는 한계가 있다. 자연의 동적인 상태, 즉 성장, 발전, 진행 등을 황금분할의 관점에서 분석하기 위해서는 황금나선구조의 이해가 필수적이다. 황금비율을 내재한 직사각형은 아래의 그림 1에서와 같이 정사각형 A, B, C, D, E, F, G ... 등으로 무한히 나눌 수 있다. 이러한 정사각형들 A, B, C, D, E, F ... 등으로의 진행은 이론적으로는 점 Q를 향해 무한소로 진행되어 갈 수 있으며 각각의 사각형들은 서로 황금분할로 분할되어 있다. 그 예로 정사각형 A는 전체 사각형의 나머지 부분(사각형 EBCF)과 황금분할을 이루고 있고 정사각형 B는 사각형 HCFL과 황금분할을 이루고 있다. 황금분할을 내재한 직사각형의 Q를 중심으로 각 정사각형에 내재한 1/4원(호)을 그려 나가면 그림 1과 같은 나선형구조의 호들이 연결된 형태를 보여줄 것이다. 이 호들의 연결된 형태를 황금나선(Golden Spiral)이라 하며 그 진행은 무한대로 뻗어나갈 수 있다. 이 황금나선의 연결된 각 호들의 상호비율을 측정해 보면 황금비율을 내재하고 있는 사실을 쉽게 알 수 있다. 사진에서는 3등분할 기법을 이용하여 이러한 황금비율을 유지할 수 있다. 사진에서 3등분할을 통하여 생기는 교차점의 가까이에 배치한 피사체는 강한 인상을 준다. 또, 교차점에 대해서 경사지게 위치하면 균형이 잡혀 보이고, 교차점 중 3개소에 둔 요소는 대담한 삼각형의 구도를 만든다. 이러한 편안함을 주는 황금비를 사진을 촬영할 때 적용하기 위해 화면을 가로, 세로 3등분하면 두 개의 수직선과 수평선이 나오고 네 개의 교차점이 생기게 되는데 이 교차점 중 한 곳에 피사체를 배치하면 황금분할에 가까운 3등분할의 화면이 구성된다. 아래의 그림 2에서는 사진에서 황금분할을 이용하여 피사체의 위치를 놓을 교차점과 그림 3에서는 이러한 3등분할 기법을 적용한 실제 사진의 예를 보여준다.

본 논문에서는 사진을 찍을 때 이러한 구도가 안정된 느낌을 줄 수 있도록 피사체의 위치를 분석하여 구도를 자동으로 보정하도록 유도함으로써 안정감 있는 사진을 제공할 수 있는 기술을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 사진의 구도 및 사진에서 필요한 데이터를 추출하기 위한 관련연구를 서술한다. 3장에서는 본 논문에서 제안한 사진의 구도를 자동으로 보정하는 기법에 대해 서술하고, 4장에서는 제안한 알고리즘을 적용한 실험결과에 대해 설명한다. 최종적으로 5장에서는 결론에 대해 서술한다.

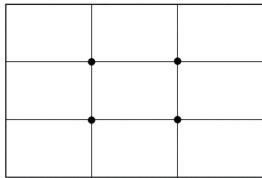


그림 2. 3등 분할과 교차점.

Fig. 2. Rule of third and its intersection.

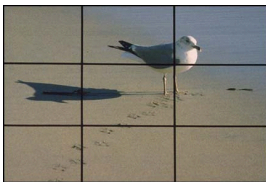


그림 3. 3등 분할 적용 사진 예시.

Fig. 3. Photography example applied the rule of third.

II. 관련연구

디지털 카메라 혹은 모바일 기기에 탑재되어 있는 카메라를 이용하여 찍은 사진을 보면 안정적이지 못한 경우가 대부분이다. 이는 사람의 눈으로 보는 피사체를 사진이라는 시각형 영역 안에 담는 과정에서 피사체의 위치를 안정적이지 않게 느껴 구도를 보정하여 촬영을 해야 하지만 일반적으로 사진을 전문적으로 공부하지 않은 비전문가가 이를 일일이 구도를 잡아서 촬영을 하는 일은 쉽지 않다. 좋은 구도의 사진을 얻기 위해 후처리 방법으로 사진 구도를 위해 saliency map정보를 이용하거나, image retargeting을 사용하여 구도를 안정시키는 방법도 연구되었다[3, 4]. 또한 사진을 좀 더 전문가들과 비슷하게 찍기 위한 기법으로 몇몇의 Photography Composition이 있다. 그 방법들로는 Rule of Thirds, Cropping, Viewpoint, Experimentation, Balancing Elements, Leading Lines, Symmetry and Patterns, Background, Depth, Framing 등과 같은 것이 있다[2]. Balancing Elements에서는 사진내의 객체를 추천하는 비율로 배치하여도 배경이 단조로우면서 공허할 수 있으므로 객체간의 간격이 중요하게 다루어진다. Leading

Lines는 사람의 눈이 사진 속에 존재하는 중심적인 라인을 따라 영상을 이해하는 경향이 있기에 사진을 찍을 때에 고려해야 한다는 것이다. Symmetry and Patterns는 자연영상이나 인공물에 대칭과 무늬가 많이 존재하는데 이와 같은 것은 사람의 시각을 끄는데 효과적이다. 따라서 사진을 찍을 때에 영상에 존재할 수 있는 성질과 패턴을 고려하여 사진을 구성하는 것이다. Background는 객체가 단순한 배경 앞에 있다면 시각을 끄는데 효과적이므로 객체를 강조하고 싶을 경우에는 배경을 단순하게 가져가는 방법이 좋다. Depth는 전경과 배경, 중간 전경의 깊이를 효과적으로 이용하여 원하는 객체를 Depth를 통해서 강조하고 전체적인 분위기를 안정감 있게 가져간다. Framing은 강조하여 표현하기를 원하는 객체를 강조하는 방법이다.

지금까지 서술한 방법들은 사진을 구성할 때에 매우 중요이며, 사람의 눈을 통해 사진을 보는 과정에서 큰 영향력을 가지는 구성요소이다. 하지만 사진을 촬영하는 순간에 디지털 카메라 또는 모바일기기 등에서 이와 같은 요소들을 가미하여 이미지를 보정하기에는 연산능력과 메모리의 제한으로 인해 올바른 결과를 얻을 수 없다. 따라서 사진 구도 중에서 촬영을 통해 이미지를 얻은 다음 후처리로 영상을 처리하여 심미성을 가장 좋게 해주는 방법인 3등분할 기법과 Cropping을 이용하고자 한다. 이 두 가지 방법을 통해 사진을 좀 더 전문가가 느끼는 느낌에 가깝게 구성하고, 사진속의 객체가 사진을 보기 편안한 위치에 놓이도록 함으로써 심미성을 극대화 시킬 수 있도록 처리한다.

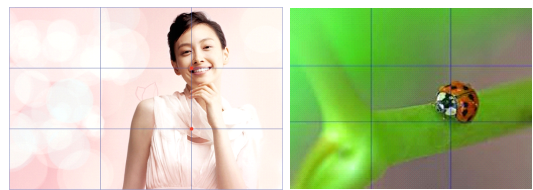


그림 4. 3등 분할 기법 적용 사진.

Fig. 4. Photography applied the rule of third.

위의 그림 4에서와 같이 3등분할 기법을 적용한 두 개의 예제가 있다. 3등분할 기법은 매우 널리 사용되는 사진 구도 방식 중 하나로서 영상은 동일한 너비가 되도록 2개의 세로선과 가로선을 통해 9개의 부분으로 나뉜다. 여기서 가로와 세로 2개의 선분이 교차되는 지점을 power points라고 하며, 전문사진사들은 중심 객체를 power points에 위치시키도록 권장한다[1, 2]. 그림 4에서 두 개의 이미지는 어떻게 3

등분할로 나뉘어져 있으며 피사체가 어디에 위치하고 있는지를 쉽게 알 수 있다. 이러한 구도의 사진은 보는 사람들에게 집중력을 높일 수 있으며 전체적인 사진의 구도에서 편안함을 줄 수 있다.

3등분할 기법이 적용되지 않았을 경우와 적용되었을 경우를 비교한 아래 그림 5의 두 사진을 비교해 보면 피사체를 사진의 정중앙에 위치한 것과 구도보정을 통해 재 위치한 사진이다. 촬영자가 강조하고자 하는 것이 피사체일 경우 구도 보정을 통해 피사체를 위치시킨 오른쪽의 사진에 시선이 더 머물게 된다.

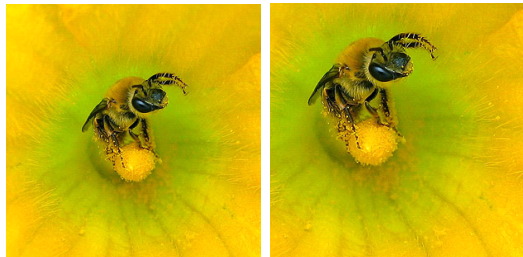


그림 5. 3등분할 기법 적용 전(좌)과 후(우).
Fig. 5. Original(left) and after applying rule of third(right).

3등분할과 함께 적용될 Cropping은 영상에서 관심이 높은 영역을 잘라서 해당영역으로 영상을 구성하도록 하는 전형적인 방법이다. 일반적으로 관심영역(Region of Interest : ROI)에 대한 처리기법으로 널리 사용되어 왔으며 아래의 그림 6은 Cropping이 적용되어 영상에서 ROI를 중심으로 구성한 것이다.



그림 6. Cropping의 예.
Fig. 6. Example of cropping.

아래의 그림 7의 두 사진은 직관적으로 볼 수 있듯이 구도가 올바르지 못함으로서 생기는 불안정한 느낌을 알 수 있다. 영상을 3등분할 기법에 맞춰 재구성함으로써 심미성을 높이는 방법과 함께 이러한 구도를 안정적이고 세련되게 자동으로 보정하는 것이 본 논문의 목적이다.



그림 7. 구도가 불안한 사진의 예.
Fig. 7. Photography example of unstable composition.

이미지의 segmentation[5-11]은 의학용 이미지분석이나 이미지 수정과 같은 응용에서 이미지 처리를 하는 과정에서 사용된다. 컬러이미지의 segmentation기법으로 interactive한 Grow Cut[6]방식에 대해 살펴본다. 이러한 과정은 그림 8에서와 같다. 각각 (a) 원본 이미지, (b) 사용자 지정 분할시작점, (c) 분할된 영상을 볼 수 있다.

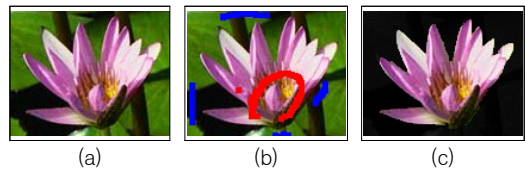


그림 8. 컬러이미지의 세그멘테이션.
Fig. 8. Segmentation of color image.

Grow Cut방식에서는 사용자가 설정한 영역을 기준(seed pixel)으로 'bacteria'가 퍼지기(grow)시작하여 전체 이미지의 영역으로 퍼지게 된다. 아래의 그림 9에서와 같이 여러 과정을 반복하면서 각 cell은 이웃하고 있는 cell에 'attack'을 시도하면서 segmentation과정을 진행하게 된다.



그림 9. 박테리아의 진화 단계.
Fig. 9. Bacteria evolution steps.

Grab Cut[7]은 graph cut 방식을 기반으로 한 이미지 segmentation 기법으로 영상에서 분리해내고자 하는 객체에 대한 초기 정보를 사용자가 입력해주면 해당하는 범위 내에서 객체를 정확하게 분리해주는 알고리즘이다. 사용자가 객체 영역에 사각형의 윈도우를 설정하여 알고리즘을 실행하게

된다. 대부분의 segmentation 기술은 이미지에 포함된 edge와 영역정보를 사용한다. Grab Cut은 이미지에 포함된 영역과 경계정보를 segmentation을 수행하는데 있어서 혁신적인 segmentation기법이다. Grab Cut은 segmentation수행을 위해 graph cut방식을 이용한다. 대부분의 segmentation 기술과 마찬가지로 Grab Cut 역시 이미지내의 정보를 캡슐화하여 사용한다. 일반적인 segmentation 기술은 edge 정보와 영역정보 중 하나만 사용하지만 Grab Cut방식은 이들 모두를 사용하여 최적의 segmentation 결과를 찾아낸다. Segmentation을 수행하기 위해 이미지의 픽셀을 표현하는 곳에 하나의 그래프가 그려지고 추가로 두 개의 특별한 node인 sink node와 source node가 함께 생성된다. 아래의 그림 10에서와 같이 이미지 내에서 배경과 전경을 labeling하여 구분한다.

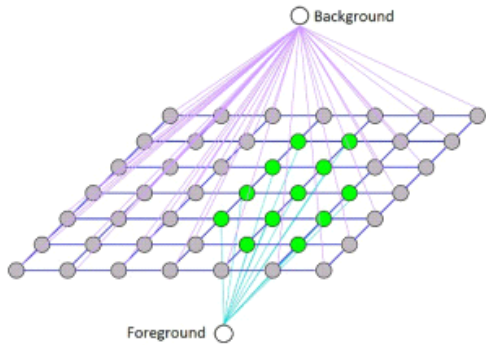


그림 10. 이미지에서의 배경과 전경.
Fig. 10. Background and foreground on image.

아래의 그림 11에서는 사용자가 직접 영역을 지정하여 배경과 전경을 segmentation 하는 과정을 볼 수 있다.

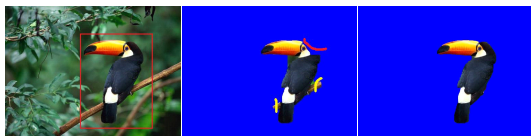


그림 11. Grab Cut알고리즘을 이용한 이미지 Segmentation의 예.
Fig. 11. Image segmentation using grab cut.

Min Cut영상에서 색상을 Segmentation하기 위해 영상의 각 픽셀을 의미하는 각각의 node에서 거리에 대한 가중치 (w)를 부여하여 clustering하는 기법 중 Minimum Cut[8] 방식은 영상의 색상 값에서의 거리를 이용하여 가중치 w 에

대하여 Adjacency Matrix와 Weight Matrix를 작성하여 이러한 weight를 기준으로 색상을 분리한다.

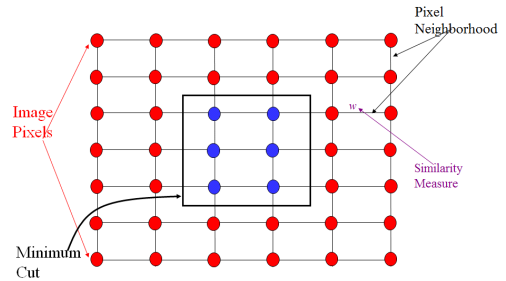


그림 12. 이미지 세그멘테이션과 Minimum cut.
Fig. 12. Image Segmentation & Minimum Cut.

그림12는 Minimum Cut으로 분리된 색상을 clustering 하는 과정을 보여준다. 이렇게 분리한 Minimal Cut은 아래의 그림 13와 같이 색상의 유사도에 따라 clustering을 반복하여 일정한 cluster를 이루는 블록으로 나누게 된다.

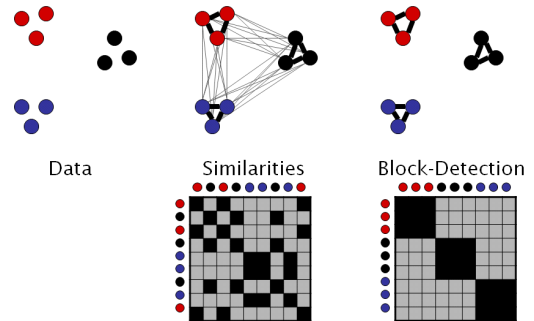


그림 13. Minimum cut을 통한 클러스터링.
Fig. 13. Finding the Minimum Cuts : clustering overview.

III. 제안하는 알고리즘

본 논문에서는 입력영상을 받아서 입력영상에 대해 Saliency Map[12-15]을 구성하고, Color Segmentation[9]을 통해 영상 내에서 사람의 눈에 보이지 않는 주요한 ROI를 찾는다. 이러한 결과에 본 논문에서 제안하는 시나리오(Scenario)를 적용시켜 각 시나리오의 조건에 맞도록 3등분할 기법을 적용하여 Cropping함으로써 각 영상의 성격에 따라 각기 다른 구도 보정을 통해 안정성을 극대화시킨 영상을 얻을 수 있다. 그림 14은 제안하는 알고리즘의 전체 흐름도를 보여준다.

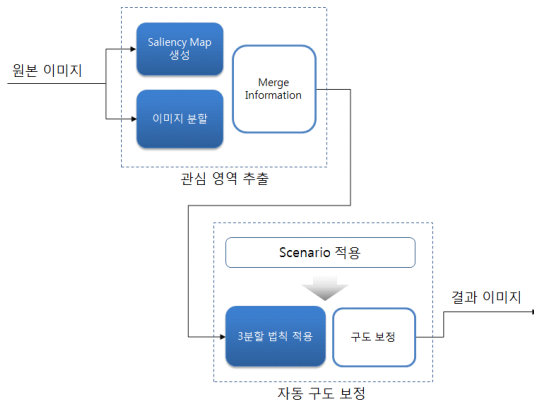


그림 14. 전체 알고리즘 흐름도
Fig. 14. Flowchart of whole algorithm.

Saliency Map을 구성하는 방법은 사용자가 관심을 가진 영역을 찾는 하향식 접근 방법과 영상에 대한 사전지식 없이 주어지는 시각자극의 특징을 기반으로 관심영역을 추출하는 상향식 접근 방법이 있다.

본 논문에서는 영상의 Color, Luminance, Motion, Texture, Orientation 등 관심영역의 perception에 영향을 주는 다양한 feature중에서 사용하기에 가장 적합한 feature를 사용하여 Saliency Map을 도출한다. 대부분 존재하는 Saliency Map을 구하는 방법은 다음과 같다. 첫 번째 단계는 특징을 추출하는 것이다. 즉 다양한 Low-Level Feature Map을 구성한다. 다음 단계는 Saliency Map을 계산하는 것이다. Saliency Map은 앞의 Feature Map들로 계산된다. 마지막으로 계산으로부터 핵심적인 영역이 관심영역으로 설정된다. 이러한 접근들은 기본적으로 Biological에 기초하거나 Computational에 기초하여 접근한다. Biological적인 접근방법은 밝기, 색, 방향의 Feature Map을 구성하여 Bottom Up 방식으로 결합하거나 합하여 계산한다. Computation을 기초로 접근하는 방식은 기본적으로 Feature Map들을 사용하지만 정규화를 Graph에 기초하여 접근한다. 이것은 부분적으로 Biological한 방법과 Computational한 방법을 조합한 것이다. 다양한 feature들에 대하여 다중 피라미드 영상을 구성하고, 영상의 상호간의 차이(Center Surround Difference)를 구하면 특징 값들이 나타내는 영상의 값을 확인할 수 있다. 이 값들은 영상에 존재하는 feature에 따라 다양한 값으로 나타나므로, 하나의 값만을 이용하기 보다는 다중으로 결합하여 최적의 Saliency Map을 도출해야 한다. 본 논문에서는 복합적인 feature를 사용하여 각각의 특징을 반영한 map을 생성하며, 자동 구도

보정 알고리즘에 가장 적절한 Saliency Map을 이끌어낸다. 아래의 그림 15에서는 Saliency Map의 일반적인 도출과정의 알고리즘을 보여준다.

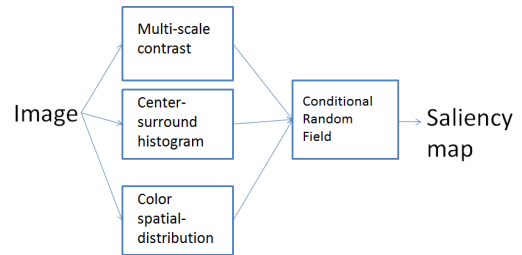


그림 15. 관심영역 지도 알고리즘.
Fig. 15. Saliency map algorithm.

가장 사람의 시각에 영향을 주는 요소인 밝기 값은 국부적인 밝기가 Saliency Map을 구성하는데 주요 기준이 되기 때문에 아래의 식 (1)에 의해서 명암도 영상을 만들고 이 영상이 Feature Map으로 사용한다.

$$I = (r + g + b) / 3. \tag{1}$$

색상정보를 서로 전달해 주는 인간의 신경회로는 3가지 추상체들에 의한 정보를 '적/녹', '황/청'의 반대 쌍의 색상정보로 바꾸어 전달한다. 입력영상에서 색상 Feature Map을 얻기 위해 RGB 컬러모델의 적색, 녹색, 청색을 r, g, b라고 할 때 R(적색), G(녹색), B(청색), Y(황색)의 4개의 채널을 만들어낸다. 본 논문에서는 위 4개의 채널을 사용하여 두 개의 색상 Feature Map을 아래의 식 (2)를 통해 만든다.

$$\begin{aligned} RG(x,y) &= R(x,y) - G(x,y), \\ BY(x,y) &= B(x,y) - Y(x,y). \end{aligned} \tag{2}$$

여기서 RG와 BY는 '적/녹', '황/청'의 반대쌍의 색상정보를 의미한다. 각각의 Feature Map은 주변으로부터 가장 두드러지게 드러나는 부분이 강조되는 특징을 가지게 되는 Conspicuity Map으로 변환된다. 아래의 식은 Conspicuity Map을 얻기 위한 Gaussian Pyramid를 수행하는 과정 식 (3)을 보여준다.

$$\begin{aligned} P_j(0) &= F_j, \\ P_j(i) &= F_j(i-1) * G. \end{aligned} \tag{3}$$

각각의 특징을 F라 하면, 9개의 크기 가우시안 피라미드 P는 Feature Map F에 Gaussian Filter G를 사용하여 점진적으로 subsampling하고 lowpass filtering하여 얻는다. 각각의 특징 P에 대해서, 6개의 중간 다중 크기 Conspicuity Map들을 아래의 식 (4)에 의해서 구한다.

$$\begin{aligned} M_{1,1} &= |P(2) - P(5)|, M_{2,2} = |P(2) - P(5)|, \\ M_{3,3} &= |P(3) - P(6)|, M_{4,4} = |P(3) - P(7)|, \\ M_{5,5} &= |P(4) - P(7)|, M_{6,6} = |P(4) - P(8)|. \end{aligned} \quad (4)$$

유일한 특징을 가지는 아래의 식 (5)을 통해 얻는다.

$$C_j = \sum_{k=1}^6 \mathcal{N}(M_{j,k}). \quad (5)$$

시각적 주의 모델은 각각의 Conspicuity Map에 대한 가중치를 결정하고 조합하여 Saliency Map을 구성한다.

$$S = \sum_{i=1}^n w_i C_i. \quad (6)$$

여기서 n은 Conspicuity Map들의 수이다. 본 논문에서는 Saliency Map을 구성하기 위해 사용되는 Conspicuity Map의 중요성에 따라 동적 가중치인 w를 동적으로 결정한다. 아래의 그림 16에서는 배경과 객체가 뚜렷한 색상의 차이를 보일 때에 성능이 좋은 LUV공간으로 변환하여 mean difference를 Saliency Map과 앞선 과정에서 기술한 방법을 이용하여 밝기 값을 가지고 center surround difference를 통해 도출한 Saliency Map의 일반적인 예를 보여준다.

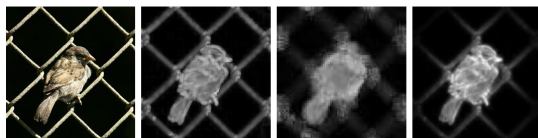


그림 16. 각 단계별 Saliency Map의 예제.
Fig. 16. Example of saliency map at each pass.

사진영상에서 나타나는 관심영역 중에 가장 중요한 관심영역은 바로 사람의 얼굴이다. 따라서 사람의 얼굴은 무엇보다도 관심영역을 중에서 높은 값을 가져야 하므로, 별도로 Face Detection[16-18]을 수행하여 사람의 얼굴영역을 검

출한다. 다음은 Face Detection을 통해서 얼굴 영역을 검출하는 과정을 보여준다.

이미지에서 사람의 피부색을 검출하는 기술은 얼굴 영역 검출하는 기법에 있어서 매우 중요하고 유용한 방법이다. 이러한 skin color를 추출하기 위해 input image가 RGB방식일 경우 YCbCr색상 모델로 변환하여 사용한다. RGB색상 모델의 경우 빛 상태가 변화에 따라 skin color를 검출하는 과정에서 오류가 발생할 확률이 높기 때문이다. RGB 색상 모델은 아래의 식 (7)을 통해 YCbCr방식으로 변환시킬 수 있다.

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B, \\ Cb &= -0.169R - 0.332G + 0.500B, \\ Cr &= 0.500R - 0.419G - 0.081B. \end{aligned} \quad (7)$$

검출된 skin color의 skin pixel의 분산형태와 Cb, Cr의 histogram의 분산형태를 아래 그림 17에서와 같이 확인 할 수 있다.

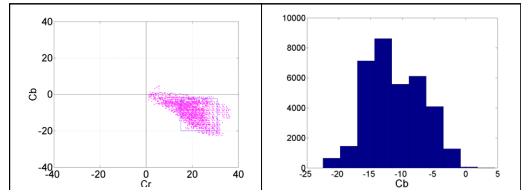


그림 17. YCbCr color에서의 Skin pixel과 Cb.
Fig. 17. Skin pixel and Cb on YCbCr color model.

얼굴영역을 검출하여 Color Segmentation을 통해 결과 이미지를 얻어내고, 이렇게 얻어진 결과 이미지의 small region들을 제거하여 만들어진 binary image를 Edge Detection기술을 이용하여 기존에 얻어진 이미지와 통합하는 과정을 거치면 edge영역에 속하지 않은 작은 영역을 제거할 수 있다. 작은 영역의 부분을 제거한 결과 이미지를 이용하여 skin color로 판단되는 위치의 좌표를 선택하고 Color-Based Segmentation 기법을 이용하여 직사각형의 형태의 영역을 지정한다. 이 때 각 선택영역은 고유의 window size를 가지고 있다.

Skin color로 검출된 영역 중 얼굴아래에 위치한 목이나 팔, 손과 같은 영역의 경우 Box-Merge algorithm을 이용하여 인접한 box영역의 경우 큰 box의 영역으로 합쳐주는 과정을 거친다. 그러한 결과를 아래의 그림 18에서 보여준다.

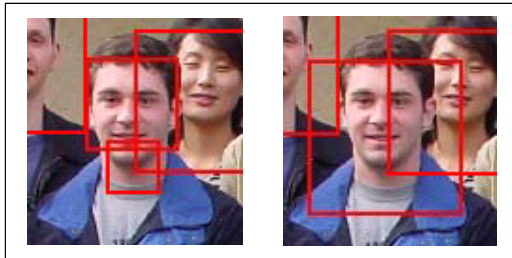


그림 18. 인접한 박스들과 병합된 이미지 영역.
Fig. 18. Image Selection : Merging of Adjacent Boxes.

이렇게 해서 검출된 얼굴 영역에 Image Matching Algorithm을 적용하여 각 영역에 rank를 부여하고 이 중 거리와 window size를 감안하여 rank값이 다른 영역의 rank에 비해 유사성이 떨어지는 영역을 제거하여 결과적으로 얼굴 영역을 검출할 수 있다. 아래의 그림 19에서는 rank를 감안하여 얼굴의 유사성과 관련이 없는 영역을 제거한 결과를 보여준다.



그림 19. 지리적 고려사항을 통한 연관성 순서들.
그림 19. Correlation Ranking after Geographical Consideration.

Saliency Map을 이용하여 영상에서 관심영역을 찾고 일정 영역 크기의 블록에 대해 local maximum을 찾았지만, 실제적으로 관심영역이 얼마나 되는지를 판단할 기준이 없다. 아래의 그림 20에서는 다양하게 설정되는 ROI를 표현한 것이며, local maximum을 중심으로 설정된 것과 실제 객체를 정확하게 ROI로 잡은 영역을 박스로 보여주고 있다.

이와 같이 올바른 ROI를 설정하기 위해서는 Saliency Map에서 ROI를 도출하는 방법이 필요하다. 본 논문에서는 Color Segmentation기법과 Saliency Map의 두 가지의 결과를 융합하여 ROI를 검출한다.



그림 20. ROI 후보영역.
Fig. 20. Candidate region of ROI

Color Segmentation은 색상간의 밀집도와 집합도 정도를 이용하여 유사한 색상들을 하나의 그룹으로 모아주어 영상의 영역을 분할하는 방법이다. 아래의 그림 21은 3개의 cluster를 기준으로 Color Segmentation이 된 결과를 보여주고 있다.



그림 21. Color Segmentation의 결과.
Fig. 21. Result of color segmentation.

Color Segmentation은 색상간의 밀집도와 결합정도도 이용하여 segmentation하는 방식으로 본 논문에서는 색상의 분리를 임의로 k개의 평균점을 찍은 후, 그 점들로부터 가까운 것들로 k의 그룹을 나누고, 다시 각각의 그룹들 안에서 새로운 평균점을 정하여 그룹을 나누는 것을 반복하는 K-Means Clustering Algorithm[19]을 적용시켜 이러한 cluster들을 segmentation하였다.

K-Means algorithm은 n개의 객체들의 집합을 k개의 군집으로 분해하는 거리에 기반을 둔 clustering기법이다. 이러한 군집의 유사성은 군집에서 군집의 무게중심으로 볼 수 있는 객체들의 평균값을 측정하여 기준점에서 가까운 곳의 데이터들을 하나의 군집으로 묶는다. 아래의 그림 22에서는 3개의 군집으로 clustering하는 수행과정을 보여준다.

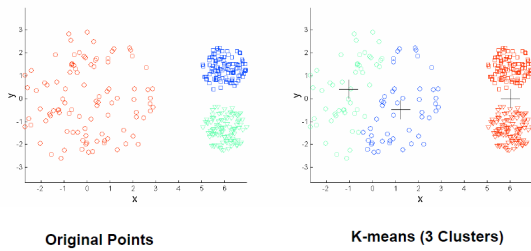


그림 22. K-means Algorithm의 수행과정.
Fig. 22. Process of K-means Algorithm.

1. 임의의 k개의 군집수를 결정하고, 각 군집에 초기치 또는 군집 중심을 1개씩 할당하여 위치를 설정한다.
2. 각각의 데이터에 대해 k개의 위치까지의 거리를 구하고 가장 가까운 군집에 소속시킨다(유클리디안 거리를 이용).
3. 군집으로 나누어진 데이터를 기준으로 새로운 군집 중앙의 위치를 최솟값을 갖도록 재설정한다.
4. 새롭게 구한 군집 중앙의 위치가 기존과 동일하면 알고리즘을 종료하고 다르면 두 번째부터 재수행한다.

이러한 알고리즘에 의한 영상은 cluster 개수 값인 k에 의해 결과가 상이하게 나타날 수 있다. 즉, 자동 구도 보정에서 사용하기 위해서 적절한 상수의 값을 선택해야하며, Saliency Map과 결합하여 최적의 ROI를 도출 할 수 있는 값을 이끌어내야 한다. 아래의 그림 23은 상수 k값에 따른 영상의 결과이다.



그림 23. k=2 일 때 Algorithm의 반복 횟수별 이미지.
Fig. 23. Each iteration image applying algorithm at k=2.

cluster의 개수가 2개일 때 이미지의 edge를 구별할 수 있도록 확연하게 식별할 수 있었으며 이에 따라 본 논문에서는 자동 구도 보정에 사용하기 위한 적절한 값으로 k=2로 설정하여 수행하였다. Color Segmentation을 수행한 결과는 Saliency Map결과와 mapping하여 최적의 ROI를 찾게 된다. 아래의 그림 24에서는 Color Segmentation과 Saliency Map을 mapping한 결과를 보여준다.

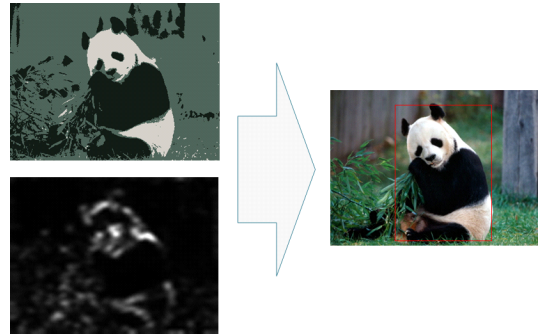


그림 24. Saliency Map + Color Segmentation의 mapping결과 ROI.
Fig.24. ROI of combining saliency map and color segmentation.

본 논문에서 제안한 방법으로 영상에서 ROI를 구한 결과를 이용하여 Rule of Thirds의 기준에 적합하도록 cropping 과정을 거치게 된다. 단순히 이때 얻은 ROI의 위치에 가장 가까운 교차점만을 찾아서 cropping하였을 경우 의도하지 않았던 구도의 보정된 결과 영상을 얻는 경우가 있다. 따라서 영상이 본질적으로 가지는 요소들을 최대한 보존하기 위해 각 영상의 특성에 맞는 시나리오를 적용하여 구도 보정을 위한 cropping을 진행하여야 한다.



그림 25. 시나리오가 적용되지 않은 예제.
Fig. 25. Example without scenario.

또한, 얻어진 ROI에 대해 영상의 가로와 세로의 크기를 대조하여 비율을 측정하여 영상이 가질 수 있는 안정성과 심미성을 최대한 고려하여 Rule of Thirds를 적용한다. 하지만 영상은 객체의 종류, ROI의 크기, 사람의 얼굴 인식 영역 등의 다양한 요소에 의존적이다. 따라서 다양한 영상의 적절한 구성의 결과영상을 가져오기 위하여 여러 가지 scenario를 기반으로 영상을 cropping하는 방법을 사용한다. 그림 25는 Face Detection으로 올바르게 얼굴 영역을 ROI로 선택하였지만 신체가 결과 영상에 전부 포함되지 않은 문제가 발생한다. 따라서 이와 같은 영상은 적절한 시나리오를 선택하여 처리해주어야 한다.

Face Detection의 경우 Rule of Thirds에 맞게

cropping하지만 신체 영역을 보호하여 사진의 안정성을 높이는 방법을 사용한다. 아래 그림 26은 일반적인 Rule of Thirds가 적용된 영상과 시나리오가 추가적으로 적용된 영상의 차이를 보여준다.



그림 26. 원본영상, 시나리오가 적용되지 않은 영상, 시나리오 적용영상.
Fig. 26. Original, without scenario, with scenario.

영상에서 2개의 ROI나 얼굴이 검출되었을 경우에는 두 얼굴의 거리와 영상의 사이즈를 비교하여 처리하는 방법을 사용한다. 두 얼굴이 근접하게 붙어있을 경우에는 두 얼굴이 가지는 ROI의 평균점을 찾아서 적용해주며, 두 얼굴이 영상 내에서 멀리 떨어져 있을 경우 영상 내에 두 얼굴을 보존하면서 자동 구도 보정을 적용한다. 아래의 그림 27에서는 이와 같이 검출된 ROI가 떨어져있는 경우 근접한 경우의 사진의 예를 보여준다.



그림 27. 얼굴이 떨어져 있는 경우와 근접한 경우.
Fig. 27. far-off and near between faces.

영상 내에서 주요하게 작용하는 지평선과 같은 선분을 검출하였을 경우 선분을 Rule of Thirds로 분할된 선분에 위치시킴으로써 영상이 가지는 심미성과 안정성을 극대화 한다. 이 방법의 경우 영상 내에 독점적으로 주요하게 나타내는 선분을 찾아 영상을 최대한 보존하면서 보정 과정을 적용해야한다.

아래의 그림 28은 수평선의 선분을 검출하여 원본이미지와 시나리오를 적용시킨 결과이미지를 비교하였다.

앞서 서술한 바와 같이 사람의 얼굴이 가장 중요한 ROI이지만 사람의 얼굴이 영상의 전반적인 영역을 차지하고 있을 경우에는 영상을 ROI에만 맞춰서 cropping하게 되면 부자연스러운 영상 결과를 얻을 수 있다. 따라서 이와 같은 경우는 영상을 그대로 유지하는 방법을 사용하여 영상의 안정성을 도모한다.

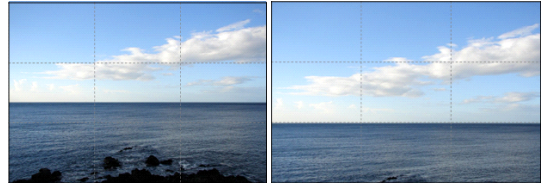


그림 28. 수평선을 3등분할 법칙에 적용한 시나리오.
Fig. 28. Applying rule of third algorithm to horizontal line.

아래의 그림 29는 이러한 시나리오가 적용되지 않아서 과도하게 cropping된 결과이미지를 보여준다.

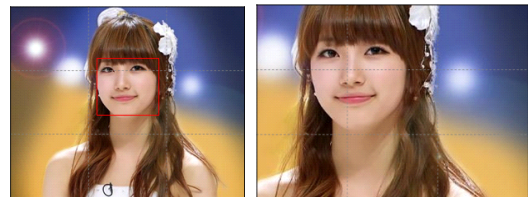


그림 29. 시나리오가 적용되지 않은 경우의 과도한 cropping.
Fig. 29. Excessive cropping without scenario.

앞에서 제시한 자동 구도 보정의 기술을 이용하여 안정적인 구도의 사진을 찍을 수 있도록 카메라 디스플레이 창에 피사체의 위치를 표시한다. 카메라 디스플레이 창에 피사체의 위치를 표시하기 위해서는 saliency map에서 객체를 찾아낸 후에 아래의 그림 30에서와 같이 피사체의 윤곽선까지 검출해야 한다.



그림 30. 피사체의 윤곽선 검출.
Fig. 30. Extraction of object contour.

그림 31에서는 카메라로 실제촬영 시 카메라의 구도보정을 유도하도록 피사체의 윤곽선을 검출하여 촬영의 구도를 보정하도록 구도에 맞는 피사체의 위치를 윤곽선으로 제시하여 주는 예를 보여준다.



그림 31. 사진 촬영 시 자동 구도 제시의 예.
Fig. 31. Auto composition suggestion on photography shooting.

피사체의 윤곽선을 검출하기 위해서 스네이크(Snake)알고리즘으로 알려진 Active Contour Model(ACM)[20]을 사용한다. 스네이크 알고리즘은 수학적 연산을 통해 이미지 곡면위에서 에너지를 최소화하는 곡선으로 구해지며, 객체의 윤곽선을 표현하는 포인트들이 움직이는 객체의 윤곽선을 지속적으로 따라 가도록 제어하여 배경과 피사체를 분할 할 수 있다. 따라서 ACM은 기본적으로 영역이 움직리드는 형태로 나타나며, 피사체의 영역인 ROI영역에 대해 ACM을 적용하면 ROI영역은 피사체의 형태로 변형되면서 segmentation 된 결과를 얻을 수 있다. Snake알고리즘으로 물체의 외곽을 추출한 후 앞에서 언급한 rule of thirds에 의해 화면에 표시 함으로써 사용자가 보다 좋은 구도의 사진을 찍을 수 있도록 유도한다.

IV. 사진의 자동 구도 제시 결과

자동 구도 제시의 결과에 대한 실험을 위해 여러 샘플 이미지들에 대하여 Saliency Map, Color Segmentation 과정 및 윤곽선 검출 기법을 통하여 자동 구도 제시를 위한 윤곽선이 추가된 이미지를 비교하였다. 실험결과 자동 구도를 제시한 위치에 피사체의 윤곽선이 알고리즘 적용 후의 이미지에 나타난 것을 확인할 수 있다.

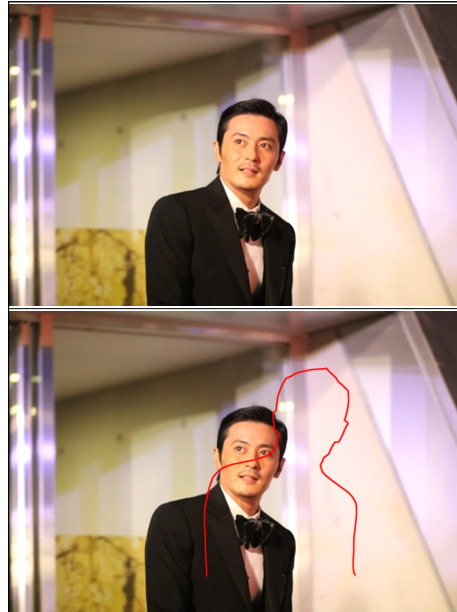


그림 32. 첫 번째 실험 결과 이미지.
Fig. 32. Experiment result image.



그림 33. 두 번째 실험 결과 이미지.
Fig. 33. Experiment result image.

그림 34에서는 본 논문에서 제안한 자동 구도 제시 방법을 통한 결과이미지가 3등분할 법칙이 잘 적용되었는지를 비교하기 위하여 추가적으로 3등분할 교차점의 위치 확인을 위한

격자 선을 적용한 이미지를 보여준다. 실험 결과 원본이미지의 피사체에 대한 윤곽선이 알고리즘 적용 후 결과이미지에서 3등분할 법칙이 적용된 위치에 정확하게 표시 되는 것을 볼 수 있다. 이를 활용하여 실제 카메라를 이용한 사진 촬영 시에 사용자에게 좀 더 나은 구도를 위한 피사체의 위치를 윤곽선을 통하여 카메라의 디스플레이에 제시할 수 있다.

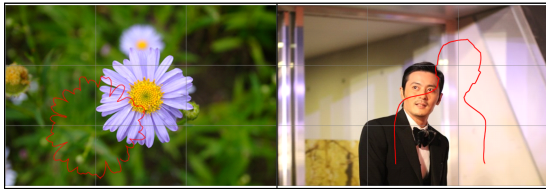


그림 34. 실험 결과에 대한 분석 이미지.
Fig. 34. Analyzing image of experiment result.

V. 결론

본 논문에서는 영상에 대한 편집을 통한 구도보정이 아닌 실제 디지털 카메라나 모바일기에서 바로 구도를 보정하여 촬영을 할 수 있는 자동 구도 제시 기술을 제안하였다. 카메라로 촬영하는 일반적인 사용자들은 전문 사진가들이 권장하는 3등분할 가이드라인이 제공됨에도 불구하고, 대부분 사진의 구도에 대한 배경지식이 없다는 문제점이 있다. 사진의 주요 객체를 교차점에 위치시켜 3등분할로 촬영된 영상은 더욱 시선을 끌고 안정된 균형감을 제공한다. 따라서 3등분할에 맞춰 피사체가 놓일 위치를 사용자에게 알려주기 위해 본 논문에서 제안한 구도보정 기술과 Edge Detection 기법을 활용하여 자동 구도 제시 기술을 구현하였다. 자동 구도 제시 기술을 통해 사용자는 카메라로 촬영할 때에 3등분할의 교차점 가운데 가장 근접한 위치에 피사체를 위치시켜 시선을 집중시키고 안정적인 영상을 얻을 수 있다. 제안하는 방법은 원본이미지와 구도를 보정할 수 있는 피사체의 윤곽선을 이용하여 자동으로 구도를 제시하여 사용자가 전문 사진가들에 가까운 안정된 구도의 사진을 촬영할 수 있도록 하여 결과영상의 주관적 품질과 안정된 균형감을 높일 수 있다.

참고문헌

[1] Grill T., Scanlon M, "Photographic Composition," Watson-Guptill, 1990.
[2] Krages B, "Photography: The Art of Composition,"

Allworth Press, 2005.

- [3] L. Liu, R. Chen, L. Wolf, and D. Cohen-Or, "Optimizing photo composition," *Comput. Graph. Forum*, Vol. 29, No. 2, pp. 469-478, 2010.
- [4] J. Chen, H. Zhao, Y. Han and X. Cao, "Visual saliency detection based on photographic composition," *proceedings of ICIMCS*, pp. 13-16, ACM New York, 2013.
- [5] X. Bai and G. Sapiro, "A geodesic framework for fast interactive image and video segmentation and matting," *ICCV 2007*, pp.1-8, 2007.
- [6] V. Vezhnevets and V. Konouchine, "Grow-Cut - Interactive Multi-Label N-D Image Segmentation," *Proc. Graphicon*. pp.150-156, 2005.
- [7] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: Interactive foreground extraction using iterated graph cuts," *ACM Trans. on Graphics*, Vol.23, pp.309-314, 2004.
- [8] A. Delong and Y. Boycov, "A scalable graph-cut algorithm for n-d grids," In *Proceedings of CVPR*, pp.1-8, 2008
- [9] C. Pantofaru and M. Hebert, "A Comparison of Image Segmentation Algorithms," *Tech. Report CMU-RI-TR-05-40*, Robotics Institute, Carnegie Mellon University, September, 2005.
- [10] J. Wang and M. F. Cohen. "An iterative optimization approach for unified image segmentation and matting," In *Proceedings of IEEE ICCV 2005*, pp.936-943, 2005.
- [11] S. Y. Kim, "Color image segmentation based on edge salience map and region merging," *Journal of the Korea society of computer and information*, Vol. 12, No. 3. pp.105-113, 2007.
- [12] Y.-F. Ma and H.-J. Zhang. "Contrast-based image attention analysis by using fuzzy growing," In *ACM International Conference on Multimedia*, 2003.
- [13] X. Hou and L. Zhang. "Saliency detection: A spectral residual approach," *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[14] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu. "Global contrast based salient region detection," In CVPR, pp.409-416, 2011.

[15] J. Harel, C. Koch, and P. Perona. "Graph-based visual saliency," Advances in Neural Information Processing Systems 19, MIT Press, pp.545-552, 2007.

[16] I. Craw, D. Tock, and A. Bennett, "Finding face features," 2nd European Conf. Computer Vision pp.92-96, 1992.

[17] A. Lanitis, C. J. Taylor, and T. F. Cootes, "An automatic face identification system using flexible appearance models," Image and Vision Computing, Vol.13, No.5, pp.393-401, 1995.

[18] T. K. Leung, M. C. Burl, and P. Perona, "Finding faces in cluttered scenes using random labeled graph matching," Proc. 5th IEEE International Conference on Computer Vision, pp.637-644, 1995.

[19] V. Faber, "Clustering and the Continuous k-Means Algorithm", Los Alamos Science, No. 22, pp.138-144, 1994.

[20] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models", International Journal of Computer Vision, pp.321.331, 1987.

저 자 소 개



최 용 섭
 2011: 서울디지털대학교
 일본학과 졸업
 2011~현재: 강원대학교
 컴퓨터정보통신공학과 석사과정
 1997~현재: ZakCorporation 대표
 2010~현재: (주)CHP커뮤니케이션 이사
 관심분야: 영상 신호 처리, 패턴 인식
 Email : bluetel7@naver.com



박 대 현
 2007: 강원대학교
 컴퓨터정보통신공학과 공학사
 2009: 강원대학교
 컴퓨터정보통신공학과 공학석사
 2009~현재: 강원대학교
 컴퓨터정보통신공학과 박사과정
 관심분야: 영상 신호 처리, 컴퓨터비전
 Email : dhpark509@kangwon.ac.kr



김 윤
 1993: 고려대학교 전자공학과 공학사
 1995: 고려대학교 전자공학과 공학석사
 2003: 고려대학교 전자공학과 공학박사
 1995~1999: LG-Philips
 LCD 연구소 주임연구원
 2002~2003: (주)달리텍 선임연구원
 2004~현재: 강원대학교
 컴퓨터정보통신공학과 교수
 관심분야: 영상 신호 처리
 멀티미디어 통신 및 처리
 Email : yooni@kangwon.ac.kr