

인간의 시각적 주의 능력을 이용한 컴퓨터 시각 시스템

최 경 주, 이 일 병
연세대학교 컴퓨터과학과
전화 : 02-365-4598 / 핸드폰 : 011-494-9866

Computer Vision System using the mechanisms of human visual attention

Kyungjoo Cheoi, Yillbyung Lee
Dept. of Computer Science, Kyungwon University
E-mail : kjcheoi@csai.yonsei.ac.kr

Abstract

As systems for real time computer vision are confronted with prodigious amounts of visual information, it has become a priority to locate and analyze just that information essential to the task at hand, while ignoring the vast flow of irrelevant detail. A method of achieving this is to using human visual attention mechanism. In this paper, short review of human visual attention mechanisms and some computational models of visual attention were shown. This paper can be used as the basic data for researches on development of visual attention system that can perform various complex tasks more efficiently.

I. 서 론

인간의 시각 기관을 통하여 들어오는 수많은 시각 정보들은 인간의 시각적 주의 기제에 의하여 특정 영역이나 특정 물체에 대해서 주의를 집중하게 되고, 그 부분에 대하여 더욱 자세한 분석이 일어난다. 실제적으로 컴퓨터 시각 시스템은 엄청난 양의 시각 정보를 받아들이게 되므로, 시스템은 주어진 문제에 가장 필요한 정보만을 선별적으로 받아들이고, 그 외의 정보는 무시함으로써 처리하고자 하는 문제를 보다 빠르면서도 효율적으로 처리할 수 있어야 한다. 이러한

한 이유 때문에 실시간으로 인식과 분석을 수행하는 컴퓨터 시각 시스템은 인간의 시각적 주의 기제의 역할을 시스템 내에 도입할 필요가 있다(그림 1).

인간의 시각적 주의는 망막에 입력되는 모든 영상 정보 중 일부만을 선택하는 능력으로서, 현 작업에 의미있는 특징만을 선별하여 주의를 특정 영역이나 물체에 집중시킴으로써, 시각 정보 처리 능력을 극대화하는 작업으로 복잡도 문제의 해결과 처리 용량의 한계를 극복한다는 측면에서 중요한 역할을 수행하는, 아직 공학에서 충분히 활용되지 못하고 있는 매우 중요한 동물의 지적 능력이다. 인간의 시각적 주의 기제가 행하는 기능을 실질적으로 컴퓨터 시각 시스템이 갖도록 하기 위해서는 시간상 또는 공간상으로 관심있는 영역이 설정되어져야 하고, 이를 위해 컴퓨터 자원을

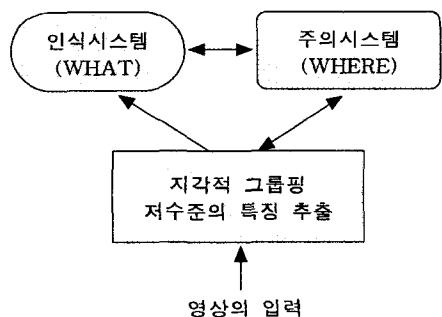


그림 1 일반적인 컴퓨터 시각 시스템에서의 주의 시스템의 작동 방식

재배치하여 주어진 작업에 맞도록 데이터 해상도를 변

경해야 한다. 또한 각 단계별로 얻어진 부분적인 결과에 기반을 두고 다음 전략을 개선하는 작업이 필요하다. 대부분의 액티브 비전(active vision)이나 시각적 주의에 관련된 모형들은 입력 영상의 일부를 필터링하여 계산의 복잡도를 낮추는 반면, 형태불변의 표준적인 객체 표상 방법을 제공하는데, 대체적으로 이런 것 만을 가지고는 입력되는 시각 정보 중 지각적으로 중요하다고 생각되는 부분을 추출하기에는 불충분하기 때문에 일반적인 시각적 주의 기법으로 사용되지 못하고 있다.

본 연구는 인간의 시각적 주의 기제를 이용하여 영상 입력 장치로 입력되는 수많은 영상 내의 정보 내에서 지각적으로 중요하다고 생각되는 부분을 자동적으로 탐지해내는 시스템 개발에 관련된 연구를 수행하는 것으로, 컴퓨터 시각 시스템에서 사용되는 시각적 주의 기법에는 어떤 것이 있는지 기존의 몇몇 시각적 주의 모형을 통해 알아본다.

II. 시각적 주의 모형

기존의 시각적 주의에 관련된 연구들은 크게 여러 기본 특징들을 결합하여 전체 영상 중에서 각 영역의 중요도를 가리키는 특징을 추출하고 이를 주의집중에 이용하는 상향식 방법과, 수행하는 작업, 찾고자 하는 물체 등의 이미 알고 있는 지식을 주의 집중 단계에 이용하는 하향식 방법의 2가지 방법에 기반을 두고 연구가 진행되었는데, 상향식 방법의 경우 수행할 작업에 적절한 모든 기본 특징을 추출하는 데 따른 어려움이 있으며, 하향식 방법은 수행할 작업에 따른 주의집중 과정을 학습할 수 있는 방법과 인식단계와의 부분적인 상호작용이 필요하다. 대부분의 유용한 모형들은 거의 상향식 방법을 채택하고 있으나, 절차 이 두 가지 방법을 결합하고자 하는 방법이 시도되고 있다. 본 논문에서는 이러한 방법론을 기반으로 한 몇몇 시각적 주의 모형을 소개한다.

Koch와 Ullman의 모형[4][5]

돌출지도(Saliency Map)와 승자전취망(Winner Take All Network)으로 이루어진 모형(그림 2)으로서, 이 모형에서는 방향성, 색상, 명암도와 같은 특징요소들을 표현하는 여러 개의 특징지도를 정의하였으며, 이런 특징들은 병렬적으로 추출된다. 각 특징지도에서 측면 억제기법이 사용되어 주변장소와 크게 다른 현저한 지점을 찾아내어 여러가지 다른 특징을 나타내는 지도들로부터 연관성 있는 하나의 정보를 만들어 돌출지도를 생성한다. 돌출지도에서 가장 현저한 지점을 선택하기

위해 WTA 알고리즘을 사용하여 수차례 반복연산을 통해 가장 높은 값을 가진 값을 추출하고 이렇게 선택된 지점의 활성화를 억제함으로써 다른 위치로의 주의 이동(shift)을 가능하게 할 수 있으며, 돌출지도에서 decay factor 도입하여 현재의 초점을 만들어내는 특정지도의 가중치를 근접성(proximity), 유사성(similarity)에 기준을 두고 억제 또는 활성화시키는 방법을 사용하기도 하였다. 하지만 승자전취망은 특징이 같은 2개의 대상이 있는 경우 수렴하지 못하는 결점을 가지고 있다.

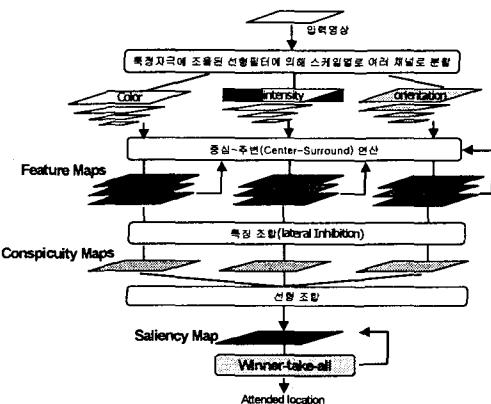


그림 2 돌출지도와 승자전취망을 사용한 모형[5]

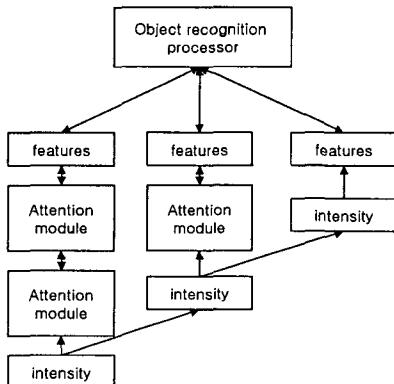
네오코그니트론[3]

네오코그니트론은 문자인식에 대표적인 선택적 주의집중 모형으로서 인공신경회로망을 사용한 최초의 주의 모형이라 할 수 있다. 문자와 같이 비록 아주 간단하기는 하지만 실질적인 데이터를 가진 응용분야에서 문제를 푸는데 목적을 두었다. 이 모형은 계층적이고도 다중의 신경회로망으로서 반복적, 자기조직화, 경쟁학습 방법을 가진다. 하나의 출력 신경세포모형이 하나의 패턴과 대응되는 출력층을 갖는 일종의 로그깊이를 가진 계층 망을 구성하고, 어떤 영상이 주어졌을 때 가장 큰 출력값을 갖는 신경세포모형을 선택하여 이 신경세포모형과 연결된 양의 피드백 경로를 형성한다. 몇 회의 수행 후에 다른 패턴과 연관된 경로들은 약화되고 해당 신경세포모형과 연관된 경로만이 남는다. 입력 패턴이 어떤 것이든 학습된 것이면 주의를 줄 수 있다는 데 장점이 있지만 네트워크가 너무 복잡하여 반복할 때마다 많은 레이어를 거치므로 주의가 해쳐야 할 영역으로의 수렴 속도가 너무 느리다는 단점을 가지고 있다.

Sandon의 피라미드식 모형[9]

인간의 시각적 주의 능력을 이용한 컴퓨터 시각 시스템

입력 영상으로부터 피라미드를 생성하여 다중 스케일 분석을 한 모형으로, 명암도, 색상과 같은 개략적인 특징 대신 입력 영상에 예지 탐지기와 같은 영상 처리 연산자를 사용하여 좀 더 복잡한 영상에 응용할 수 있도록 하였다. 특별히 특징지도는 예전 모형과 달리 직접 주의지도(attention map)에 통합되지 않고 interest 연산자에 의해 먼저 처리된다. 이러한 interest 연산자가 적용되면 특징지도의 내용이 모두 합해져서 가장 높은 값을 가지는 지점이 WTA방에 의해 추출된다.



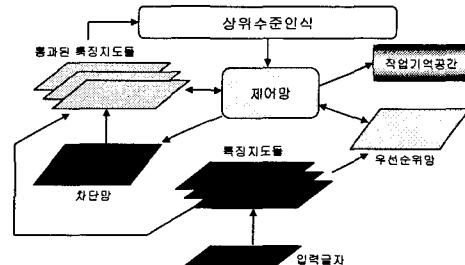
Pre-Attentive Model

시각적 주의가 가지는 최종 목표가 영상 분할이나 텍스쳐 분할 등이 아니라 하더라도, 시각적 주의는 이러한 작업과 밀접한 관련이 있다. 생물학적으로 모사한 시스템들(Biologically Plausible Systems)은 입력된 영상을 주의 전과정으로서 어떠한 고차원(high-level) 정보하나 없이 병렬적으로 처리한다. 그러므로, 이러한 처리 과정을 전처리 단계로서 주의 시스템에 도입할 수 있다. Malik와 Perona은 생물학적인 증거에 기반을 하여 텍스쳐를 인식할 수 있는 시스템 만들었는데[6], 입력영상을 DOG (Difference Of Gaussian), DOOG(Directional Differences of Offset Gaussians)와 같은 방향성을 가진 커널과 컨볼류션(convolution)을 통해 특징지도를 만들고 WTA 방법에 의해 반응이 악한 부분을 억제시킨다.

Ahmad의 VISIT[1]

인간이 주의를 주는 위치를 바꾸는 데 걸리는 시간이 평균 40-60msec이므로 주의 모형도 8-12단계 내에 주의를 바꿀 수 있어야 한다는 생각으로 모형의 수행 속도를 고려한 모형으로 그림2와 같이 나타낼 수 있다. 차단망은 전체 영상에서 특정영역을 지정하는데

주의가 가해지는 영역의 좌표를 입력받고, 특징지도에 남아있는 나머지 지점에 대한 활성화를 억제한다. 상위수준인식 모듈은 차단망에 의해 통과된 특징지도만 받아들이며, 우선순위망은 현작업과의 관계도에 따라 영상에서의 각 위치의 순위도를 정하는데, 관계도는 특징도에 서로 다른 가중치를 부여함으로써 정해진다. 우선순위망사이의 정보의 흐름을 조정하는데 있어서는 가중치를 부여할 때 탐지하고자 하는 목표물을 미리 알고 있다는 가정하에 우선순위를 결정한다. 특징지도의 값을 가중치를 두어 합한 후, 로페스 필터링을 한 후, 그 결과에 임계치를 두어 쓰레스홀드시킨다. 제어망은 주의 집중에서의 제어문제를 해결한다. 이 모형은 전체적으로 인간과 같이 병렬적인 처리단계와 순차적인 처리단계가 존재하며 인공신경회로망으로 구현되었고, 하드웨어로 구현될 수 있다는 장점이 있지만 앞서 설명한 모형들처럼 인위적으로 만들어진 자료에만 적용하여 실영상을 다루기엔 모자란 부분이 많다.



shifter circuits theory의 계산적 모형[8]

가중치의 동적 조정과 연상기억장치를 도입한 모형으로서 위치와 크기 변형에 딜 민감한 대상 표현을 위해 가중치를 동적으로 조정하여 주어진 영역에서의 정보 흐름을 상위 영역으로 몰라 보낸다. Koch 방법에 의해 시각장에서 주의를 받는 대상을 찾아내고, 다시 상위 수준으로 정보를 전파하여 연상기억이 일어나도록 한다. 이 모형에서는 상향식 방법과 하향식 방법을 결합하고 있는데 상향식 방법으로 입력된 영상에 로페스 필터링을 수행하여 잡음을 제거하고, 벌어진 캡 부분 채워 희미하게나마 흥미있는 물체가 있는 장소를 추출해낸다. 하향식 방법으로 현재 주의가 가해진 영역을 흡필드 망으로 보내어 에너지 함수에 의해 하향식 요소와 상향식 요소가 서로 동적으로 표현되도록 한다. 여기서 흡필드 망은 연상기억장치로 사용된다.

Milanese의 모형[7]

기본특징들을 다중 스케일의 특징지도 형태로 인코딩하였고, 자극의 강도(saliency) 뿐만 아니라, DAM (Distributed Associative Memory)에 의한 하향식 주의 지도(top-down map)를 생성하여 주의 이동 단계에서 사용하여 실생활 문제에 적용하였다.

FeatureGate 모형[2]

FeatureGate 모형은 국부적으로 연결되어지는 피라미드 형태의 인공신경회로망으로써 특징지도와 찾고자 하는 목표물에 대한 정보를 알고 있으면 그것을 이용해 찾고자 하는 목표물의 특징과 가장 잘 부합이 되는 특징을 가진 장소를 찾아낸다. 각 계층마다 주의 차단(gate)케이트라는 것이 있어서 그것이 정보가 각 계층별로 이동하는 흐름을 통제하는데, 상향식 시스템에서 제공받는 특징과 하향식 시스템에서 제공받는 지정된 목표물에 대한 특징을 사용하여 어떤 특징을 가진 지점을 통과시킬 것인지를 결정한다. 중간에 주의가 잘못갔는지를 탐지할 수 있는 활동수준(activation level)이 있다는 것, 또한 논리적으로 분산병렬 처리 알고리즘이라는 것 등이 다른 모형들과 다른 점이며, 특징 또한 시각적인 것 뿐만 아니라 비시각적인 것도 포함될 수 있다.

III. 정리

지금까지 몇몇 시각적 주의 시스템에 관련된 연구들을 살펴보았다. 이러한 대부분의 시각적 주의 모형들은 Treisman의 특징 모듈들[10]에 해당하는 특징지도를 계산하는 단계와 이들을 조합한 결과로 각 위치에서 특징들이 얼마나 현저히 변하는가를 표현하는 둘출지도나 우선 순위망을 계산하는 단계를 가지고 있으며, 둘출지도나 우선 주의가 주어질 후보, 혹은 후보 영역을 선정하기 위해 승자전취망이나 차단망 또는 차단회로를 사용하거나 측면억제를 통하여 주의 후보를 선정한다. 그 외에도 주의 이동 등의 전체적인 제어를 위한 중앙처리기(Central Representation), 제어망 등을 제안하거나 제어 목적의 단위들을 설정하여 사용하고 있다. 그러나 제안된 대부분의 방법들이 이론적인 측면에만 치중되어 있고, 현재 개념 정립 단계에 있는 실정이며, 대부분 시뮬레이트된 데이터에만 적용되거나 특정 문제에 한해서만 동작하여 아직은 실제 문제에 제대로 적용되지 못하고 있는 경우가 많음을 알 수 있었다. 앞으로 지금까지 개발된 시스템들을 토대로 여러 문제에 보다 효율적으로 적용될 수 있도록 하는 시각적 주의 시스템 개발에 관련된 연구가 진행되어져야 할 것이며, 본 논문은 이를 위한 기초자료로 활용

될 수 있을 것으로 사료된다.

참고문헌(또는 Reference)

- [1] S.Ahmad, *VISIT : An Efficient Computational Model of Human Visual Attention*. Ph.D. Thesis, University of Illinois at Urbana - Champaign. Also as Technical Report TR-91-049, International Computer Science Institute, Berkely, CA, 1991.
- [2] K.Cave, *The feauregate model of visual selection*, Psychological Research, Vol.62, pp.182-194, 1999.
- [3] K.Fukushima, *Neural Network Model of Selective Attention in Visual Pattern Recognition and Accociative Recall*, Applied Optics, Vol.26, No.23, pp.4985-4992, 1987.
- [4] L.Itti, *Models of Bottom-up and Top-down Visual Attention*, Ph.D. Thesis, California Institute of Technology, Pasadena, California, 2000.
- [5] C.Koch and S.Ullman, *Shifts in Selective Visual Attention : Towards the Underlying Neural Circuitry*. In L.M. Vaina(edt), *Matters of Intelligence*, Reidel Publishing, pp.115-141, 1987.
- [6] J.Malik, and P.Perona, *Preattentive Texture Discrimination with Early Vision Mechanisms*. Journal of Optical Society of America, Vol.7, No.5, 1990, pp.923-932.
- [7] 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 Conf.of Computer Vision and Pattern Recognition, 1994.
- [8] B.Olshausen, D.Essen and C.Anderson, *A neurobiological model of visual attention and Invariant pattern recognition based on dynamic routing of information*, Journal of NeuroScience, Vol.13, pp.4700-4719, 1993.
- [9] P.Sandon, *Simulating Visual Attention Journal of Cognitive NeuroScience*, Vol.2, No.3, pp.213-231, 1990.
- [10] A. Treisman and G.Gelade, *A Feature integration Theory of Attention*, Cognitive Psychology, Vol.12, pp.97-136, 1980.