

다양한 움직임 영역의 크기에 적응적인 시각 주의 탐색 시스템

최병근, 최경주*
충북대학교 전자정보대학 컴퓨터학과
e-mail : icefini@chungbuk.ac.kr

A NOVEL VISUAL ATTENTION SEARCHING SYSTEM ADAPTED ON VARIOUS MOTION OF RECT SIZE

Byung Geun Choi, Kyung Joo Cheoi
Dept. of Computer Science, Chungbuk National University

요 약

본 논문에서는 동영상을 대상으로 하는 시각 주의 탐색에 대한 새로운 시스템에 대하여 설명한다. 제안하는 시스템은 기존의 공간 주의 모델에 새로운 시간 특징 추출 모듈을 추가함으로써 색상 및 명암, 형태, 방위와 같은 공간 특징 외에 움직임과 같은 시간 특징을 추가로 사용한 시각주의 탐색 모델이다. 기존 시스템과 가장 큰 차이점으로 공간 특징의 가중치 결합 방법과 움직임 특징 추출방법, 공간과 시간 특징 간 결합방법에 있다. 시스템의 성능평가를 위하여 다양한 환경의 영상을 대상으로 실험하였고 제안하는 시스템은 영상에서 사람이 시각적으로 중요하게 인지하는 영역과 부합되는 결과를 보였다.

1. 서론

능동 시각 시스템은 입력된 영상을 처리할 때 영상의 정보 중 먼저 관심이 가는 영역을 동적으로 찾아 우선적으로 처리함으로써 처리 시간을 줄이는 것을 중요하게 생각한다.

인간의 시각 주의에 대한 연구는 이러한 능동 시각 시스템 연구를 위한 기반이 될 수 있다. 인간의 시각 주의란, 시각 체계에 입력된 정보 중 그 일부를 선택하거나 (selecting) 여과하는(filtering) 기제를 일컫는 것으로, 현 작업에 의미있는 특징만을 선별하여 주의를 특정 영역이나 물체에 집중시킴으로써 시각정보처리 능력을 극대화하는 작업이라 할 수 있다.

이러한 시각 주의 연구는 기존에 많이 진행되었는데, 본 논문에서 제안하는 시스템은 특징 기반 시각주의 시스템으로서 기존의 특징기반 공간주의 시스템에 새로운 시간 특징 추출 모듈을 추가하여 공간 특징과 시간 특징을 모두 고려하여 시각 주의를 추출한다.

기존의 특징기반 시각 주의 시스템에는 몇 가지 한계점이 존재한다. Itti[1]는 WTA(Winner-Take-All)를 사용하여 시선처리를 함으로 경쟁에서 밀린 값은 추출이 되지 않는 한계점이 있었고, Park[2]의 시스템에서는 공간 현저함 지도에서 하나의 특징지도만 두드러지게 나타나는 한계점이, 홍혜민[3]의 시스템에서는 고정된 가중치를 사용

하여 다양한 영상의 환경에 정적인 결합을 하는 한계점이 있었다. 이진영[4]의 시스템에서는 FSBM(Full Search Block Matching)방법을 사용하여 움직임 벡터를 추출하였는데, 전체영상에서 블록의 차이 값이 가장 적은 것을 찾음으로 물체의 크기에 맞는 움직임을 추출하지 못한다는 한계점이 있었다. 또한 Zhai[5]는 SIFT(Scale Invariant Feature Transform)와 RANSAC 알고리즘을 사용하여 시간 현저함 지도를 생성하였는데, 움직이는 물체가 클 경우 작게 움직여 눈에 띄지 않더라도 현저함 값이 높아져 버리는 한계점이 있었다.

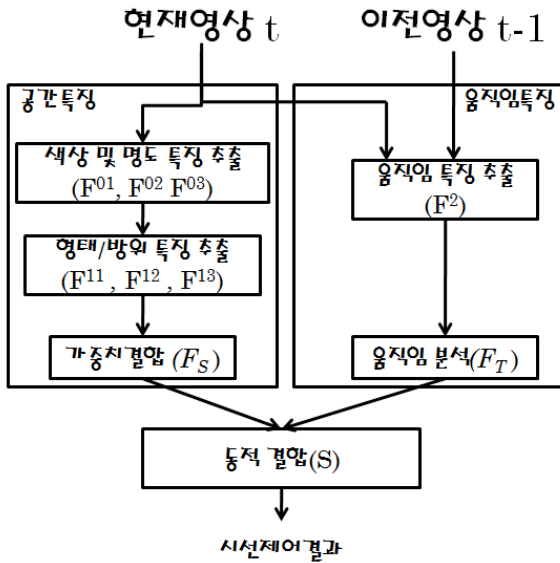
본 논문에서는 이러한 한계점을 극복하여 다양한 움직임 영역의 크기에 적응적으로 시간 특징을 추출하고, 공간 특징과 시간 특징을 동적으로 결합함으로써 사람의 시각 주의 시스템과 보다 더 부합되는 시스템을 제안한다.

2. 제안하는 시스템

그림 1은 제안하는 시스템의 구성을 나타낸 것이다. 주의 시각 탐색에 사용되는 시공간 현저함 지도는 공간 현저함 지도와 시간 현저함 지도의 동적결합으로 생성하며 공간과 시간 현저함 지도는 공간 특징 추출과 시간 특징 추출을 통하여 생성된다.

공간 특징 추출은 색상 및 명도 특징 지도를 생성하여 이 지도들의 형태 및 방위 특징을 추출하여 동적 가중치 결합으로 공간 현저함 지도를 생성한다. 색상 및 명도 특징지도는 현재영상 t 를 입력으로 받아 색상 특징 지도 2개(F^{01} , F^{02})와 명도 특징 지도

*교신저자 email : kjcheoi@chungbuk.ac.kr



(그림 1) 제안하는 시스템 구성

(F^{03}) 1개를 추출하고 Park[2]의 색상 및 명도 특징 추출 방법을 변형하여 사용한다. Park[2]의 색상 및 명도 특징 추출 방법은 특징 지도의 0이하 값을 버렸는데 0이하의 값도 중요한 정보를 가진다고 생각하여 0이하의 값을 버리지 않도록 수정하였다. 형태 및 방위 특징 추출은 색상 및 명도 특징 지도에 Park[2]이 형태 및 방위 특징 추출에 사용하였던 8개의 방향 ($\theta = \{0/8\pi, 1/8\pi, \dots, 7/8\pi\}$)을 가지는 DOoG(Difference of Oriented Gaussian)로 색상의 형태 및 방위 특징 지도 (F^{11}, F^{12})와 명도의 형태 및 방위 특징 지도(F^{13})를 추출한다. 형태 및 방위 추출은 전체에서 국지적으로 다른 형태와 방위를 가지는 물체나 영역이 주변보다 더 눈에 띄게 된다. 공간 현저함 지도(F_s)는 색상 및 명도의 형태 및 방위 특징지도들을 식(1)과 같은 동적 가중치 결합 방법으로 생성한다. 동적 가중치는 특징 지도의 활동량을 기반으로 부여되는데, 지도들의 활동량을 모두 더한 '전체 지도활동량'을 '전체 지도 활동량에서 현재 특징 지도의 활동량을 뺀 값'으로 나누어 활동량이 높은 지도일수록 더 많은 가중치를 가지도록 하였고 지도간의 연관성 또한 가지게 하였다.

$$F_s = F^{1k} * \frac{\sum_{j=1}^3 Diff(F^{1j})}{[\sum_{j=1}^3 Diff(F^{1j})] - Max(F^{1j})} \quad (1)$$

where, $Diff(F^{1k}) = Max(F^{1k}) - avg(F^{1k})$

$Max(F^{1k}) = F^{1k}$ 지도의 최대특징값

$avg(F^{1k}) = F^{1k}$ 지도의 평균값

시간 특징 추출은 움직임 특징을 Lowe[6]의 SIFT를 응용하여 추출하고 움직임 특징을 분석하여 시간 현저함 지도를 생성한다.

움직임 특징은 이전영상 ($t-1$)과 현재영상 (t)를 입력으로 받아 SIFT에 넣어 두 영상의 흥미있는 점을 추출하고, SIFT의 결합방법[6]을 사용하여 영상 간 일치하는 점들을 찾아내 움직임 벡터를 가지는 해(F^2)로 구성한다. SIFT에서 추출되는 흥미있는 점들은 대부분 영상에서 물체나 영역의 경계선(edge)에 있는 점이고, 이 점의 주변 정보를 분석하여 두 영상 간 일치하는 점을 찾기 때문에 두 영상간 정합하는 흥미있는 점들을 분석하면 물체의 크기에 적응적으로 움직이는 영역을 찾을 수 있다.

시간 현저함 지도(F_t)는 이렇게 추출한 움직임 특징 값에서 움직인 거리와 방향이 같은 움직임 벡터들의 영역을 사각형으로 그리고, 사각형으로 된 영역을 움직임 벡터의 움직인 거리 값으로 현저함 값을 부여한다.

시공간 현저함 지도(S)는 공간 현저함 지도와 시간 현저함 지도를 (2)와 같은 동적결합방법을 통해 생성한다. (2)에서, KS 와 KT 는 각각 공간 현저함 지도와 시간 현저함 지도에 부여하는 가중치를 나타낸다. 가중치는 식(3)을 사용하여 결정하는데, 시간 현저함 지도의 움직임의 크기와 움직임이 존재하는 영역의 크기, 움직임들 간의 활동량을 고려하여 시간 현저함 지도의 활동량이 높으면 시간 현저함 지도에 더 많은 가중치를 부여하고 반대의 경우 공간 현저함 지도에 더 많은 가중치를 부여한다.

$$S = KT * F_t + KS * F_s \quad (2)$$

$$KS = 1 - KT;$$

$$KT = \frac{\text{count}(t - \text{motion})}{\text{count}(t)} / 3 + \frac{\text{avg}(\text{avgMotion})}{3} + \frac{\text{count}(\text{motion} - \text{maxMotion})}{\text{count}(\text{motion})} / 3$$

where,

t : 시간특징지도블록수

$\text{count}(a)$: a 의 element수

motion : 움직임 블록수

maxMotion : 최대값 움직임 블록수

$\text{avg}(a)$: a 의 평균값

avgMotion : 움직임블록 - 최대값 움직임 블록의 수

3. 실험 및 결과

시스템 성능 평가를 위하여 다양한 환경의 영상을 실험영상으로 수집하였다. 실험 영상은 동영상을 사용하였는데, 먼저 실험영상을 공간 특징의 복잡도에 따라 높고 낮음으로 크게 2가지로 분류하였고, 이렇게 분류된 영상은 다시 시간 특징의 복잡도에 따라 움직이는 물체 크기, 움직이는 물체의 방향, 움직이는 물체의 개수가 어떠냐에 따라 3가지로 세분류하였다.

분류된 실험 영상 중 공간 특징 복잡도가 낮으면서 움직이는 물체가 크고, 움직이는 물체가 하나인 영상에 대한 실험결과가 그림 2에 보여진다. 실험 영상은 학교 안으로 진입하는 자동차 톨게이트를 찍은 것이다. 실험영상의 특징을 살펴보면, 정지영상만을 생각했을 때 공간특징만을 고려하므로 오른쪽에 위치한 적색 톨게이트 건물과 하늘색 톨게이트 지붕, 녹색 나무, 하얀색 차량에 주의가 가고, 시간 특징을 고려하면 좌측에서 우측으로 이동하는

하얀색 차량에 가장 주의가 먼저 간다. 이를 모두 고려하여 시공간 특징을 모두 적용하면 움직이는 하얀색 차량에 가장 높은 주의가 가고, 그 다음으로 적색 톨게이트에 주의가 간다.

제안하는 시스템의 공간 현저함과 시간 현저함 지도의 추출 결과를 살펴보면, 그림 2(c)는 공간 특징들을 추출한 후 이들을 모두 가중치 결합한 공간 현저함 지도를 나타내고, 그림 2(d)는 시간특징을 추출한 후 분석한 시간 현저함 지도를 나타낸다. 또한 그림 2(e)는 공간 현저함 지도와 시간 현저함 지도를 동적으로 결합한 시공간 현저함 지도로, 입력영상의 공간 및 시간 특징을 모두 고려한 최종 현저함 지도이다. 이를 기반으로 주의 강도가 높은 순서대로 시선 처리 한 결과가 그림 2(f)에 나타나 있다.

그림 2(c)의 공간 현저함 지도를 보면, 공간 시각 주의는 지도에서 가장 밝은 부분인 적색 톨게이트가 가장 높게 나타났고, 그림 2(d)의 시간 현저함 지도를 보면 시간 시각 주의는 움직이는 흰색 택시에서 가장 높게 나타났다. 공간 현저함 지도와 시간 현저함 지도를 동적 결합한 그림 2(e)를 통해 공간 현저함 지도와 시간 현저함 지도에서 보였던 주의 영역이 적절한 가중치를 두어 결합되었음을 알 수 있다.

시공간 현저함 지도로 시선 처리를 한 그림 2(f)는 그림 2(e)의 시공간 현저함 지도를 기반으로 주의 강도가 가장 높은 영역부터 일정한 부분을 순서대로 1부터 6까지 표시한 것으로서, 제안하는 시스템은 시공간적으로 현저함이 가장 높은 영역인 움직이는 흰색 택시를 우선적으로 찾아냈고, 그 다음으로 현저함이 높은 적색 톨게이트를 찾아내었다. 이는 제안하는 시스템이 사람의 시각 주의 처리와 부합하는 결과를 보인다고 할 수 있다.

4. 결론

제안하는 시스템에서는 기존의 공간 시각 주의 시스템에서 존재하던 한계점 중 하나였던 추출된 특징 지도 중 가장 활동량이 높은 지도 하나만 선택되는 특징 지도 결합 방법을 모든 공간 특징 지도를 가중치 결합함으로써 모든 공간 특징지도의 값을 가질 수 있도록 하였으며, 시간 특징 추출에서의 한계점 중 하나였던 움직이는 물체의 크기에 비적응적인 움직임 특징 추출 방법을 새로운 움직임 벡터 분석방법을 사용함으로써 다양한 움직임 크기에 적응적인 특징을 찾을 수 있게 되었다.

또한 기존의 시스템에서 시간 특징에서 움직임 영역의 크기를 고려하지 않은 채 가중치 값을 결정하고, 영상의 환경을 고려하지 않은 채 고정된 가중치로 공간과 시간 현저함 지도를 결합하였던 결합을 제안하는 시스템에서는 움직임 영역의 크기를 고려한 동적 결합으로 극복할 수 있었다.

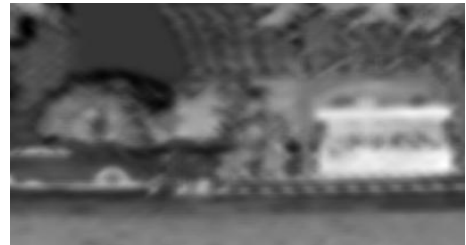
이로써 제안하는 시스템은 기존의 시스템이 가지는 성능을 모두 만족하면서 더 많은 정보 기반으로 시각 주의를 탐색하기 때문에 사람의 시각 처리 시스템과 더 부합하는 성능을 보인다고 할 수 있겠다.



(a) 현재 영상 t



(b) 이전 영상 t-1



(c) 공간 현저함 지도



(d) 시간 현저함 지도



(e) 시공간 현저함 지도



(f) 시선 처리 결과

(그림 2) 실험 결과

참고문헌

- [1] L. Itti and 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.
- [2] M. Park, K. Cheoi, "Selective Visual Attention System Based on Spatiotemporal Features," LNCS, Vol. 5068, pp.203-212, 2008.
- [3] 홍혜민, "인간의 네 가지 시각경로를 이용한 주의시각 기반 시각정보처리체계", 연세대학교 석사 학위 논문, 2006.
- [4] 이진영, "Selective Visual Attention System Based on Motion Information for Active Vision System", 고려대학교 석사 학위 논문, 2008.
- [5] Y. Zhai, M. Shah, "Visual Attention Detection in Video Sequences Using Spatiotemporal Cues," 2006.
- [6] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, 2004.