

# 잡영 영상에서의 선택적 주의 기반 목표물 탐지 시스템

최경주<sup>0</sup> 이일병

연세대학교 컴퓨터과학과

{kjcheoi<sup>0</sup>, yblee}@csai.yonsei.ac.kr

## A Selective Attention Based Target Detection System in Noisy Images

Kyungjoo Cheoi<sup>0</sup> Yillbyung Lee

Dept. of Computer Science, Yonsei University

### 요 약

본 논문에서는 선택적 주의에 기반한 잡영 영상에서의 목표물 탐지 방법에 대해 기술한다. 특히 제안하는 방법은 목표물에 대한 아무런 지식을 사용하지 않고, 단지 입력되는 영상의 상향식 단서만을 사용하여 목표물을 탐지해냄으로써 여러 다양한 분야에 일반적으로 사용될 수 있다. 제안하는 시스템에서는 몇 가지 기본 특징들이 입력된 영상에서 바로 추출되며, 이러한 특징들이 서로 통합되어 가는 과정에서 목표물 탐지에 유용하지 않은 정보는 자연스럽게 걸러지며, 유용한 정보는 추가되고 부각되어진다. 간단한 영상부터 복잡한 자연영상에 이르는 다양한 잡영 영상을 대상으로 실험하여 제안하는 시스템의 성능을 평가하였다.

### 1. 서 론

영상 분석에 관련된 대부분의 전형적인 방법들은, 우선 전체 영상으로부터 고차원적인 특징을 추출하고, 추출된 각각의 특징집합들이 미리 저장된 대상(object)에 대한 프로토타입과 순차적으로 비교되어진다. 따라서 이러한 방법들은 속도가 굉장히 느리며, 또한 특정 응용분야에 제한적인 성능을 가지고 있다. 이에 비해 생물학적인 시스템은 모든 시각적인 장면을 병렬적으로 모두 처리한다기 보다는 보다 복잡한 고차원적인 처리를 위해 주의를 집중되는 일정한 영역들을 순간적으로 포착하여 그 부분을 순차적으로 처리해가는 전략을 사용한다[1]. 이러한 선택적 주의 기능으로 인해 영장류는 시스템의 제한된 시각장면을 실시간으로 인지할 수 있는 것이다.

목표물 탐지에 빠르고도 효과적인 컴퓨터 시각 알고리즘이 많이 개발되어졌지만, 이들은 특정 목표물에 아주 잘 조율이 되어져 있으며, 따라서 조율된 목표물이 아닌 다른 대상을 탐지해내지 못하는 경우가 대부분이다. 이러한 이유로 인해 다른 응용분야로의 확장이 어려운 실정이다. 또한 대부분의 알고리즘들이 입력되는 영상에 잡영이 끼어있을 경우 목표물 탐지에 실패한다. 이러한 배경 아래, 본 논문에서는 탐지하고자 하는 목표물에 대한 어떠한 선지식도 사용하지 않는 일반적인 목표물 탐지 시스템으로써, 잡영에 강한 시스템을 제안한다. 제안하는 시스템은 위에서 언급한 시각적 주의의 상향식 처리 방법에 기반을 두고 있다.

본 논문의 구성은 다음과 같다. 먼저, 2장에서 기존의 연구에 관해 간략히 기술하며, 3장에서 제안하는 목표물 탐지 시스템에 관해 설명한다. 이어 4장에서 실험결과에 대해 논하고 5장에서 결론을 맺는다.

### 2. 기존연구

\* 본 연구는 과기부 뇌신경정보학사업의 지원을 받음.

시각적 주의는 현재 공학에서 주로 나누어서 처리되고 있는 전경과 배경의 분리와 인식의 문제를 효율적으로 결합할 수 있으나 아직 공학에서 충분히 활용되지 못하고 있는 매우 중요한 동물의 지적 능력이라 할 수 있다. 시각적 주의 개념을 사용하는 컴퓨터 시각 시스템에서 특히 해결해야 할 기본 문제는 어떻게 시스템으로 하여금 주의를 집중하게 하도록 할 것인가이다. 어떠한 특징이 주의를 집중하는데 있어서 중요한 것인가?

인간의 주의 집중 방식에 관한 수많은 패러다임들이 많은 심리학자들에 의해 연구되어져왔다[2,3,4,5,6,7]. 그 중에서 많은 연구자들이 주의를 가해지는 방식이 2개의 서로 다른 기능을 하는 위계적인 단계, 전-주의 단계와 주의 단계로 이루어져 있다는 가설에 힘을 주고 있다. 이 가설에 의하면 초기의 전-주의 단계에서는 시각장(Field of View)에 놓여진 모든 물체들이 병렬적으로 활성화시키고(상향 지향적 처리), 시간의 경과와 함께 이러한 넓은 범위가 주의 단계에서 점차 특정된 대상이나 영역으로 집중되어 향후 그 부분만이 세밀한 스케일로 분석되어진다는 것이다(하향 지향적 처리).

컴퓨터 시각에서 생물학적인 증거에 기반을 둔 목표물 탐지 시스템이 개발되었다[6,8,9,10,11,12]. 이 중, [8]과 [9]의 시스템은 단지 간단한 인공영상 및 알파벳과 아주 간단한 영상에만 시스템을 적용하여 다른 복잡한 컬러 실 영상에 적용하기에는 한계가 있다. [10]에서는 실 영상에 적용하려고 노력은 하였지만, 잡영 영상은 고사하고 복잡한 컬러 실 영상에 대한 실험결과가 아주 미흡하며, [11]과 [12]에서는 좀 더 많은 잡영 영상에 대한 시스템 분석이 요구된다.

제안하는 시스템은 입력되는 시각 자극의 특징만을 사용하여 잡영 영상에서 목표물을 탐지하며, 입력된 영상의 상향식 단서(bottom-up cue)만을 사용함으로써 시스템의 기본 구조를 크게 변경시키지 않고도 다른 응용분야에 적절히 사용할 수 있다.

### 3. 제안하는 목표물 탐지 시스템

#### 3.1 특징추출

제안하는 시스템에서는 제일 먼저 입력되는 원영상이 서로 독립적인 3개의 특징지도로 분할된다. 각각의 지도는 저차원적인 특징집합에 의해 계산된 특정한 속성을 가지고 있다. 본 시스템에서는 하나의 무색상 특징지도와 2개의 색상 특징지도가 생성되었다.

무색상 특징지도  $F^1$ 은 명암도(밝기) 정보를 이용하여 생성되어지는데,  $R, G, B$ 가 각각 적색, 녹색, 청색을 나타낸다고 할 때, 명암도 정보는 입력되는 영상의 색상정보를 사용하여  $I=(R+G+B)/3$ 을 통해 추출할 수 있으며, 이 정보가  $F^1$ 이 된다.

2개의 색상지도는 시각피질에 있는 동일한 형태를 지닌 수용야를 가지고 있는 세포들에게 보여지는 2가지 형태의 color-opponency 필터를 기반으로 하여 생성되어진다. 먼저,  $r=R-(G+B)/2$ ,  $g=G-(R+B)/2$ ,  $b=B-(R+G)/2$ ,  $y=R+G-2((R-G)+2)$ 와 같이 다른 색상과 조급도 회색되지 않은 순수한 색상에 조율된 채널  $r, g, b, y$ 를 생성한다. 여기서,  $r, g, b, y$ 는 각각 순수한 색상에 조율된 적색, 녹색, 청색, 황색 채널을 나타내며, 각 채널은 흑색과 백색 자극에 대해 0을 출력한다. 이렇게 생성된 채널  $r, g, b, y$ 를 사용하여 적/녹 opponency에 따른  $F^1$ 을  $F^1=r-g$ 로, 청/황 opponency에 따른  $F^2$ 를  $F^2=b-y$ 로 계산하여 2개의 색상지도 생성한다.

이렇게 생성된 모든 특징지도들은 각각의 특징지도로부터 방위를 탐지하고 주변과 크게 다른 영역을 부각시키기 위하여 식(1)으로 업데이트되어진다.

$$F_{x,y}^k = \sum_{\theta} \left[ \sum_{m,n} F_{m,n}^k \cdot h_{x-m,y-n}(\theta) \right]^2 \quad (1)$$

우선, 서로 다른 특징추출기법에 기인한 across-modality 상이점을 배제하기 위해서 생성된 3개의 독립적인 특징지도들 모두 0~1 사이의 값으로 정규화한다. 그리고 난 후, 각 특징지도는 8가지 방위를 가진  $h(\theta)$  필터(단,  $\theta \in \{0, \pi/8, 2\pi/8, \dots, 7\pi/8\}$ )와 회전되는데,  $h(\theta)$  필터는 식(2)와 같이 계산된다.

$$h_{x,y}(\theta) = | K_1 \cdot G_{x,y}(\sigma, r_1 \cdot \sigma, \theta) - K_2 \cdot G_{x,y}(r_2 \cdot \sigma, r_1 \cdot r_2 \cdot \sigma, \theta) | \quad (2)$$

여기서,  $G(\cdot, \cdot, \cdot)$ 은 방위를 가진 2차원 가우시안 함수이며,  $K_1, K_2$ 는 양수,  $r_1$ 은 2개의 가우시안의 이심률,  $r_2$ 는 ON과 OFF 가우시안 간의 폭 비율을 뜻한다. 회전 후, 회전된 결과를 제공하여 콘트라스트를 증가시킨 후, 방위별로 나뉘어진 지도들을 하나로 통합한다.

#### 3.2 특징통합

목표물을 탐지하기 위하여 서로 독립적인 3개의 특징지도들은 하나의 출력지도로 통합되어진다. 출력지도란, 시적인 환경에 놓여있는 특이할만한 물체들에 대한 정보

를 가지고 있는 2차원 지도를 뜻하는 것으로써 Koch와 Ullman[5]이 진-주의 단계에서의 선택의 문제를 해결하기 위하여 처음 도입하였다. 서로 독립적인 각각의 특징지도 하나로 통합하기란 쉬운 일이 아닌데, 본 논문에서는 특징지도를 구성하고 있는 뉴런들의 통계적 정보와 국부적인 경쟁력 특성을 이용한 아주 간단한 통합 방법을 제안한다. 제안하는 방법은 그 수는 적으면서도 의미 있는 활동양이 높은 영역이 존재하는 지도를 지지하고, 그렇지 않은 지도는 억제한다.

먼저, 각 특징지도는 식 (3)과 같이 LoG 함수로 생성된 넓은 크기의 2차원 필터와 회전되고, 원래 입력된 영상과 더해진다.

$$F_{x,y}^k = \sum_{m,n} (F_{m,n}^k \cdot LoG_{x-m,y-n}) + F_{x,y}^k \quad (3)$$

이 과정이 몇 차례 반복되고, 결과적으로 계산된 지도는 식 (4)와 같이 계산되어 3개의 출력지도로 만든다.

$$S_{x,y}^k = \frac{SF_{x,y}^k - MIN}{MAX - MIN} \quad (4)$$

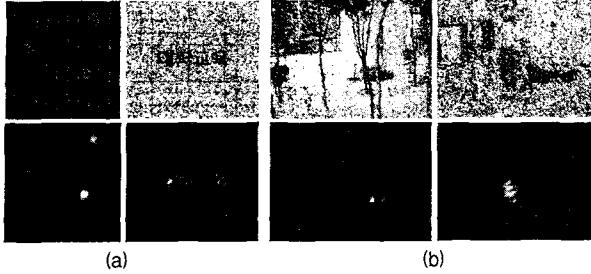
$$SF_{x,y}^k = F_{x,y}^k \times (MAX^k - AVE^k)^2$$

여기서,  $MAX^k$ 는  $F^k$ 를 구성하는 특징값 중 가장 큰 값을,  $AVE^k$ 는  $F^k$ 의 평균값을 뜻하며,  $MAX$ 과  $MIN$ 은 각각  $SF^1, SF^2, SF^3$ 를 구성하는 모든 특징값 중 가장 큰 값과 가장 작은 값을 뜻한다. 마지막으로, 3개의 출력지도  $S^k$ 는 모두 합해져 하나의 출력지도  $S$ 를 만든다.

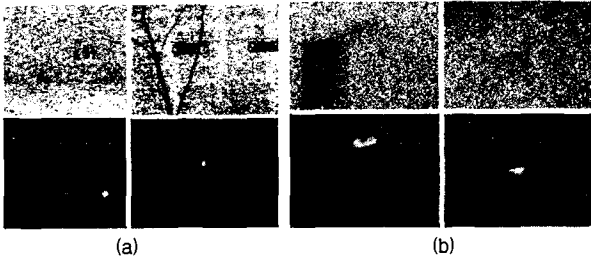
### 4. 실험결과

제안하는 시스템의 성능을 평가하고자 인간의 시각탐색에 관한 심리학적 실험에 사용되는 간단한 인공영상에서부터 외부환경에서 흔히 볼 수 있는 복잡한 실영상에 이르는 다양한 컬러 영상을 대상으로 실험을 수행하였다. 각각의 실험영상은 모두 상당한 양(40%~100%)의 잡음을 가지고 있는데, 이 때 잡음은 가우시안 분포(그림 1(a), 그림 2(a)) 또는 균일한 분포 형태(그림 1(b), 그림 2(b))를 보이며, 잡음이 색상을 가질 수도 있으며(그림 1), 또는 색상이 없을 수도 있다.

간단한 인공영상의 경우, 목표물(target)이 다른 방해물(distractor)들과 방위, 색상, 모양, 크기, 밝기면에서 다른 영상을 사용하였으며, 이런 영상에 대해 모두 원하는 목표물을 제대로 탐지하였으며, 그림 1(a)는 이러한 간단한 인공영상에 대한 실험결과를 보여준다. 복잡한 실영상의 경우, 각각의 영상은 간판, 신호등, 교통 표지판, 현수막, 우체통 등과 같은 목표물이 존재하며, 국부적인 영역에 있어서 조명이 균일하지 않다거나, 목표물이 아닌 기타 다른 대상과도 같은 방해물들이 존재한다. 그림 1(b)~그림 2는 제안하는 시스템을 통해 얻어진 몇몇 실영상에 대한 결과를 보여준다.



[그림 1] 색상이 있는 잡영이 낀 영상에 대한 실험결과 : (a) 가우시안 분포의 잡영 : (왼쪽) 50%의 잡영이 낀 영상에서 모양이 다른 3개의 도형 탐지 (오른쪽) 70%의 잡영이 낀 영상에서 가운데에 있는 푸른 글자 탐지 (b) 균일한 분포의 잡영 : (왼쪽) 30%의 잡영이 낀 영상에서 붉은 바구니 탐지 (오른쪽) 90%의 잡영이 낀 영상에서 붉은 우체통 탐지



[그림 2] 색상이 없는 잡영이 낀 영상에 대한 실험결과 : (a) 가우시안 분포의 잡영으로써 40%의 잡영이 낀 영상에서의 목표를 탐지 : (왼쪽) 붉은 비상 삼각대 탐지 (오른쪽) 녹색볼이 켜진 신호등 탐지 (b) 균일한 분포의 잡영으로써 60%의 잡영이 낀 영상에서의 목표를 탐지 : (왼쪽) 녹색의 비상 탈출구 신호등 탐지 (오른쪽) 황색의 개 탐지

5. 결 론

본 논문에서는 잡영 영상에서 목표물을 탐지하는, 선택적 주의 기반 목표물 탐지 시스템을 제안하였다. 제안하는 시스템은 기본적으로 특징추출과 특징통합의 두 단계로 이루어져 있다. 특징추출 단계에서는 컬러와 강도와 같은 아주 간단한 특징지도들이 입력영상으로부터 추출되고, 추출된 독립적인 특징지도들이 각각의 특징지도로부터 방위를 탐지하고 주변과 크게 다른 영역을 부각시키기 위하여 방위를 가진 ON-중심, OFF-주변 연산자에 의해 처리된다. 이렇게 계산된 각각의 특징지도들은 특징통합 단계서의 각 지도를 구성하고 있는 특징값들의 국부적인 경쟁 관계와 통계적 정보를 이용하여 서로 통합되어진다. 이러한 통합과정을 통하여 목표물을 탐지하기에 유용하지 않은 특징은 자연스럽게 걸러지고, 그렇지 않은 정보는 부각되게 된다.

제안하는 시스템의 전체적인 성능을 평가하기 위하여 크게 2가지 부류의 잡영 영상을 사용하였는데, 그 첫 번째가 인간의 시각탐색에 관한 심리학적인 실험에 사용되는 간단한 인공영상이며, 두 번째가 서로 다른 범위에서 촬영된 외부환경에서 흔히 볼 수 있는 실 영상이다. 실

험결과에서 보여지듯이, 제안하는 시스템은 잡영에 아주 강함을 알 수 있으며, 기존의 목표물 탐지 시스템이 탐지하고자 했던 목표물 외의 다른 대상을 찾는데 많은 어려움을 가지고 있는 반면, 본 논문에서 제안된 시스템은 입력된 영상의 상향식 단서만을 사용하고 목표물에 대한 어떠한 선지식을 사용하지 않음으로써 일반적인 목표물 탐지기로 사용될 수 있는 가능성을 보여주었다.

6. 참고문헌

- [1] M.Bear, B.Connors, and M.Paradiso, Neuroscience exploring the brain, Williams and Wilkins, USA 1996
- [2] J.Duncan, J.Humphreys, Visual search and stimulus similarity, Psychological Reviews, Vol.96, pp.433-458, 1989.
- [3] O.Gallat, P.Gaussier, and J.Cocquerez, A Model of the Visual Attention to Speed up Image Analysis, Proc. of Intl. Conf. on Image Processing, Vol.1 pp.246-250, 1998.
- [4] J.Jonides, Further toward a model of the mind's eye's movement, Bulletin of the Psychonomic Society, Vol.21, No.4, pp.247-450, 1983.
- [5] B.Julesz, J.Bergen, Textons, the fundamental Elements in Preattentive Vision and Perception of Textures, Bell Systems Technical Journal, Vol.62, pp.1619-1643, 1983.
- [6] C.Koch, S.Ullman, Shifts in Selective Visual Attention : Towards the Underlying Neural Circuitry, Human Neurobiology, Vol.4 pp.219-227, 1985.
- [7] A.Treisman and G.Gelade, A Feature Integration Theory of Attention, Cognitive Psychology, Vol.12, pp.97-136, 1980.
- [8] M.Mozer, The Perception of Multiple Objects : a Connectionist Approach, MIT Press, Cambridge, MA 1991.
- [9] P.Sandon, Simulating Visual Attention, Journal of Cognitive NeuroScience, Vol.2, Nol.3, pp.213-231, 1990.
- [10] R.Milanese, H.Wechsler, S.Gil, J.Bost and T.Pun, Integration of Bottom-up and Top-down Cues for Visual Attention Using Non-Linear Relaxation, Proc. IEEE Conference of Computer Vision and Pattern Recognition, pp.781-785, 1994.
- [11] L.Itti, C.Koch, A saliency-based search mechanism for overt and covert shifts of visual attention, Vision Research, Vol.40, No.10-12, pp.1489-1506, 2000.
- [12] L.Itti, C.Gold, and C.Koch, Visual Attention and Target Detection in Cluttered Natural Scenes, Optical Engineering, Vol.40, No.9, pp.1784-1793, 2001.