

## 독립성분해석을 이용한 Saliency map 모델 구현

손준일\*, 이민호\*\*, 신장규\*\*

### Implementation of saliency map model using independent component analysis

Jun-il Sohn\*, Minho Lee\*\*, Jang-Kyoo Shin\*\*

#### 요 약

논문에서는 임의의 시각계에서 인간과 유사한 시각 응시점을 선택하기 위한 Saliency map 모델을 제안한다. 제안하는 모델은 영상의 에지 정보를 시각 응시점 결정을 위한 특징 기저로 이용한다. 자연 정지 흑백 영상에서 상호 독립적인 에지 성분들을 찾는 데 가장 좋은 방법이라고 알려진 독립성분해석 방법을 이용한다. 인간 시각계에서 시각 수용체의 비균일 분포를 고려하기 위해 카메라와 같은 시각 센서로 받은 영상을 직접 이용하는 대신에 입력 영상으로 다중 해상도를 갖는 계층 영상을 이용한다. 컴퓨터를 이용한 시뮬레이션으로부터 제안한 Saliency map을 이용하여 주어진 임의의 이미지에 대한 시각 응시점을 구한다.

#### Abstract

We propose a new saliency map model for selecting an attended location in an arbitrary visual scene, which is one of the most important characteristics of human vision system. In selecting an attended location, an edge information can be considered as a feature basis to construct the saliency map. Edge filters are obtained from the independent component analysis(ICA) that is the best way to find independent edges in natural gray scenes. In order to reflect the non-uniform density in our retina, we use a multi-scaled pyramid input image instead of using an original input image. Computer simulation results show that the proposed saliency map model with multi-scale property successfully generates the plausible attended locations.

## 1. 서 론

인간의 안구는 동시에 입력되는 많은 시각 정보를 두뇌로 전달한다. 하지만 인간의 두뇌는 이

들 정보에서 관심의 대상인 물체에 제한적으로 보다 더 많은 처리를 수행한다. 즉 인간의 시각 시스템은 선택적 주의 집중에 의해서 시각 수용체로 도달되는 많은 물체들 중에서 어느 특정한 부분에 집중한다. 선택적 주의 집중은 시각계의 자극으로 수용되는 수많은 정보들 중에서 뇌가 어떤 정보를 처리해야 하는지를 선택함에 있어서 가장 중요한 역할을 하며 지능을 갖춘 능동 시각 시스템에 필수적인 기능이다. 인간의 뇌에서의 시각 정보처리 과정에 대한 연구는 크게 top-down 방식과 bottom-up 방식으로 구분할 수 있으며 선택적 주의 집중을 수행하는 과정에는

\* 경북대학교 센서공학과  
(Dept. of Sensor Engineering, Kyungpook National University)

\*\* 경북대학교 전자전기컴퓨터공학부(School of Electrical Engineering and Computer Science, Kyungpook National University)  
<접수일자 : 2001년 4월 6일>

이 두 방식이 동시에 작용한다<sup>[1]</sup>. 고도의 인지 정보 처리 과정에 의존하는 top-down 방식은 인간의 뇌의 기능에서 가장 복잡한 활동 중의 하나이다. 따라서, 이러한 기능을 자세히 세부적으로 모델링 하는 것은 매우 어렵다. 이에 비해 bottom-up 방식은 다양한 자극을 포함하는 입력 영상에 대하여, 다른 자극들 보다 두드러진 특정 부분에 대해서 집중하는 저 차원적인 선택 주의 집중 기능이다. 이러한 뇌의 선택적 주의 집중에 관한 연구는 활발히 진행되고 있는데 그 중 두드러지는 것으로, Yagi는 이미지의 특정 부분에 대해 간단히 평균을 취하여 선택 위치를 얻었으며 안구의 신경세포의 비균일한 분포를 고려하였다<sup>[2]</sup>. 그러나 이러한 방법은 선택 위치를 결정하는데 다소 임의적인 경향이 있다. Itti와 Koch는 이보다 더 인간의 뇌와 유사한 Saliency map 모델을 제안하였다<sup>[1]</sup>. 그들은 명암과 방향성 그리고 색깔 정보를 기초로 하여 자연 이미지(natural scene)에 대한 Saliency map 모델을 제안하였다. 그러나 이 모델에서도 이들 세 가지 기본 성분들이 과연 인간의 두뇌에서 일어나는 이미지 처리를 제대로 설명하기에 충분한가에 대해선 의문이다. 한편 Barlow는 인간이 이미지로부터 특징을 추출하는 것은 중복도(redundancy)를 줄이는 과정에 의한 결과이며 특징 추출 과정은 가능한 다른 것들로부터 통계적으로 독립성을 유지하는 것이라 제안하였다<sup>[3]</sup>. Olshausen과 Field는 자연 영상의 표현에 있어서 희박성(sparseness)을 최대를 하는 신경 회로망을 고려하였고 추출된 이미지의 특징 요소가 극부적이고 방향성을 가짐을 보였다<sup>[4]</sup>. 이러한 특징은 실제로 striate 피질에서의 수용체와 유사하다. Bell과 Sejnowski는 자연영상의 독립 요소는 에지 필터이며, ICA는 이미지로부터 특징을 추출하는 가장 좋은 방법으로써 이것은 결국 이미지의 에지를 찾는 것이라 제안하였다<sup>[5]</sup>. 그러나 이들이 구한 에지 필터는 망막의 비균일적인 시세포의 분포를 고려하지 않았으므로 완벽하게 인간의 특징 추출을 구현했다고 할 수 없다.

본 논문에서는 새로운 형태의 saliency map 모델 구현을 위해 망막에서의 시세포의 비균일한 분포를 반영한 계층 형태의 다중 해상도 자연 영상을 고려한다. 이들 다중 해상도 영상은 5개의 다른 해상도 이미지를 이용하여 구성된다. 이러

한 다중 해상도 영상으로부터 특징을 추출하기 위해, 독립성분해석 알고리즘에 의해 각 해상도에 대한 에지 필터를 구하였다. 컴퓨터를 이용한 시뮬레이션에서 주어진 흑백의 정지 자연 영상에 대해 saliency map 모델을 보이고 실제로 선택된 영역들이 많은 에지 정보를 포함하고 있음을 보인다.

## II . 독립 성분 해석 알고리즘을 이용한 Saliency map 모델

Saliency map 모델은 CCD카메라를 통하여 입력되는 흑백 정지 영상에 대하여 선택 위치를 순차적으로 발생시켜 시선 이동 경로를 생성한다. 본 연구에서는 뇌에서의 이미지 처리 과정인 중복도 감소를 구현하기 위하여 ICA 알고리즘을 사용하였다<sup>[6]</sup>. 독립성분해석 알고리즘은 혼합신호로부터 원래의 독립신호를 분리하는 알고리즘으로서 신경망의 출력신호에 대한 엔트로피 또는 log-likelihood가 최대가 되거나 혹은 출력신호 상호간의 중복도나 상호 정보량이 최소가 되도록 신경망의 연결 강도를 학습함으로써 구현되어지는데 일반적으로 혼합된 음성신호의 분리나 이미지를 구성하는 독립적인 요소를 찾는 데 이용된다.

그림 1은 독립 요소  $s_i$ 가 혼합 행렬  $A$ 에 의해 혼합된  $x_i$ 로부터 신경 회로망으로 구현되어지는 역 혼합 행렬에 의해 독립요소를 분리하는 과정을 나타낸다.

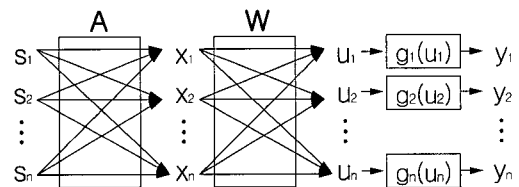


그림 1. 독립신호 분리를 위한 신경회로망의 구조  
Fig. 1. The structure of neural network for independent component analysis

그림 1에서  $N(N=1,2, \dots, n)$ 은 영상을 구성하는 독립 성분의 개수이며 결국 ICA 필터의 개수

와 일치하는데 특별히 정해진 것은 없으나 영상의 경우 대개 144 정도를 사용한다.<sup>(6)</sup>

일반적으로 log-likelihood는 다음 식과 같이 주어진다.<sup>(6)</sup>

$$L(\mathbf{u}, \mathbf{W}) = \log|\det(\mathbf{W})| + \sum_{i=1}^N \log p_i(u_i) \quad (1)$$

여기서  $\mathbf{W}$ 는 신경망의 연결 강도이며  $u_i$ 는  $i$ 번째 노드의 신경망 출력을 의미한다. 또한  $p_i(u_i)$ 는  $u_i$ 의 확률 밀도 함수(pdf)를 나타낸다.

이 식으로부터 신경회로망의 학습 알고리즘은 natural 그라디언트법에 의하여 다음과 같이 주어진다.<sup>(7)</sup>

$$\Delta \mathbf{W} = \alpha \frac{\partial L(\mathbf{u}, \mathbf{W})}{\partial \mathbf{W}} \mathbf{W}^T \mathbf{W} = [\mathbf{I} - \varphi(\mathbf{u})\mathbf{u}^T] \mathbf{W} \quad (2)$$

여기서

$$\varphi(\mathbf{u}) = -\frac{\frac{\partial p(\mathbf{u})}{\partial \mathbf{u}}}{p(\mathbf{u})} = \left[ -\frac{\frac{\partial p(u_1)}{\partial u_1}}{p(u_1)}, \dots, -\frac{\frac{\partial p(u_N)}{\partial u_N}}{p(u_N)} \right]^T$$

를 의미한다.

비선형 함수  $g_i(u)$ 를  $g_i(u) = \tanh(u_i)$ 로 정의할 경우  $\varphi(u) = 2\tanh(u)$ 와 같이 되며 최종적으로 신경망의 학습 알고리즘은 다음과 같이 된다.

$$\Delta \mathbf{W} = [\mathbf{I} - 2\tanh(\mathbf{u})\mathbf{u}^T] \mathbf{W} \quad (3)$$

각 해상도의 ICA 필터는 해당 해상도의 이미지로부터 추출한 랜덤 패치(patch)를 이용하여 구한다. 즉, 여러 개의 자연 영상 전체에 대해 여러 해상도로 변환하고 각 해상도 별로 랜덤 패치를 추출하여 각각의 학습 데이터를 구한 후, 식 3에 의해 각 해상도의 ICA 필터를 구한다.

한편, 안구의 망막 시세포는 중심으로 갈수록 밀집되어 분포하며 중심에서 멀어질수록 드물게 분포되어 있다. 이러한 망막의 구조는 현재 많이 연구되고 있는 log-polar 사상으로 모델링 할 수 있다. 하지만 log-polar 사상은 많은 계산량을 필요로 하며 필터의 적용이 어렵다. 따라서 본 연구에서는 다중 해상도를 갖는 계층 이미지로 구성하여 망막의 특성을 고려하면서 동시에 정방형의 이미지 필터를 적용할 수 있게 하였다.

다중 해상도를 갖는 계층 영상은 입력 영상을 응시점을 중심으로 여러 개의 정방형 영역으로 나누고 가장 가운데 영역을 가장 높은 해상도로 표현하고 밖으로 나갈수록 낮은 해상도로 표현한다. 영상에서 정방형 영역의 개수를 많이 하면 자연스러운 다중 해상도 영상을 얻을 수 있으나, 계산량을 고려하여 본 연구에서는 5개의 영역으로 나누었고, 각 영역의 해상도는 중심의 해상도를 1로 가정할 경우 그 다음 영역은 중심 영역 해상도의 절반인 1/2, 그 다음 영역은 1/4과 같이 순차적으로 낮추어 1/16까지 표현하였다. 이렇게 구한 입력 영상은 중심으로 갈수록 시세포가 밀집해 있는 망막의 구조를 모방 할 수 있다. 그림 2는 ICA 에지 필터를 이용하여 Saliency map 모델을 구현하는 과정을 보인다. 우선 ICA 에지 필터는 몇 개의 흑백 자연영상을 다중 해상도 이미지로 변환하고 이들 이미지로부터 무작위로 샘플을 뽑아 독립성분해석 알고리즘을 적용하여 각 해상도에 대한 ICA 필터들을 구하며, 이는 결국 에지 필터의 역할을 한다. Saliency map은 주어진 흑백의 정지 자연영상을 다중 해상도의 이미지로 변환하고 선택위치를 중심으로 높은 해상도에서 낮은 해상도 순으로 계층 이미지를 구성한다. 이러한 계층 이미지에 미리 구한 ICA 필터를 각 해상도 별로 적용한다.

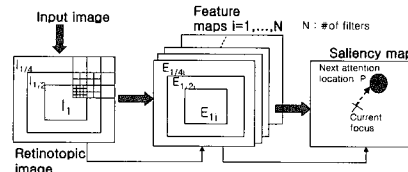


그림 2. ICA 필터를 이용한 Saliency map 모델

Fig. 2. The Saliency map model using ICA filters

식 4는 독립성분해석 알고리즘을 이용하여 구한 에지 필터들로부터 Saliency map 모델을 계산하는 것을 나타낸다.

$$E_{rj} = I_r * ICs_{rj}, \text{ for } r = 1, 1/2, 1/4, 1/8, 1/16 \quad (4)$$

for  $j = 1, \dots, N$  where  $N$  denotes the number of edge filters

식 4에서  $E_{rj}$ 는 각 해상도에 대해  $j$ 번째 ICA

에지 필터에 의해 추출된 feature map 중  $r$  해상도의 영역을 의미하며  $I$ 는 계층적 형태의 다중 해상도를 갖는 영상으로서,  $I_r$ 은 입력된 다중 해상도 영상 중  $r$  해상도 영역을 의미한다.  $ICs_{rj}$ 는  $r$  해상도의 이미지로부터 구한  $j$ 번째 ICA 에지 필터를 의미한다. 여기서 구한  $E_{rj}$ 로부터 Saliency map,  $S$ 는 식 5와 식 6에서 보여지는 것처럼 모든 영역  $(x, y)$ 에 대해 필터를 적용한 모든 feature map을 더하여 가장 큰 값을 가지는 점을 선택함으로써 구할 수 있다. 선택된 Saliency 영역  $P$ 는 입력된 다중 해상도 영상에서 가장 많은 에지 정보를 포함하므로 가장 중요한 영역이라 할 수 있다.

$$S(x, y) = \sum E_{rj}(x, y) \quad \text{for all } j \quad (5)$$

$$P = \{(x, y): \max(S(x, y) \quad \text{for all } (x, y))\} \quad (6)$$

### III. 실험 및 결과

Saliency map 모델을 구현하기 위하여 이 실험에서는 우선 그림 3에서 보여지는 4개의 흑백의 자연영상 이미지를 사용하였다. 이들 이미지에 대해 1:1, 1:2, 1:4, 1:8, 1:16 해상도를 갖는 이미지를 구하고 각 해상도들의 이미지로부터 가로 세로  $12 \times 12$  크기의 이미지 샘플 17,000개를 무작위로 추출하여  $144 \times 1$  크기의 벡터로 변환하고 ICA 알고리즘에 따라 신경 회로망을 학습하였다. 이때 신경 회로망의 가중치 행렬은  $144 \times 144$ 의 크기를 가지며 ICA 필터는 이 가중치 행렬에서  $1 \times 144$  크기의 각 행을  $12 \times 12$ 로 변환하여 구한다. 그림 4는 이러한 독립 성분분석 알고리즘을 이용한 에지 필터를 구하는 과정을 나타내며 각 해상도에 대해 이 같은 과정을 수행하여 각각의 해상도에 대한 에지 필터를 구하게 된다.

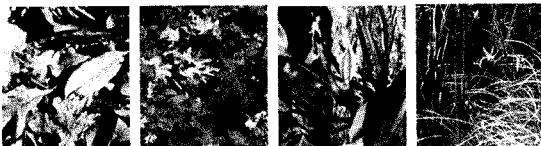


그림 3. ICA 필터를 위한 흑백 자연영상 이미지  
Fig. 3. The natural scenes with gray scale for ICA filters

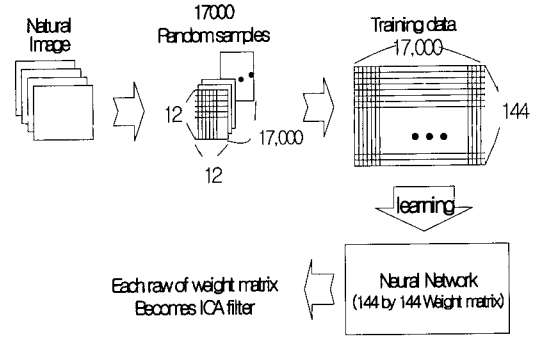


그림 4. 자연영상으로부터 에지 필터를 구하는 과정

Fig. 4. The process of obtaining the edge filters from natural scene

그림 5는 각 해상도별로 구한 144 개의 독립성분분석 알고리즘을 이용하여 에지 필터를 나타낸 것이다.

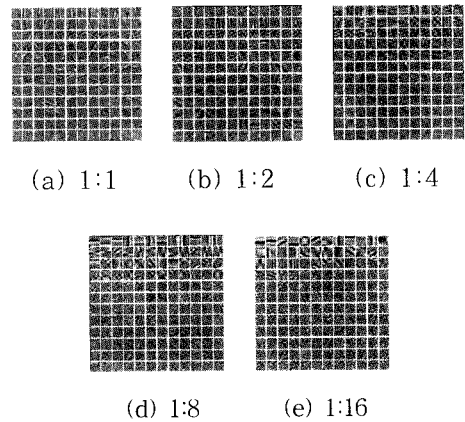


그림 5. 각 해상도에 대한 ICA 에지 필터  
Fig. 5. ICA edge filters for each resolution

그림 6은 단일 해상도를 갖는 테스트 이미지에 대해 ICA 에지 필터를 이용하여 구한 선택 경로를 나타낸 것으로 선택된 경로가 주로 영상의 일 부분에 국한되어 있음을 알 수 있다. 그림 7은 각 해상도에 대한 144개의 ICA 에지 필터를 이용하여 식 5와 식 6에 의해 두 개의 다중 해상도 테스트 이미지에 대한 선택 경로의 결과를 나타낸 것으로 영상 전체에 대해 에지 성분이 많은 영역 위주로 경로가 형성됨을 알 수 있다.



그림 6. 단일 해상도의 테스트 이미지에 대한 선택 경로

Fig. 6. The attended location for the single resolution test image



그림 7. 제안한 Saliency map 모델에 의한 테스트 이미지의 선택경로

Fig. 7. The attended location for test image using the proposed saliency map model

#### IV. 결 론

이 논문에서는 여러 해상도의 흑백 자연영상들로부터 독립성분해석 알고리즘을 이용하여 각 해상도에 대한 에지 필터를 구하였으며 주어진 영상에 이들 에지 필터를 적용하여 많은 에지 정보를 포함하는 영역의 순차적인 정보를 갖는 saliency map 모델을 구현하였다. 흑백 자연영상을 이용한 컴퓨터 시뮬레이션을 통하여 제안된 saliency map 모델이 실제 이미지에서 비교적 많은 에지 정보를 갖는 영역을 우선적으로 나타냄을 보였다. 이러한 saliency map 모델은 선택적 주의집중 능력을 갖는 지능형 능동 시각 시스템에서 전체 이미지에서 어느 위치를 우선적으로 볼 것인가 하는 선택위치를 결정지을 수 있을 것이다.

#### 감사의 글

이 연구는 뇌과학연구 개발 사업(과학기술부)과 한일 국제 공동 연구 과제의 공동 지원에 의해 수행되었습니다.

#### 참 고 문 헌

1. L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, pp. 1254-1259, 1998.
2. T. Yagi, N. Asano, S. Makita and Y. Uchikawa, "Active vision inspired by mammalian fixation mechanism", *Intelligent Robots and Systems*, pp. 39-47, 1995.
3. H. B. Barlow and D. J. Tolhurst, "Why do you have edge detectors?" *Optical society of America Technical Digest*, 23, 1992.
4. B. A. Olshausen and D.J. Field, "Natural image statistics and efficient coding", *Network Computation in Neural Systems*, 7(2), 1996.
5. A. J. Bell and T.J. Sejnowski, "The independent components of natural scenes are edge filters", *Vision research*, Pre-prints.
6. T. W. Lee, "*Independent component analysis-theory and application*", Kluwer academic publishers, 1998.
7. S. Amari, "*Neural learnign in structured parameter spaces-natural Riemannian gradient*", *Advances in Neural Information Processing Systems*, 9, MIT Press, 1997.

著 者 紹 介

**손 준 일(Jun-il Sohn, 孫峻壹)**

1971년 10월 15일생.  
 1997년 대구대학교 정보통신공학과 졸업(공학사)  
 1999년 대구대학교 대학원 정보통신공학과 졸업  
 (공학석사)  
 1999년 ~ 현재 경북대학교 대학원 센서공학과  
 박사과정  
 주관심분야 : 신경회로망, 적응신호처리

**이 민 호(Minho Lee, 李敏鎬)**

1965년 12월 19일생  
 1988년 2월 경북대학교 공과대학 전자공학과  
 (학사)  
 1995년 8월 한국과학기술원 전기 및 전자공학과  
 (석사, 박사)  
 1995년 9월 ~ 1998년 2월 한국해양대학교전  
 기공학과 (전임강사, 조교수)  
 1998년 3월 ~ 현재 경북대학교 전자전기공학부  
 /센서기술연구소(전임강사, 조교수)  
 주관심분야 : 신경회로망, 신호처리, 지능정보처  
 리시스템

**신 장 규(Jang-Kyoo Shin)**

학 력  
 1978 - 서울대학교 (공학사 - 전자공학)  
 1980 - 한국과학기술원 (공학석사 - 전자공학)  
 1991 - Colorado State University (공학  
 박사 - 전자공학)  
 주요경력  
 1980 - 1981 경북대학교 전자공학과 조교  
 1981 - 1984 경북대학교 전자공학과 전임강사  
 1985 - 1991 Colorado State University  
 연구조교  
 1984 - 1994 경북대학교 전자공학과 조교수  
 1995 - 1997 일본 토요하시 기술과학대학 교  
 환교수  
 1994 - 1999 경북대학교 전자전기공학부 부교수  
 1999 - 현재 경북대학교 전자전기공학부 교수  
 전문분야  
 1. III-V 화합물 반도체 표면물성  
 2. 자기센서 및 다공질 실리콘 광센서 등의 센  
 서개발  
 3. 집적회로 기술을 이용한 지능화 시스템 개발