

논문 2007-44SC-1-7

이동로봇의 물체인식을 위한 질의 기반 시각 집중 알고리즘

(Query-based Visual Attention Algorithm
for Object Recognition of A Mobile Robot)

류 광 근*, 이 상 훈**, 서 일 홍***

(Gwang Geun Ryu, Sanghoon Lee, and Il Hong Suh)

요 약

본 논문에서는 로봇이 태스크와 관련된 부분에 시각 집중을 하도록 하기 위해서 기존의 상황식 주목 알고리즘을 확장한 질의 기반 시각 집중 알고리즘을 제안한다. 질의 기반 시각 집중 알고리즘은 로봇이 수행 할 태스크와 관련한 물체를 질의하면 그 물체의 속성을 분석하여 여러 종류의 도드라짐(Conspicuity) 영상 지도에 적용될 가중치 값을 작성한다. 그리고 가중치를 이용하여 도드라짐 영상 지도들을 합성한 Saliency 영상 지도를 작성하여 기존의 주목 알고리즘과 비교 평가를 수행하였다. 여기서는 일례로서 질의 물체의 속성을 색으로 사용하였다.

Abstract

In this paper, we propose a query-based visual attention algorithm for effective object finding of a vision-based mobile robot. This algorithm is developed by extending conventional bottom-up visual attention algorithms. In our proposed algorithm various conspicuity maps are merged to make a saliency map, where weighting values are determined by query-dependent object properties. The saliency map is then used to find possible attentive location of queried object.

To show the validities of our proposed algorithm, several objects are employed to compare performances of our proposed algorithm with those of conventional bottom-up approaches. Here, as one of exemplar query-dependent object property, color property is used.

Keywords: Visual Attention, 시각 집중, Object Recognition, Mobile Robot, Attentive Image Feature

I. 서 론

로봇이 태스크를 성취하기 위해서는 올바른 환경 인식이 요구된다. 로봇의 환경 인식은 센서 정보를 기반으로 이루어지는데 예를 들어 주행 태스크는 소나나

레이저와 같은 거리 측정 센서가 사용된다. 그런데 물체 인식이 필요한 태스크는 소나나 레이저 센서만으로는 불가능하며 이를 위해 일반적으로 비전 센서를 사용한다. 상위 수준의 태스크 일수록 이동 로봇에게 비전 센서로 취득하는 환경정보가 중요하다. 이것은 비전 센서는 사람의 눈과 비슷하기 때문에 물체 인식이나 거리 측정까지 할 수 있을 정도로 활용 범위가 크고 영상에 내제된 정보량이 많기 때문이다. 예를 들어 영상 내부의 여러 물체들을 구분하는 알고리즘인 분할(Segmentation)은 라플라시안 필터와 가버 필터를 이용하여 각각 외곽선과 방향성 정보를 얻을 수 있는데 여기서 중요한 것은 비전 센서로 들어온 영상 한 장에서 두 개의 정보를 얻었다는 점이다. 이와 같이 영상으

* 학생회원, *** 정회원, 한양대학교 정보통신학과
(College of Information and Communications,
Hanyang University)

** 정회원, BK21 수요지향적 정보기술 전문인력양성
사업단

(Advanced IT Education Program on Industrial
Demand, a Part of Brain Korea 21 Project at
Hanyang University)

※ 이 논문은 서울시 산학연 협력사업의 지원을 받아
연구되었습니다.

접수일자: 2006년10월13일, 수정완료일: 2007년1월2일

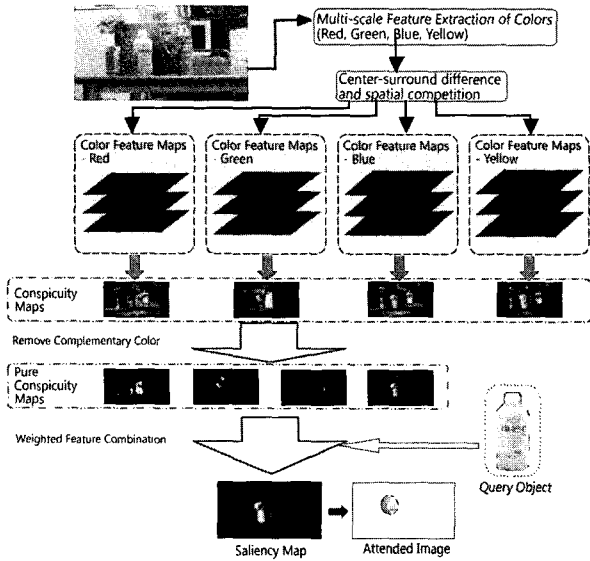


그림 1. 질의 기반 시각 집중 알고리즘 시스템 구조도
 Fig. 1. Query-based Visual Attention Algorithm System Structure.

로부터 취득 할 수 있는 정보의 양이 매우 많으며, 취득을 원하는 정보마다 다양한 가공 방법이 존재한다. 로봇의 특성상 실시간성이 중요한데 전체 영상 이용하여 정보를 얻는 것은 비효율 적일 수밖에 없다. 따라서 현재 수행해야 할 태스크와 관련된 정보만을 모을 수 있다면 효과적인 영상처리가 가능하다.

인간은 자신을 둘러싼 주변 환경으로부터 몸에 있는 센서를 통해 한 번에 천개 이상의 정보를 동시에 입력 받는다. 하지만 수많은 정보를 동시에 처리 할 수는 없기 때문에 자신이 해야 할 행동과 관련된 정보만을 모아서 처리를 하게 된다^[17]. 관련정보란 예를 들어 인간이 읽기라는 행동을 수행하고 있을 경우 책의 한 페이지에 그려진 삽화는 읽는 행동과 관계가 있지만 책이 놓여 있는 책상은 관계가 없다. 이와 같이 현재 행동과 관련된 정보만을 모으는 과정을 주목이라고 한다. 로봇도 인간과 마찬가지로 한 번에 처리할 수 있는 정보에는 한계가 있고 고수준의 태스크를 수행 할수록 정보의 양이 많아지기 때문에 반드시 주목 알고리즘이 필요하다. 특히 정보량이 많은 비전센서에 주목이라는 알고리즘을 적용하여 태스크와 관련된 정보만을 모으는데 이용하도록 한다.

컴퓨터 비전분야에서의 주목 알고리즘으로 Linda Lanyon의 물체 기반 주목 모델(Object-based Attention Model)^{[3][4]}과 Itti의 상향식 시각 집중 알고리즘(Bottom-up Visual Attention Algorithm)^[9]이 있다. 전

자는 인간의 주목과 관련한 두뇌활동을 모델링하여 시각적 자극에 따라 망막의 세포가 흥분하여 시점 이동을 수행하는 과정을 구현한 알고리즘이다. 이 알고리즘의 가장 큰 특징은 양방향성이다. 시간이 지남에 따라 망막의 자극이 강화되어 주목한 곳을 되먹임(Feedback)하여 다음에 주목할 위치에 영향을 끼친다. 후자의 상향식 시각 집중 알고리즘은 영상의 하급특징(Low-level Feature)를 이용하여 자극도를 계산하는 알고리즘이다. 상향식 시각 집중 알고리즘은 영상 내에 존재하는 여러 물체들 중에 정보량이 많은 부분을 우선적으로 주목하도록 한다. 정보량은 하급특징의 자극도와 비례한다. 그러나 알고리즘을 그대로 사용하면, 로봇이 수행할 작업과 연결하여 주목을 하기에는 적합하지 않다. 예를 들어 로봇의 사용자가 컵을 가져오라고 했을 경우 부엌 옆에 더 복잡한 무엇인가가 있으면 주목을 사용했을 경우 부엌보다 다른 복잡한 무엇인가를 주목할 가능성이 있다.

본 논문에서는 로봇의 태스크와 관련된 부분을 주목하기 위해 로봇이 태스크 수행을 위해 요구하는 질의 물체를 중심으로 주목을 하는 ‘질의 기반 주목 알고리즘’을 제안한다. 질의 물체의 속성을 분석하여 기존의 주목 알고리즘을 구성하는 하급 특징에 가중치를 주어서 질의 물체의 위치에 주목할 수 있다.

II. 질의 기반 시각 집중 알고리즘

본 논문이 제안하는 알고리즘은 앞서 언급한 상향식 주목 알고리즘을 기반으로 한다. 기존의 상향식 주목 알고리즘은 영상에서 자극도가 높은, 즉 정보량이 많은 부분을 알 수는 있지만 원하는 위치에 주목 할 수는 없었다. 즉, 책상위에 물체를 찾아야 할 경우 인간은 당연히 책상 위를 주목하지만 기존 주목 알고리즘으로는 책상 근처에 정보량이 더 많은 부분이 있을 경우 그곳을 주목하는 일이 발생한다. 따라서 로봇의 태스크에 관련한 주목을 보장하기 위해서 기존 주목 알고리즘의 구성 성분 중에서 태스크와 관련된 속성들을 강화시키고 비관련 성분들은 약화시키는 방법을 사용하여 정보량을 변화시킴으로써 최종적으로 주목할 위치를 가변하도록 한다.

질의 기반 주목 알고리즘의 시스템 구조도는 그림 1과 같으며 다음절부터는 알고리즘의 구현 방법을 설명한다.

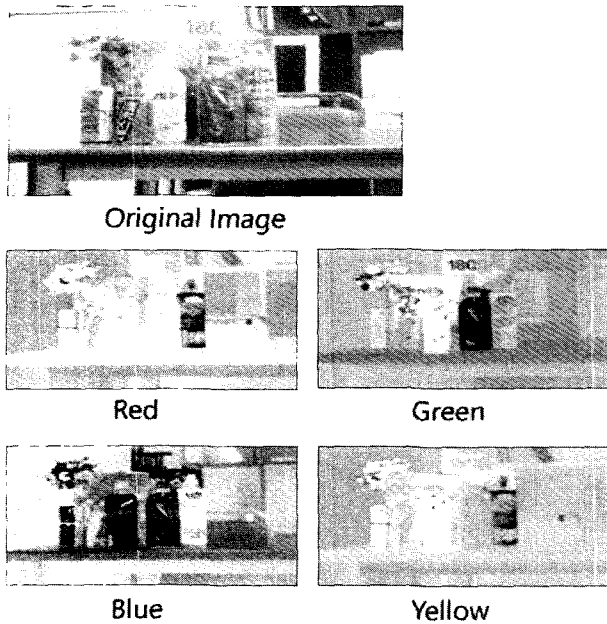


그림 2. 하급 특징 맵의 예시
Fig. 2. Examples of low-level feature maps.



그림 3. 가우시안 피라미드
Fig. 3. Gaussian Pyramid.

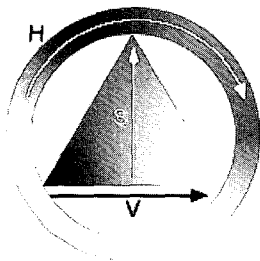


그림 4. HSV 컬러 공간
Fig. 4. HSV Color Space.

가. 하급 특징 맵과 가우시안 피라미드 구성

첫 번째 단계에서는 영상의 하급 특징을 추출하여 가우시안 피라미드를 이용한 크기조정영상(Scale Map)들의 집합을 만든다. 크기조정영상은 정보는 같고 크기만

다른 영상을 말한다. 영상의 하급 특징으로 색상 정보를 이용하는데, 적색(R), 녹색(G), 청색(B), 황색(Y)을 사용한다. 각 정보를 얻기 위한 수식은 다음과 같다.

$$R = r - (g + b)/2 \tag{1}$$

$$G = g - (r + b)/2 \tag{2}$$

$$B = b - (r + g)/2 \tag{3}$$

$$Y = (r + g)/2 - |r - g|/2 \tag{4}$$

식 (1)~(4)를 이용하여 얻은 영상들은 시각 집중 알고리즘의 입력이 되며, 각 영상들을 맵이라고 부르기도 한다. 그림 2는 각 하급 특징 맵들의 예시를 나타낸다. 이 맵들을 9개의 Scale을 가진 가우시안 피라미드(Gaussian Pyramid)^[2]로 구축해야 한다. 각 Map에 대한 가우시안 피라미드는 $I(\sigma)$, $R(\sigma)$, $G(\sigma)$, $B(\sigma)$, $Y(\sigma)$, $O(\sigma, \theta)$ 와 같이 표현하며, 인수 σ 와 θ 는 $\sigma \in \{0 \dots 8\}$, $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ 를 나타낸다. 예를 들어, σ 가 0이면 원본 영상이고 8일 경우, 원본 영상의 1/256배의 크기를 가진 영상이다. 가우시안 피라미드는 임의의 영상을 저대역 필터를 거친 뒤, 서브-샘플링을 수행하여 만든다. 그림 3은 가우시안 피라미드의 예시이다.

나. 질의 물체 및 가중치

태스크를 수행하기 위해 필요한 물체의 속성을 주목 알고리즘에 반영해야 한다. 따라서 질의 물체의 속성을 정확하게 분석할 필요가 있으며, 본 논문에서는 질의 물체의 속성으로 색상 정보를 이용한다.

질의 물체를 구성하는 색의 분포를 분석하기 위해 HSV 컬러 공간을 이용한다. 그림 4에 그것의 모델을 도시하였다. HSV 컬러 공간은 RGB에 비해 빛의 변화에 덜 민감한 특성을 갖고 있다. 그 이유는 RGB는 빛의 세기가 변하면 RGB값 각각에 영향을 주어 전혀 다른 색이 될 수 있지만, HSV 컬러 모델은 빛을 받으면 색의 강도는 변할지 모르나, 색 자체의 분류는 변하지 않기 때문이다.

HSV 색상 모델을 이용하여 질의 물체의 색의 분포를 계산한 뒤, 가장 많이 차지하고 있는 색을 1로 정하고 나머지 색의 비율을 계산하여 식 (1)~(4)에 각각 반영한다.

다. 특징 맵(Feature Maps)

가우시안 피라미드를 이용하여 특징 맵을 구축한다. 특징 맵은 하급 특징 맵들을 구성하고 있는 각 픽셀들이 주변 픽셀들과 어느 정도 차이가 있는지를 계산하여 영상으로 만든 것이다. 주변에 비해 두드러지는 픽셀을 찾아내기 위해서는 중심 차분 연산(Center-surround Difference Operation(Θ))을 이용한다. 특징 맵은 다음의 식을 이용하여 구할 수 있다.

$$R(c,s)=|R(c)\Theta R(s)| \tag{5}$$

$$G(c,s)=|G(c)\Theta G(s)| \tag{6}$$

$$B(c,s)=|B(c)\Theta B(s)| \tag{7}$$

$$Y(c,s)=|Y(c)\Theta Y(s)| \tag{8}$$

인수 c 와 s 는 $c \in \{2,3,4\}, s = c + \delta, \delta \in \{3,4\}$ 를 나타낸다. 중심 차분 연산은 서로 다른 크기의 영상을 같은 크기로 맞추는 뒤 각각에 대응하는 픽셀의 밝기값 차를 구하는 연산이다. 그리고 그것의 절대값으로 맵을 구성한다. 예를 들면, 인수 c 는 2이고 s 는 5일 경우 가우시안 피라미드의 2번째와 5번째 영상을 같은 크

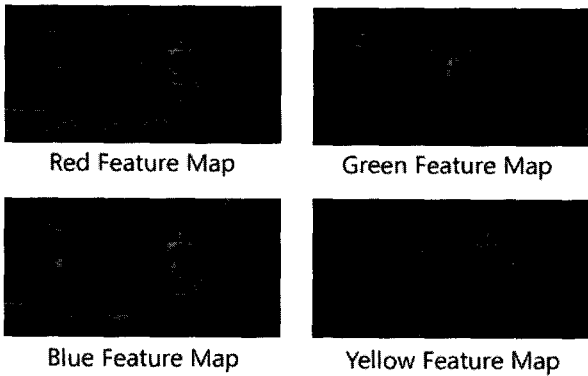


그림 5. 특징 맵의 예시
Fig. 5. Examples of Feature Maps.

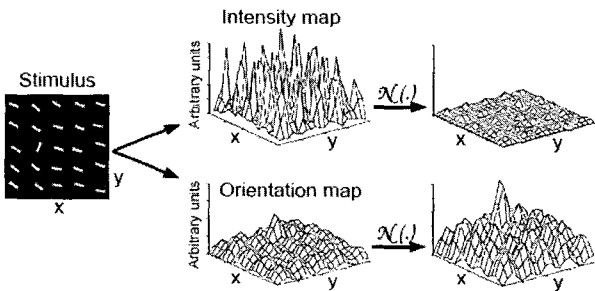


그림 6. 비선형 정규화
Fig. 6. Nonlinear Normalization.

기로 맞추는 뒤 차영상을 구한다. 그림 6은 특징 맵의 예시이다.

라. 도드라짐 맵(Conspicuity Maps)과

순수 도드라짐 맵(Pure Conspicuity Map)

도드라짐 맵을 만들기 위해서 특징 영상에 비선형 정규화(Nonlinear Normalization: \mathcal{N})^[7]를 적용한다. 비선형 정규화를 이용하면 앞에서 구한 특징 맵에서 두드러지는 부분을 강화시키고 그 외의 부분을 억제할 수 있다. 비선형 정규화를 적용하기 위해서는 다음의 순서대로 진행한다.

- ① 8bit 흑백영상의 밝기 값을 [0..M]의 값으로 정규화 한다. 본 논문에서는 M을 255로 한다.
- ② 영상에서 가장 밝은 부분인 M의 위치를 찾는다. 그리고 영상의 지역 최대값(Local Maxima)의 평균인 \bar{m} 을 계산한다. 지역 최대값이란 주변 픽셀에 비해 값

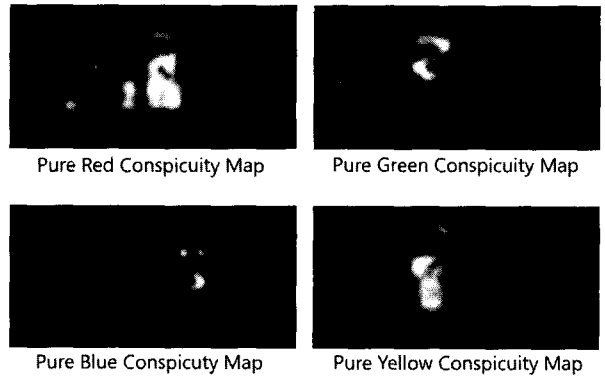
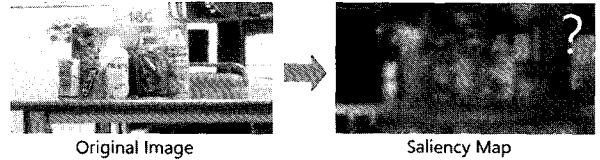


그림 8. 보색이 제거된 도드라짐 영상 예시
Fig. 8. Examples of Pure Conspicuity Maps.

1. Bottom-up Visual Attention Algorithm



2. Query-based Visual Attention Algorithm

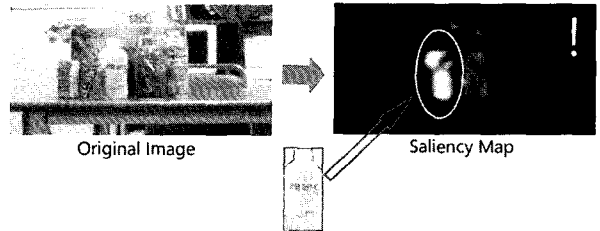


그림 9. 상향식 시각 집중 알고리즘과 질의 기반 주목 알고리즘의 결과 비교
Fig. 9. Compare the result of bottom-up visual attention algorithm and our algorithm.

이 큰 픽셀들을 말한다.

③ 영상의 모든 픽셀에 $(M - \bar{m})^2$ 을 곱한다.

비선형 정규화의 효과는 그림 6과 같이 표현할 수 있다. 방향성 정보로 구성된 맵에서 방향성이 튀는 부분을 중심으로 더 강조되고 다른 부분은 억제된 것을 볼 수 있다. 모든 특징 영상에 비선형 정규화를 적용한 뒤 Across-scale Addition(\oplus)을 수행하면 도드라짐 맵을 구할 수 있다. Across-scale Addition이란 서로 크기가 다른 맵 두개를 이용하여 각각의 픽셀마다 합을 구하는 것으로 최종적인 영상의 해상도는 가우시안 피라미드의 중앙의 크기와 동일하다. 수식으로 표현하면 다음과 같다.

$$\bar{R} = \bigoplus_{c=2}^4 \left(\bigoplus_{s=c+3}^{c+4} N(R(c,s)) \right) \quad (9)$$

$$\bar{G} = \bigoplus_{c=2}^4 \left(\bigoplus_{s=c+3}^{c+4} N(G(c,s)) \right) \quad (10)$$

$$\bar{B} = \bigoplus_{c=2}^4 \left(\bigoplus_{s=c+3}^{c+4} N(B(c,s)) \right) \quad (11)$$

$$\bar{Y} = \bigoplus_{c=2}^4 \left(\bigoplus_{s=c+3}^{c+4} N(Y(c,s)) \right) \quad (12)$$

그림 7은 도드라짐 영상의 예시를 나타낸다.

기존의 상향식 주목 알고리즘과 달리 본 논문에서 제안하는 알고리즘은 색상 정보의 주목에 대한 기여도를 높이기 위하여 각 도드라짐 영상에서 보색을 제거하는 과정을 거친다. 보색을 제거한 도드라짐 영상을 순수 도드라짐 맵이라 하였으며, 예시를 그림 8에 도시하였다. 그림에서 보는 것과 같이 자극 되는 부분이 많이 감소

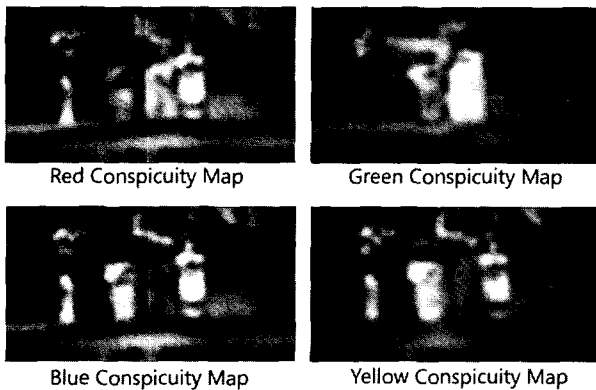


그림 7. 도드라짐 영상 예시
Fig. 7. Examples of Conspicuity Maps.

하여 본래 색의 영향이 강화된 것을 알 수 있다.

마. Saliency 영상 지도

기존의 상향식 시각 집중 알고리즘은 도드라짐 영상을 모두 합하여 다음 식과 같이 자극 맵을 구축한다.

$$S = \frac{1}{3}(N(\bar{R}) + N(\bar{G}) + N(\bar{O})) \quad (13)$$

질의 기반 주목 알고리즘의 자극 맵을 만들기 위해서는 질의 물체를 구성하는 색의 비율에 따라 가중치를 주어서 도드라짐 영상들의 자극 영상에 대한 기여도를 조절해야 한다. 질의 물체의 색에 대한 가중치를 적용하기 위하여 식(13)을 재구성 하였다. 식은 다음과 같다.

$$S = \frac{1}{K}(\alpha N(\bar{R}) + \beta N(\bar{G}) + \gamma N(\bar{B}) + \delta N(\bar{Y})) \quad (14)$$

$$K = \alpha + \beta + \gamma + \delta$$

식(13)와 식(14)를 이용하여 자극 영상을 만든 예시를 그림 9에 보였다. 그림에서 보면 알 수 있듯이 상향식 시각 집중 알고리즘에 비해 질의 기반 주목 알고리즘이 해당 물체의 자극도를 향상시킨 것을 알 수 있다.

바. 마스크 영상(Masked Image)

Saliency 영상 지도의 자극도의 크기에 따라 시점의 위치를 정하고 일정 면적에 해당하는 마스크 영상을 만든다. 마스크는 자극도가 높은 지점의 일정 영역의 영상을 잘라낸 것으로 로봇의 영상 처리를 위해 영상 정보의 크기를 줄인 것이다. 주목도에 따라 잘라낸 영상을 주목 영상(Attended Image)이라고 하는데 본 논문에서는 원형으로 잘라냈다. 자극도 주변 일정 영역은 1의 값을 갖고 나머지 영역은 0의 값을 갖는 마스크를 만든 뒤 입력 영상과 합친다. 식은 다음과 같다.

$$I(x,y) = [255 - M(x,y)(255 - I(x,y))] \quad (15)$$

마스크와 입력영상이 합쳐진 영상(I)의 예시는 그림 10과 같다.

사. 주목 영상(Attended Image) 및 물체 인식

앞 절과 같이 만들어낸 주목 영상을 물체 인식 알고리즘의 입력으로 만들어서 우리가 제안하는 주목 알고리즘이 올바르게 주목을 하였는지를 평가한다. 주목 알고리즘으로는 다수의 지역 특징점을 이용한 물체 인식 알고리즘의 하나인 Scale-invariant Feature Transform

(SIFT)^[1]를 이용하였다. SIFT 알고리즘은 데이터베이스에 저장된 물체의 SIFT 특징점과 새로 들어오는 영상의 SIFT 특징점이 일정 개수 이상 일치할 경우 물체가 존재한다고 한다. 본 논문에서는 3개 이상의 특징점이 일치하였을 경우 물체를 인식했다고 정하였다.

III. 실험

1. 다수의 지역 특징점을 이용한 물체 인식 알고리즘의 성능 실험

앞서 언급한 SIFT는 높은 물체 인식률을 바탕으로 물체인식 분야에서 많이 사용된다. 물체 인식률이 높은 대신 계산량이 많다는 단점이 있다. 따라서 주목 알고리즘을 이용하여 비전 정보를 줄임으로써 물체 인식 알고리즘의 속도 향상에도 일조할 수 있다. 실험 1에서는 SIFT 알고리즘의 강인성을 평가하여 이후 SIFT를 이용하는 실험의 보조 지표로 사용하고자 한다. SIFT의 가장 큰 특징은 데이터베이스에 들어있는 물체와 현재 보고 있는 물체가 같은 것일 경우 그 물체가 회전하였을 경우나 크기가 변하여도 인식이 가능하다는 것이다. 이렇게 회전과 크기 변화에 대해 어느 정도 강인성을 보이는지 실험을 통해 알아보았다.

실험 방법은 다음과 같다. 데이터베이스에 임의의 물체 3개를 저장해 놓고 동일한 물체를 거리와 각도에 따라 촬영한 영상을 취득하여 SIFT 매칭을 통해 결과를 비교한다. 데이터베이스에 들어있는 물체 영상은 물체를 1m 떨어진 거리에서 정면으로 촬영한 사진을 저장하였으며 입력 영상으로는 동일한 물체를 거리는 1m ~ 5m 떨어진 거리에서 1m 간격으로 이동하며 촬영하였다. 각도에 대한 강인성을 알아보기 위해 $\pm 50^\circ$ 사이로 10° 의 간격으로 움직이면서 촬영하였다. 한 물체 당 총 16장의 사진을 촬영하였으며 3물체 48장의 사진을 입력 영상으로 정하였다. 물체1은 A4박스, 물체2는 에어컨 세정제, 물체3은 커피 박스를 사용하였고, 해상도는 1024×768 이다. 실험 결과는 그림 12, 13과 같다.

그림 11의 각도를 변화시킨 물체의 인식 실험에서 매칭된 특징점 개수의 변화를 살펴보면, 0° 에서 10° 만 벗어나도 매칭된 키의 개수가 급격하게 줄어들지만 일정 수준의 특징점 수를 유지한다. 이는 매우 고무적인 것으로 비전 정보의 특성상 회전했을 때에 정면에서 바라보는 것과 영상이 많이 변할 수 있는데도 인식이 가능하다는 것을 보여준다. $20^\circ \sim 30^\circ$ 사이로 물체가 회전한 경우에도 특징점의 일치가 이루어지고 있다. 물체

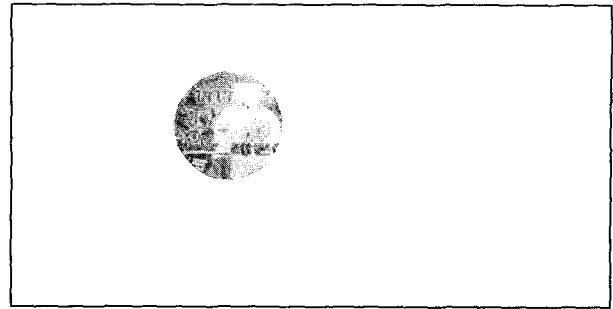


그림 10. 마스크가 씌워진 주목 영상의 예시
Fig. 10. Examples of Masked Image.

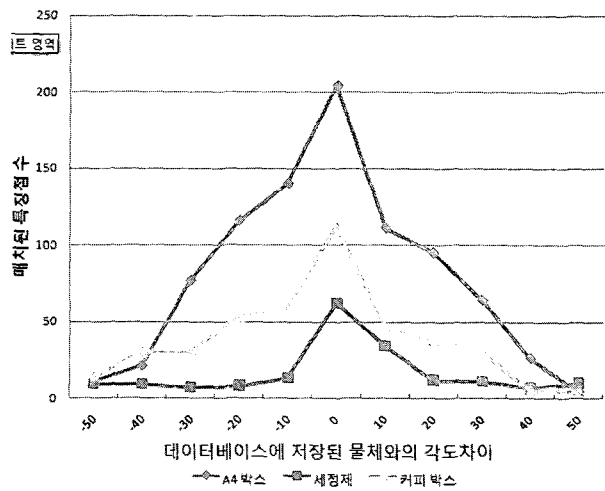


그림 11. 각도에 따른 매치된 특징점 개수의 변화
Fig. 11. The change according to the angle of a SIFT Keypoint number to be matched.

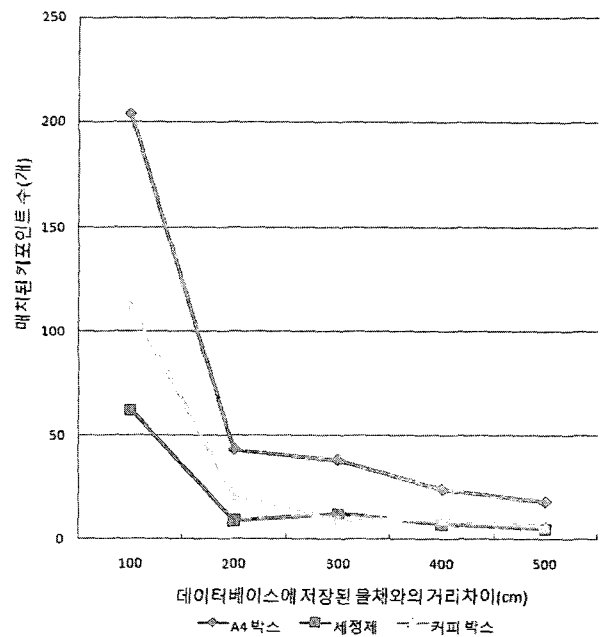


그림 12. 거리에 따른 매치된 특징점 수의 변화
Fig. 12. The change according to the distance of a SIFT Keypoint number to be matched.

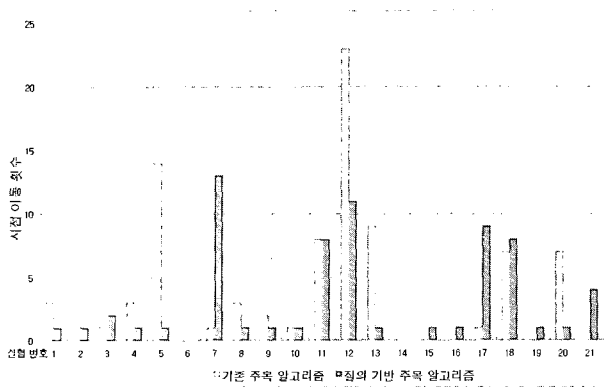


그림 13. 질의 기반 주목 알고리즘과 기존 알고리즘의 비교 실험 결과

Fig. 13. Result of comparing between bottom-up visual attention algorithm and query-based visual attention algorithm.

회전에 대한 최대 한계는 약 30° 정도라고 보인다. 40° 이상 회전했을 경우에도 어느 정도 특징점이 매치되긴 하지만 특징점의 개수가 적기 때문에 정확한 물체 인식을 보장하지 못할 것으로 판단된다.

그림 12의 거리에 따른 특징점 수의 변화를 살펴보면 100cm과 200cm의 매치된 특징점의 개수가 급격하게 감소하는 것을 볼 수 있다. 물체 자체의 면적 감소에 따른 검출되는 특징점의 수의 감소에 비추어 볼 때 충분히 예상할 수 있는 결과이지만 데이터베이스에 물체의 특징점을 저장할 경우 그 물체를 촬영한 거리도 충분히 고려해야 한다는 사실을 알 수 있다.

SIFT 특징점을 안정적으로 사용하기 위해서는 데이터베이스에 저장된 물체와의 시점 차이가 ±30° 이내이고 데이터베이스에 들어있는 물체와 실제 촬영한 물체의 거리가 100cm이내이어야 안정적인 물체 인식을 보장할 수 있다고 결론을 내렸다.

2. 질의 기반 주목 알고리즘 성능 실험

본 논문에서 제안하는 질의 기반 시각 집중 알고리즘의 시뮬레이션 프로그램을 작성하여 기존 주목 알고리즘과의 성능 비교를 수행하였다.

이 실험의 성능을 평가하기 위해서 찾고자 하는 물체를 질의로 주었을 때, 몇 번의 시점 이동 후 원하는 물체를 찾았는가를 평가하였다. 시점 이동이 적을수록 성능이 더 좋다고 할 수 있다. 시점 이동이란 Saliency Map의 자극도 순으로 주목을 이동하는 과정을 말한다. 각 주목의 이동마다 주목 영상을 생성한다.

실험환경구성은 다음과 같다. 입력 영상은 800×600×24bit의 컬러 영상을 사용하였으며 질의 물체는

표 1. 질의 기반 주목 알고리즘과 기존 알고리즘의 비교 실험 결과 정리

Table 1. Arrangement of experiment result.

	기존 방법	질의 기반 방법
평균 시점이동 횟수(검색 성공 시 종료)	6회	3.4회
3회 이내 검색 성공률	38%	62%
검색실패율	29%	10%

영상 안에 존재하는 물체들에서 무작위로 선정하였다. 실험의 목적은 물체가 2개 이상 존재하는 Complex Scene에서 질의 얼마나 빠르게 물체에 주목하는가를 알아보기 위해 영상 안에는 적어도 5개 이상의 물체를 배치했다. Complex Scene은 5장을 사용하였고 한 Scene 당 4번의 질의를 하였기 때문에 실험은 총 21회 수행하였다. 시점 이동은 30회로 제한하고 그 안에 찾지 못한 경우 실패로 정하였다. 18회는 실내에서 여러 물체를 놓고 촬영한 사진을 이용하였으며 3회는 외국의 데이터베이스에 들어있는 비전 실험에서 주로 사용하는 사진을 이용하여 실험하였다. 이 경우 질의 영상은 사진 내에 존재하는 임의의 특정 지역으로 정하였다.

실험 결과를 그림 13에 그래프로 도시하였으며 결과를 정리하여 표 1에 나타내었다.

IV. 결 론

질의 영상의 특징을 분석하여 주목 알고리즘에 적용하는 방법을 제시하고 실험을 통해 성능을 알아보았다.

표 1에서 질의 기반 주목 알고리즘이 기존 알고리즘보다 성능이 향상되었음을 알 수 있다. 3회 이내의 시점이동으로 검색에 성공한 확률도 기존 방법에 비해 약 24% 증가하였다. 이 결과를 통해 질의 영상의 색상을 이용하여 주목 알고리즘의 Saliency 영상지도의 우선순위에 올바르게 영향을 주었다는 것을 알 수 있었다.

그러나 그림 14를 보면 12번째 실험에서 질의 기반 방법의 결과가 28번째로 가장 늦게 검색을 성공한 것으로 되어있다. 이것은 질의한 영상의 색상분포가 적색, 녹색, 청색, 황색의 영역에 들지 못하기 때문에 가중치를 정확히 부여하지 못하여 Saliency 영상지도의 우선순위가 잘못 변경된 것으로 추정된다. 특히 흰색이 많이 존재하는 질의 영상의 경우 대체적으로 늦게 검색하거나 검색에 실패하는 경향을 보였다(그림 13 실험번호 14, 17).

질의 기반 시각 집중 알고리즘의 성능을 향상시키기

위해서 색상정보와 동시에 외형정보를 사용하면 색상정보가 약한 경우에도 질의 물체에 대한 자극도의 변화를 가능하게 할 수 있을 것으로 기대한다. 예를 들어 노란색 음료수 병을 찾을 경우 색과 동시에 병의 외형 정보를 외곽선을 구하는 알고리즘을 이용하여 새로운 특징 맵을 만들어 낸다. 그 외에도 무늬(Texture)를 이용할 수도 있다.

추후에는 색상 정보 외에 외형과 윤곽의 정보, 그리고 무늬들의 속성을 반영하는 새로운 도드라짐 맵을 구축하여 질의 물체의 속성을 충분히 반영할 수 있는 질의 기반 시각 집중 알고리즘을 연구할 계획이다.

참 고 문 헌

- Pattern Analysis and Machine Intelligence, Vol. 20, No. 11, pp.1254-1259, November 1998.
- [10] HAUN P. VECERA and STEVEN J.LUCK, "Attention", In V.S.Ramachandran (Ed.), Encyclopedia of the human brain, Vol. 1, pp.269-284), San Diego: Academic Press
- [1] David.G.Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision 2004, pp.1-28, January 5, 2004.
- [2] David A. Forsyth, Jean Ponce, "Computer Vision - a modern approach," Prentive Hall, 2003.
- [3] Linda Lanyong and Susan Denham, "A Model of Active Visual Search with Object-based Attention Guiding Scan Paths," Neural Networks 17, pp.873-897, 2004.
- [4] Linda Lanyon and Susan Denham, "A Model of Object-based Attention That Guides Active Visual Search to Behaviourally Relevant Locations," WAPCV 2004, LNCS 3368, pp.42-56, 2005.
- [5] Dirk Walther, Ueli Rutishauer, Christof Koch , and Pietro Perona, "On the Usefulness of Attention for Object Recognition," In Proceedings of WAPCV, 2004.
- [6] Brad C. Motter, Eric J. Belky, "The Guidance of Eye Movements during Active Visual Search," Vision Research 38, 1998, pp.1805-1815
- [7] J.J. Bonaiuto & L. Itti, "Combining Attention and Recognition for Rapid Scene Analysis," Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.
- [8] Cheng Liu, "Gabor-based Kernel PCA with Fractional Power Polynomial Models for Face Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, No. 5, May 2004.
- [9] Laurent Itti, Christof Koch and Ernst Niebur, "A Model of Saliency-based Visual Attention for Rapid Scene Analysis," IEEE Transactions on

저자 소개



류 광 근(학생회원)
 2004년 한양대학교 전자컴퓨터
 공학부 학사 졸업.
 2006년 한양대학교 정보통신학과
 석사 졸업.
 2006년~현재 한양대학교
 정보통신학과 박사 과정.

<주관심분야 : 컴퓨터비전, 지능응용시스템, 로봇
 공학>



이 상 훈(정회원)
 1994년 한양대학교 수학과 학사
 졸업.
 1997년 한양대학교 전자계산학과
 석사 졸업.
 2006년 한양대학교 전자전기제어
 계측과 박사 졸업.

2006년~현재 수요지향적 정보기술 전문인력양성
 사업단 박사 후 과정

<주관심분야 : 로봇 지능, 행동선택방법, 로봇
 S/W 프레임워크, 지능응용시스템>



서 일 홍(정회원)
 1977년 서울대학교 공학사(전자).
 1982년 KAIST 공학석사 및 박사
 (전기 및 전자).
 1985년 대우중공업 기술연구소.
 1985년~현재 한양대학교 교수

<주관심분야 : 지능응용시스템, 인공지능, 로봇공
 학>