

깊이정보를 이용한 선택적 주의 처리 기법

홍해민^o 조진수 이일병
연세대학교 대학원 컴퓨터학과
{yourmate^o, hamster, yblee}@csai.yonsei.ac.kr

Selective Attention Processing Method Using Depth Information

Hyemin Hong^o, Jinsu Jo, Yillbyung Lee
Dept. of Computer Science, Yonsei University

요 약

본 논문은 사람이나 로봇이 자율보행을 하기 위해서 시각적으로 한 곳을 집중적으로 선택하도록 하기 위한 논문이다. 동물들이 효율적인 시각 정보 수집을 위해 사용하는 선택적 주의 처리 기법을 기본으로, 깊이정보를 이용하여 여러 개의 주의시선 중 한 곳을 선택하도록 하는 방법을 제안하였으며, 이를 이용한 실험 결과 깊이정보를 이용한 선택적 주의 처리 기법이 유용함을 알 수 있었다. 향후 연구과제는 보행하는데 필요한 정보를 한 번의 계산으로 깊이에 따라 차례대로 정할 수 있는 3차원 맵을 만드는 것이 될 것이다.

1. 서 론

사람의 시각체계는 효율적인 시각정보 수집을 위해 선택적 주의 기법을 이용하는데 이는 시각정보처리 능력을 극대화시키기 위하여, 입력된 정보 중 그 일부를 선택하거나 여과해서 의미있는 특징만을 선별하여 집중시키는 것을 말한다. 그런데, 사람이나 로봇이 자율보행을 하기 위해서는 특징있는 부분들이나 사물들 중에서도 어느 한 곳을 집중적으로 선택하여 보행하는데 필요한 위치나 공간 정보를 얻는 것을 필요로 한다. 이런 정보를 얻기 위하여 사람은 눈을 통해 3-D세상을 2-D로 받아들이면서도 두 개의 눈을 이용하여 깊이정보를 얻어낼 수 있다.

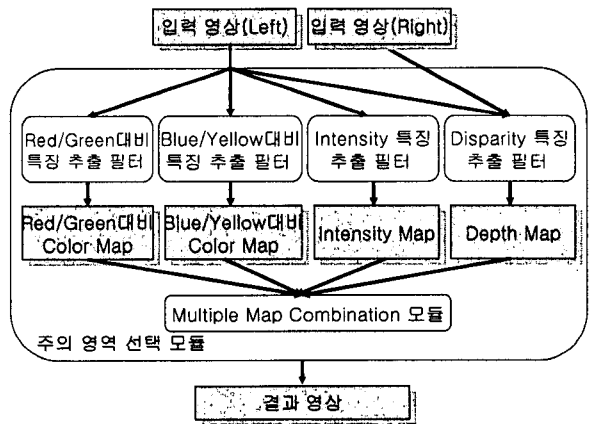
본 논문에서는 Koch와 Itti의 선택적 시각 주의 모델 중 일부를 사용하였으며, 여기에 사람과 같이 양안을 이용하여 깊이지도도를 만들고, 그 깊이정보를 이용하여 위의 문제를 해결하는 방법을 제시하고자 한다. 2장에서는 제안된 시스템과 각각의 알고리즘에 대하여 설명하였고, 3장에서는 실험 결과를 보이고 분석하였다.

2. 제안하는 시스템 구조와 알고리즘

입력 영상을 처리하는 전체적인 시스템은 [그림1]과 같다. 입력영상은 사람의 눈과 같이 두 개의 카메라를 통하여 얻은 영상을 사용하고, 주의영역 선택 모듈을 통과한 결과영상은 주의가 가는 곳 중에 한 곳을 집중적으로 선택한 영상이 나오게 된다.

주의 영역 선택 모듈은 크게 두 부분으로 나눌 수 있는데, 첫 번째는 네 가지 특징 추출 필터를 거쳐 네 가

지 특징 Map을 만들어내는 것이고, 두 번째는 이 Map 들을 Multiple Map Combination 모듈을 통과시켜 결과 영상을 만드는 부분이다.



[그림 1] 전체 시스템 구성도

첫 번째 부분은 왼쪽으로 받아진 입력 영상으로 Red/Green대비 특징 추출 필터와 Blue/Yellow대비 특징 추출 필터, Intensity 특징 추출 필터를, 양쪽 모두의 입력 영상으로 Disparity 특징 추출 필터를 거쳐, 각각 Red/Green 대비 Color Map, Blue/Yellow 대비 Color Map, Intensity Map, Depth Map 이 나오게 된다.

먼저, Red/Green 대비와 Blue/Yellow 대비는 망막의 3가지 추상체의 반응들(Red, Green, Blue에 의한)이 뇌로 전달되는 과정에서 Red/Green, Blue/Yellow 대비의

반대쌍의 색상정보로 바꾸어 전달한다는 기본개념을 기반으로 하여 만들어진 것이다. 대비가 큰 영역은 추출되며, 그렇지 않은 부분은 무시된다. 보색대비 정도를 계산한 식은 다음과 같으며, 아래 식의 r,g,b는 입력 영상의 Red/Green/Blue 채널값(0~255사이의 값)을 나타낸다.

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

$$(2) [B/Y대비] = \{b-(r+g)/2\} - \{(r+g)-2(|r-g|+b)\}$$

각 수식은, 세 개의 채널(Red, Green, Blue)에서 하나의 채널값이 나머지 두 개의 채널값과 평균적으로 얼마나 차이가 나는지 계산하게 된다.

Intensity부분은 영상의 밝기를 이용하는데, 입력 영상을 gray영상으로 변환하여 평균 밝기값을 계산한 후 전체에서 평균 이상인 부분을 0~255 사이의 값으로 정규화하는 방법을 사용한다. 아래에는 입력 영상을 gray 영상으로 변환하는 식과, 정규화 과정에 사용한 식을 나타냈다. r, g, b는 입력 영상의 R/G/B채널값, intensity는 칼라 영상의 gray 레벨값, m은 영상의 평균 intensity값, normalized는 정규화된 intensity값, max와 min은 영상 내의 최대와 최소 gray레벨 값을 의미한다.

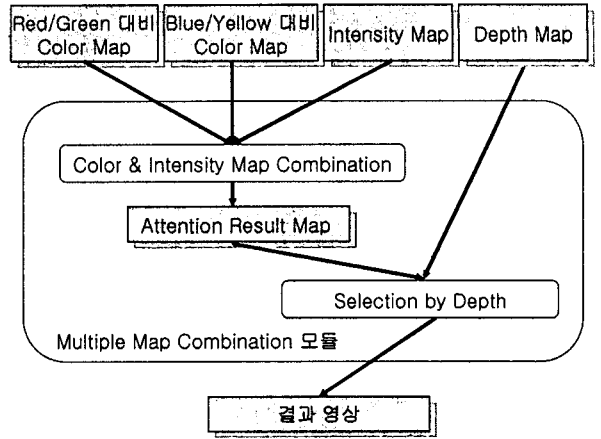
$$(1) \text{칼라 영상의 gray 레벨 변환식} \\ \text{intensity} = (r + g + b) / 3$$

$$(2) \text{normalized} = (\text{intensity}-m) * 255 / (\text{max}-\text{min})$$

Disparity부분은 입력받은 두 개의 영상으로 Depth Map을 만드는데, 이 때 사용된 알고리즘은 양안 시차를 이용하여 상대적 깊이 정보를 얻는 Stereo matching 기술을 이용하였다. Global matching과 Local matching중 계산량이 적은 Local matching을 이용하였으며, 최근 연구에 빠르며, 대체적으로 좋은 실험결과를 보인 A Fast Area-Based Stereo Matching Algorithm을 사용하였다. 이 알고리즘은 빠른 계산을 위하여 한 픽셀씩 움직일 때마다, 전체의 평균과 분산을 다시 구하지 않고, 추가 경계부분만을 더하는 방식을 이용하며, 거리 계산은 Sum of Absolute Differences 방식을 사용하였다. 식은 다음과 같으며, SAD는 거리차를 나타내는 함수, L은 left영상, R은 right영상, d는 거리, U는 추가 경계부분을 계산하는 함수이다.

$$(1) \\ \text{SAD}(x, y, d) = \sum_{i,j=-n}^n |L(x+j, y+i) - R(x+d+j, y+i)|$$

$$(2) \text{SAD}(x, y+1, d) = \text{SAD}(x, y, d) + U(x, y+1, d)$$

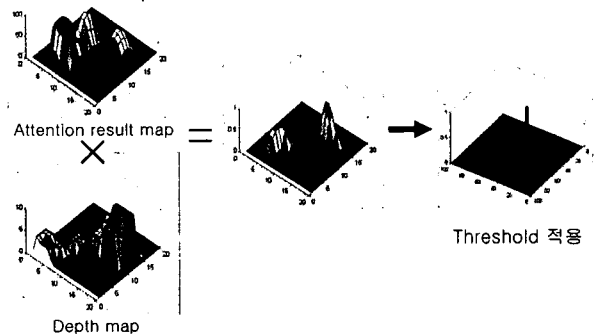


[그림 2] Multiple Map Combination 모듈

이 시스템의 두 번째 부분은 Multiple Map Combination 모듈인데, 각 필터를 거쳐 나온 4가지 결과map을 가지고 주의가 가는 영역들을 선별한 후, 최종적으로 한 곳의 영역을 선택하게 된다.

먼저 두 가지 Color Map과 Intensity Map으로 통계적 방법에 의해 Color & Intensity Map Combination을 하게 된다. 각 Map의 전역 최대값(MAPNAME_Gmax)을 구하고, 얼마다의 최대값인 L_max의 평균(averageMAPNAME_Lmax)를 구한 후, Map의 각 원소값을 업데이트하여 Attention Result Map을 만든다. 여기서 사용한 식은 다음과 같다.

$$(1) \text{new원소값} = (\text{old원소값}) * (\text{MAPNAME_Gmax} - \text{averageMAPNAME_Lmax})^2$$

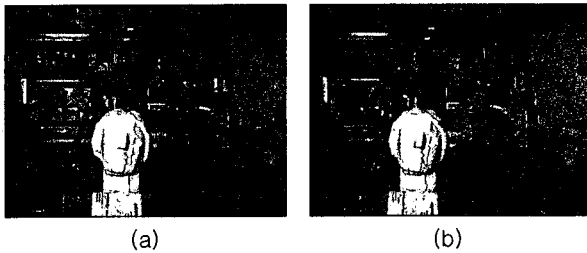


[그림 3] Selection by Depth 모듈

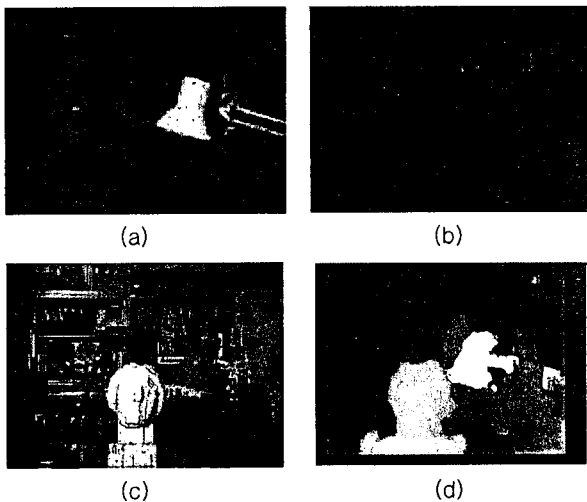
이렇게 만들어진 Attention Result Map 중 선택되었던 주의 영역은 [그림3]과 같이 Depth Map과의 연산을 통해 최종적으로 한 부분이 선택되어지는데, 연산은 Attention Result Map과 Depth Map을 곱하고, 이렇게 해서 나온 값에 threshold를 적용하여 한 영역이 선택이 되게 하도록 한다.

3. 실험 및 결과

[그림 4]에서는 두 개의 입력영상을 나타내었고, [그림 5]에서는 왼쪽 영상으로 Red/Green대비 특징 추출 필터와 Blue/Yellow대비 특징 추출 필터, Intensity 특징 추출 필터를 통과한 결과 영상과 Disparity 특징 추출 필터를 통과한 Depth Map을 나타내었다.

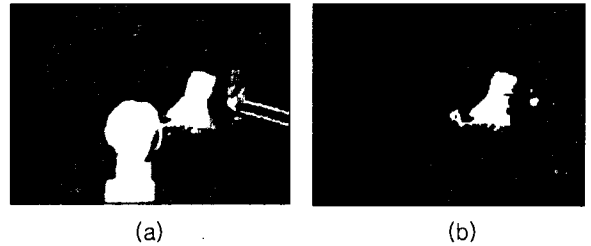


[그림 4] 두 개의 입력영상((a)-왼쪽, (b)-오른쪽)



[그림 5] 각 필터들을 거쳐 나온 결과 맵들
((a)-R/G Color Map, (b)-B/Y Color Map,
(c)-Intensity Map, (d)-Depth Map)

Attention Result Map과 Depth Map과의 연산을 통해 최종적으로 한 곳이 선택된 결과 영상을 나타내었다.



[그림 6] 결과 영상들
(a)-Attention Result Map, (b)-결과 영상)

4. 결론

실험에 의하면, 여러 개의 주의 영역이 선택되었을 때, Depth Map을 이용하여 한 곳을 선택하는데 효과적인 결과가 나왔으며, 이를 이용하여 로봇이 사람과 같이 자율보행을 하는데 있어서 도움을 줄 수 있을 것으로 보인다. 향후 연구과제는 한 곳의 응시점만을 찾는 것이 아니라, 보행하는데 있어서 필요한 정보들을 한 번의 계산으로 깊이에 따라 차례대로 정할 수 있는 3차원의 맵을 만드는 것이 될 것이다.

5. 참고문헌

- [1] A.L. Yarbus, "Eye Movements and Vision, Premium Press", New York, 1967.
- [2] C. Koch, S. Ullman, "Shifts in Selective Visual Attention : Towards the Underlying Neural Circuitry.", Human Neurobiology, Vol: 4, pp. 219-227, 1985.
- [3] L. Itti, C. Koch, E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 11, pp. 1254-1259, Nov 1998.
- [4] 최경주, 이일병, "컴퓨터 시각에서의 선택적 주의기법," 한국인지과학회 논문집. 2001. 6
- [5] L. Di Stefano, M. Marchionni, S. Mattoccia, G. Neri, "A Fast Area-Based Stereo Matching Algorithm", In Digital Image Computing: Techniques and Applications, pages 95-100, Auckland, Nov, 1997.

[그림 6]은 Multimap Combination 모듈에 의한