

다중 해상도 필터링을 이용한 시각주의

김동환^o 임도형 최경주 이일병
 연세대학교 컴퓨터정보공학부
 {raykdh^o, dhrim, yblee}@csai.yonsei.ac.kr

Visual Attention by Multi-Resolution Filtering

Donghwan Kim^o Dohyung Rim, Yillbyung Lee
 Division of Computer and Information Engineering, Yonsei University

요 약

사람시각은 입력되는 전체영상을 한 번의 필터링으로 처리하기 보다는, 여러 단계의 필터링을 통해 현저함이 두드러지는 지역을 탐지한다. 본 논문에서는 동물의 시각정보처리 과정에 기반을 두고 주의 지역을 찾아내기 위해 다중 해상도의 복수의 필터를 사용하여 선별된 각 후보지점의 이미지를 처리하는 것을 제안한다.

1. 서 론

인간의 눈은 감각기관 중에 가장 빠르고 정확하게 사물을 인지할 수 있는 감각체계를 갖추고 있다. 시각을 통해 구체화된 실세계에 대한 정보들을 처리하기 위해 뇌로 전달되며 보다 창조적이고 정교한 정보처리 결과를 양산한다. 이러한 수많은 영상정보들을 효과적으로 처리하기 위하여 주어진 시각정보 중에서 선별되어진 필요한 영상 정보만을 처리하는 과정을 필요로 하게 된다. 인간의 망막은 경계선 추출, 색상의 밝기, 명암 감지 등의 여러 가지 크기로 조율된 시각 필드[1]로 특징을 추출한다.

본 논문에서는 이러한 동물적 시각주의의 특성을 기반으로 입력된 영상의 처리를 하기 위하여, 여러 단계의 필터링을 거쳐 현저함(Saliency)이 높은 지역을 탐지하고, 주의후보지점을 찾기 위한 다중 해상도의 복수의 필터링으로 주의후보지점들을 선별한 뒤 각 후보지점들을 또 다시 색상, 강도, 방향성에 의해 처리하여 주의가 탐지된 각 영상을 최종적으로 선별 조합하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 제안하는 시각주의 시스템의 모델을 각 모듈별 구성에 대해 기술하였고, 3장에서 실험결과와 분석에 대해 설명하고 마지막으로 4장에서 결과를 제시한다.

2. 제안하는 시스템

제안하는 시스템은 인간주의의 상황식 방식인 주어진 지식단서가 없이 시각의 자극을 갖는 특징에 의해 영상 정보를 처리하는데, 주의후보 지역을 찾아내기 위하여 다중 해상도의 복수의 필터를 사용한다. 그리고 여기서 선별되어진 후보지점들은 각각 통합모듈의 특징추출 및 통합 단계[2]를 거쳐 최종적으로 영상의 현저한 영역을 획득하게 된다.

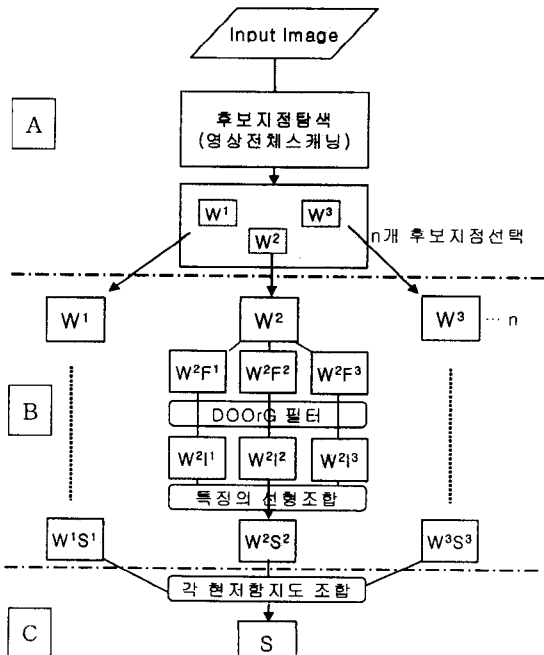


그림 1. 제안하는 시스템의 구성도

“본 연구는 과기부 뇌신경정보학사업으로부터 부분적인 지원을 받아 수행되었음.”

2.1 주의후보지점 탐색

그림 1.의 A모들은 주의 지역을 찾아내기 위한 이미지 전체 크기의 저해상도 필터를 먼저 적용하여 주의후보지점을 선택하고, 선별된 각 주의후보지점 주변의 영역을 작은 크기의 고해상도 필터에 적용한다.

여기서 후보지점은 시각주의가 두드러지는 지역을 후보군으로 n개로 선택하게 된다. 입력되는 영상에 따라 주의후보지점이 개수가 달라지는데, 일정한 크기의 윈도우를 만들어 영상 전체를 스캐닝을 한 후, 저해상도의 필터링을 하여 선택하는 것이다.

주의후보지점은 그 주변지역이 주의를 갖는 것을 의미하기 보다는, 주의와 관련된 다른 처리를 할 만한 가치가 있는 지역임을 나타낸다. 각 후보지점의 주변지역이 고해상도의 필터에 적용되어 그 지역의 현저함이 계산되는 것이다.

2.2 특징추출 및 통합

그림 1.의 B모들은 특징을 추출한 후, 계산된 각 특징맵들을 하나로 통합하는 과정이다.

인간의 시각적 주의에 영향을 미치는 색상과 명암에 관한 특징이 추출된다. 색상은 인간의 시각세포에 있는 R,G,B 들의 3가지 추상체의 반응으로 물체를 구분하며, 다른 색상과 섞이지 않은 순수한 색상으로 조율된 r,g,b,y를 생성한다.

$$\begin{aligned} r &= R-(G+B)/2, \quad g = G-(R+B)/2 \\ b &= B-(R+G)/2, \quad y = R+G-2(|r-g|+2) \end{aligned} \quad (1)$$

식(1)은 적, 녹, 청, 황 채널의 값을 나타내며, 흑, 백 자극에 대해서는 0의 값으로 출력한다. '적/녹' 과 '청/황'의 각 대립을 이용한 W^rF^2, W^gF^3 을 만든다.

$$W^rF^2 = r-g, \quad W^gF^3 = b-y \quad (2)$$

여기서 W^n 는 후보지점 영역의 개수를 뜻한다. 이렇게 생성된 기본 특징맵들은 서로 다른 3개의 독립된 특징맵이며 0~1 사이의 값으로 normalization한 후, 각각의 특징맵들로부터 방위를 탐지하고 주변과 다른 영역으로 두드러지게 하기 위해서 DOORg 필터를 이용하여 8가지 방위($\theta \in \{0, \pi/8, 2\pi/8, \dots, 7\pi/8\}$)에 조율된 8가지의 지도가 생성된다.

$$h_{x,y}(\theta) = |DOORG_{x,y}(\sigma, r_{x/y}, r_{on/off})| \quad (3)$$

$$I_{x,y}^k(\theta) = \left[\sum_{m,n} F_{m,n}^k \cdot h_{x-m,y-n}(\theta) \right]^2 \quad (4)$$

DOORg $x,y(\dots)$ 은 방위를 가진 2차원 가우시안 함수이며, 입력되는 특징맵에 대하여, $h(\theta)$ 와 회선시켜 $I^k(\theta)$ 중, 각 종류별로 가장 큰 반응이 나타나는 방위성분을 가진 지도를 찾아낸다.

여기서 생성된 중요도 지도들을 하나로 통합해야 하는데, 각각의 특징맵마다 각기 다른 영역에 주의가 가해진다고 할 수 있는데, 이를 모두 하나로 통합하여 하나의 중요도 값이 되도록 해야 한다. 각각의 중요도 지도의 특징 값들의 공간 경쟁력 특성과 통계적 정보를 바탕으로 조합하여 현저함 지도를 생성한다[2].

통합하는 방법은 우선 병렬적으로 입력된 N개의 지도를 LoG 함수로 생성된 넓은 크기의 2차원 필터로 회선한 후, 입력 받은 원래의 영상과 더한다.

$$I_{x,y}^k = \sum_{m,n} (I_{m,n}^k \cdot LoG_{x-m,y-n}) + I_{x,y}^k \quad (5)$$

$$LoG(x,y) = \left[\frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^6} \right] e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \quad (6)$$

식(6)은 LoG함수의 식이며, σ 는 가우시안의 너비를 뜻한다.

이러한 과정이 일정횟수를 반복하게 되고, 결과적으로 계산된 지도 I^k 는 다음과 같이 계산되어 3개의 현저함 지도를 만든다.

$$S_{x,y}^k = \frac{SI_{x,y}^k \cdot MinSI}{MaxSI - MinSI} \quad (7)$$

단, $SF_{x,y}^k = I_{x,y}^k \times (MaxI^k - AveI^k)^2$

식(7)에 의해 계산되어진 S^k 지도는 전반적으로 노동으로 특징값을 가진 특징맵은 그 특징 값이 저하되고, 반대의 경우는 이전보다는 높은 특징 값을 가지게 된다. 이를 통해 각 후보지점의 현저함 지도의 특징 맵을 산출할 수 있다. S^k 을 단순히 모두 합한후, 0~1 사이의 값으로 정규화시켜 하나의 현저함 지도 S를 만든다.

2.3 각 지역 현저함 지도의 통합

그림 1.의 C모들은 B모들에서 생성된 각 W^nS^n 후보지점의 현저함 지도들을 다시 통합하는 과정이다. 모여진 각 주의후보지점의 현저함 지도들은 B모들과 같은 거치게 되며, 이 때, 각 지역의 현저함이 주어진 임계치보다 클 때, 그 지역은 주의를 갖는 지역으로 표시된다. 만약 현저함이 높은 지역이 여러 곳일 경우 가장 현저한 지역

이 주의지역으로 최종 선택되는 것이다.

3. 실험 결과와 분석

제안하는 시스템의 성능을 분석하고자, 인간의 시각은 좌우로 더 넓은 가시 영역을 가지고 있기 때문에 가로비율이 세로비율보다 넓은 실영상을 이용하였다. 각 주의 후보지점들을 스캐닝하여 찾기 때문에 복잡도가 높은 실영상에서 좀더 오락지율을 조금 낮출 수가 있다.

그림 2.에서는 저해상도에서 주의 후보지점이 3가지로 선별이 된 후, 각각에 대하여 고해상도 필터링을 하게 된다. 각 주의 후보지점에서 특징추출과 특징통합과정을 거쳐 주어진 임계치에 따라 최종적으로 현저함 정도의 결과를 얻을 수 있다.

전체를 스캐닝한 후, 후보지점의 중심점을 중심으로 사이즈를 결정하기 때문에 계산량이 많아져 시간이 다소 걸리는 단점이 있다.

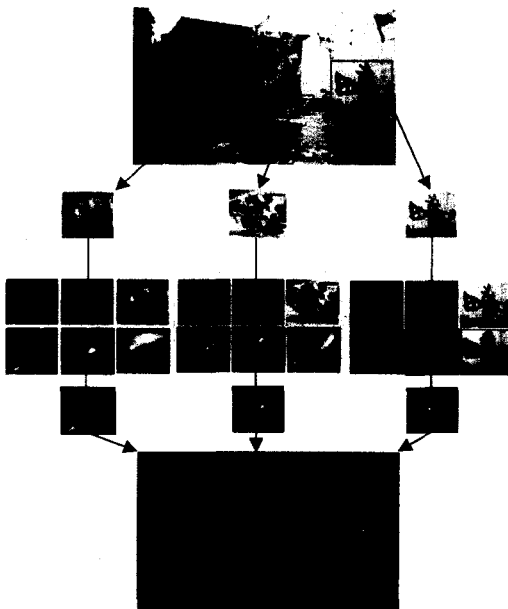


그림 2. 주의 후보지점이 3개가 나타나는 분수대 실영상에서의 실험결과

4. 결론

본 논문에서는 전체 영상을 스캐닝하여 주의가 가는 후보영역을 탐지한 후, 각 후보지점들을 다중필터링을 통하여 분석한 후, 현저함이 두드러지는 영역의 영상을 획득할 수 있게 된다. 인간의 시각정보처리 과정을 바탕으로 한, 넓은 시야영역에서 특징이 있는 후보지점들을 처

리할 수 있도록 하여 주의가 분산되는 후보지점에 대한 영역들을 각각 처리할 수 있도록 하였다.

본 시스템에서 임의로 주어진 임계치값으로 최종결과를 얻어 내었지만, 하향식 방식의 시스템과 통합하여 최종 현저함지도를 선택할 때, 임의적으로 선택할 수 있는 모델 연구가 수행되어야 할 것이며, 이를 바탕으로 시각주의를 기반한 이동 물체 추적시스템에 응용할 수 있는 가능성을 보여주었다.

5. 참고문헌

[1] Wilson, H.R, and J.R. Bergen, A four mechanism model for threshold in spatial vision, Vision Research, 19, Issue 1, 19-32, 1979.
 [2] Kyungjoo Cheoi, Yillbyung Lee, Detecting Perceptually Important Regions in an Image Based on Human Visual Attention Characteristic, LNCS, 2002.
 [3] Treisman, A.M., & Gelade, G.A., A Feature-Integration Theory of Attention. Cognitive Psychology, 12, 97-136, 1980.
 [4] Koch, C. & Ullman, S. Shifts in Selective Visual Attention : Towards the Underlying Neural Circuitry, Human Neurobiology, 4, 219-227, (1985).