

---

# 시공간적 차를 이용한 동영상의 움직임 객체 추출

김동욱\*

## Moving Object Extraction Using Spatio-Temporal Difference

Dong-wook Kim\*

### 요 약

본 논문에서는 동영상의 움직임 객체를 추출하는 새로운 기법을 제시하고자 한다. 움직임 객체를 추출하기 위해 공간영역의 그래디언트 차와 시간영역의 차를 이용하고, 이를 바탕으로 움직임 객체는 영역별 움직임 추정 기법에 의해 얻어진다. 본 논문에서 제시된 기법의 모의실험 결과 우수한 결과를 얻을 수 있음을 알 수 있다.

### ABSTRACT

In this paper, we present a new technique for extraction of moving objects in moving image sequence. The detection method of change regions is based on spatial gradient difference and temporal difference. Moving objects are extracted by motion information and prediction error of each region. In the simulation results, the proposed technique shows good performance.

### 키워드

동영상, 움직임 객체 추출, 움직임 추정

## 1. 서 론

유선 또는 무선 통신을 통한 효과적인 멀티미디어 통신을 위해 MPEG-1, 2 등을 비롯하여 여러 가지 영상 압축기법이 널리 적용되고 있다. 이러한 영상 압축 기법은 압축 그 자체의 목적과 함께 영상의 구조적 기술(description)을 이용한 비디오 분할기법을 적용하여 좀 더 효과적인 영상 전송이 달성될 필요가 한층 높아졌으며, 이러한 노력의 결과로 MPEG-4 등이 적용되고 있다. MPEG-4 에서는 비디오 객체평면(Video object plane, VOP)의 형태로 구체화되어 내용 기반의 가능성을 추가하였다[1]. 이를 통해 장면을 여러 개의 분리된 객체로 기술하고 다양한 채

널을 통해 전송이 가능하도록 시도되기도 한다. 또한 움직임 객체는 정확한 침입자 탐지 등에 적용되기도 한다[2].

디지털 영상 열, 즉 비디오의 분석은 대개 연속적인 영상 프레임사이의 변화를 바탕으로 시도된다. 이러한 해석은 장면에 있어서 입력 영상 센서와 이동 객체사이의 상대적 움직임에 의해 프레임간 차가 발생한다는 점에서 출발한다[3]. 일반적으로 움직임 객체의 추출은 움직임 영역의 분할에 주안점을 두며, 이러한 움직임 영역의 추출은 영상 압축의 경우 압축 효율을 높이기 위함이거나 객체를 좀 더 정확히 표현하기 위해 행해진다[4].

기존의 움직임 객체 분할 기법은 다음과 같은 측

면에서 문제점을 가지고 있다. 즉, 움직임 객체의 분할은 일반적으로 프레임 차를 기반으로 행해진다. 이때, 프레임 차가 일정한 값 이상인 경우 움직임 영역으로 처리하고 그렇지 못할 경우 배경 영역으로 처리되는 것이 일반적인 방법이다. 그러나 이러한 경우 조명의 영향이나 카메라 잡음 등으로 인해 일정한 문턱 값을 지정하기가 매우 곤란한 경우가 많이 발생한다.

본 논문에서는 이러한 문제점을 해결하기 위하여 시공간적 차(spatio-temporal difference) 기반의 움직임 객체 추출 기법을 제시하고자 한다. 본 논문의 구성은 다음과 같다. 제2장에서는 시공간적 차 기반의 변화영역 추출에 관해 서술하고, 제3장에서는 이러한 결과를 바탕으로 움직임 객체 추출기법을 제시하며, 제4장에서는 제시된 기법에 대한 모의 실험결과를 보이며, 마지막으로 제5장에서 결론을 맺는다.

## II. 변화영역 검출

일반적으로 비디오 시퀀스는 고정된 배경과 한 개 이상의 움직임 객체로 구성된다. 이 때 카메라의 이동이 있는 경우 배경 영역의 칼라 혹은 밝기 값도 변화한다. 본 논문에서는 카메라의 이동이 없는 경우에 한정해서 살펴보기로 한다. 카메라의 움직임이 없는 경우에 배경의 밝기 변화는 주로 잡음과 조명의 차에 의해 발생된다. 주어진 두 개의 프레임  $f_0(i, j)$  과  $f_k(i, j)$ 에 대해 다음과 같이 정의할 수 있다. 여기서,  $f_0(i, j)$ 는 시간  $t = t_0$ 일 때의 밝기 신호를 나타내며  $f_k(i, j)$ 는  $t = t_k$ 일 때의 밝기 신호를 나타낸다.

$$f_0(i, j) = I_0(i, j) + n_0(i, j) \quad (1)$$

$$f_k(i, j) = I_k(i, j) + n_k(i, j) \quad (2)$$

여기서  $n_0(i, j)$ 과  $n_k(i, j)$ 는 카메라 및 기타 여러 가지 요인에 의한 잡음신호를 나타낸다. 식 (1), (2)를 이용하여  $f_0(i, j)$ 와  $f_k(i, j)$ 의 차를 구해보면,

$$FD_{0,k}(i, j) = dI(i, j) + dn_0(i, j) \quad (3)$$

여기서  $dI(i, j)$ 는  $I_0(i, j) - I_k(i, j)$  나타내며,  $dn_0(i, j)$ 는  $n_0(i, j) - n_k(i, j)$  나타낸다. 이 때 일반적으로 두 프레임간 움직임이나 카메라의 이동이 없다고 가정한다면  $n_0(i, j)$ 는 잡음으로서 가우스 분포를 갖는 것으로 표현할 수 있다[5]. 한편,  $dI(i, j)$ 는 밝기는 조명의 밝기 또는 방향에 매우 큰 영향을 받으며, 이는 변화영역을 구하는데 있어서 문턱 값을 고정된 값으로 지정하지 못하는 요인이 된다.

서론에서도 언급하였듯이 일반적으로 움직임 영역의 추출을 위해서 일반적으로 프레임간 차를 구하여 이를 적용한다. 즉, 식 (3)을 참조하여, 다음과 같이 정의될 수 있다.

$$CR_0(i, j) = \begin{cases} 1, & \text{if } FD_{0,k}(i, j) > T_{CR} \\ 0 & \text{else} \end{cases} \quad (4)$$

여기서,  $CR_0(i, j)$ 는 프레임  $f_0(i, j)$ 와 각 화소에 대해 프레임 차가  $T_{CR}$ 보다 크면 변화 영역이며 그렇지 않을 경우 배경 영역으로 처리된다. 이 경우, 위에서도 그 문제점을 지적하였듯이, 시간의 변화에 따라 조명의 세기가 임의적으로 변화할 경우, 문턱 값  $T_{CR}$ 의 선택은 일정한 값을 적용하기에는 많은 문제점들이 따른다.

따라서 본 논문에서는 이러한 문제점을 해결하기 위해 시공간적 밝기 변화율에 기반으로 하는 새로운 평가기준을 도입하고자 한다. 즉 제시되는 평가기준  $NFD_{0,k}(i, j)$ , 즉  $NFD[f_0(i, j), f_k(i, j)]$

$$NFD[f_0(i, j), f_k(i, j)] = FDT_{0,k}(i, j) + FDS_{0,k}(i, j)$$

로 정의되며, 여기서,  $FDT_{0,k}(i, j)$

$$FD_{0,k}(i, j) = dI(i, j) + dn_0(i, j) \quad (6)$$

를 나타내며,  $FDS_{0,k}(i, j)$ 는

$$FDS_{0,k}(i, j) = \sqrt{\left(\frac{\partial I_0(i, j)}{\partial x} - \frac{\partial I_k(i, j)}{\partial x}\right)^2 + \left(\frac{\partial I_0(i, j)}{\partial y} - \frac{\partial I_k(i, j)}{\partial y}\right)^2} + dn_s \quad (7)$$

를 나타낸다. 식 (7)에서  $dn_s$ 는 잡음에 대한 변화량을 나타낸다. 식 (7)은 2차원의 각 요소에 대한 공간적 변화율에 대한 1차 미분을 구함으로써 달성된다. 따라서, 변화 영역은 (5), (6), (7)에서 정의된 파라미터를 바탕으로 다음과 같이 새로이 정의된다.

$$CR_0(i, j) = \begin{cases} 1, & \text{if } NFD[f_0(i, j), f_k(i, j)] > T_{CR} \\ 0 & \text{else} \end{cases} \quad (8)$$

식 (8)을 이용하여 변화 영역을 검출하는 경우 뒤에 가서 모의 실험 결과에서 볼 수 있겠지만, 카메라 잡음을 비롯하여 여러 가지 유형의 잡음 성분으로 인하여 국부적인 고립 점을 비롯하여 움직임 객체 및 드러난 배경을 제외한 나머지 부분에서 변화 영역으로 판정되는 화소들이 나타난다. 본 논문에서는 이러한 점들은 각 8근방 연결요소들의 화소 수를 바탕으로 일정 이하의 화소로 구성된 변화영역은 제거하였다.

또한 변화 영역 내부에 배경영역으로 처리되는 부분이 발생한다. 이러한 경우는 움직임 객체의 영역들이 여러 개의 동일한 밝기 값 또는 동일한 칼라 값을 갖는 영역들로 구성되어 있기 때문에 움직임 객체의 영역임에도 불구하고 식 (5)를 기반으로 연산을 할 때 임계값 이하로 되는 영역이 발생되기 때문이다. 이러한 경우 본 논문에서는 변화 영역으로 둘러싸인 영역에 대해서는 내부를 채움 연산을 통해 변화영역으로 대체된다. 본 논문에서는 이러한 채움 연산을 위해 수리형태학적 필터링이 적용된다.

본 논문에서는 변화영역을 검출하기 위해 변화영역을 얻는 대신 배경영역을 얻고 원 영상으로부터 배경영역을 제거함으로써 변화영역을 얻는 방법을 적용한다. 즉, 현재 입력된 프레임울 기준으로 이전에 입력된 복수개의 영상과 식 (5)와 식 (6)을 적용하고, 만일 식 (6)을 통해 배경영역으로 판정이 되면 임시 배경영역으로 지정하고, 이 값이 복수개의 연속적인 프레임에 대해 계속 배경영역으로 유지되면 배경영역으로 최종 확정하고 그렇지 않으면 움직임 영역으로 간주된다. 이러한 방법으로 배경영역을 얻고 이를 바탕으로 원 영상에서 이를 제한으로써 변화영역을 얻는다. 본 논문에서는 현재프레임을 기준으로

이전의 연속적인 2개의 프레임에 대해 이러한 개념을 적용시키고 변화영역을 검출하였다.

### III. 움직임 객체의 검출

본문 앞에서 얻어진 변화영역을 바탕으로 움직임 객체의 추출을 수행하게 된다. 일반적으로 한 장의 영상은 크게 움직임 객체, 배경, 드러난 배경 등으로 구성된다. 본 논문에서는 변화 영역을 바탕으로 움직임 객체를 추출하기 위해 다음과 같은 단계에 의해 처리된다. 즉, a) 변화 영역의 추출, b) 변화 영역의 영역 분할, c) 변화 영역 중 드러난 배경의 제거, d) 후처리 등으로 구성된다.

#### 1) 변화 영역의 영역 분할

변화 영역은 일반적으로 크게 두 가지의 요소로 구성되어 있다. 첫째는 움직임 객체이며, 또 하나는 움직임에 의해 가려졌던 부분이 나타나게 되는 드러난 배경(uncovered background) 영역이다. 물론, 장면에 새로이 나타나는 새로운 객체의 출현도 포함되지만 여기서는 제외되었다. 이러한 구성에 대해 드러난 배경의 제거를 통해 움직임 객체를 추출할 수 있을음을 쉽게 짐작할 수 있다. 본 논문에서는 이러한 목적을 달성하기 위해 먼저 변화 영역에 대해 영역분할을 행한다. 이는 다음 단계에서 드러난 배경의 유무를 판단하기 위한 화소들의 그룹화를 의미한다.

#### 2) 드러난 배경의 제거

일반적으로 변화영역에 포함된 드러난 배경은 움직임 객체가 이전 프레임에서 현재 프레임으로 이동함으로써 발생하는 것으로써 이전 프레임에서 움직임 객체가 존재하던 위치이다. 본 논문에서는 이러한 특성을 이용하여 영역기반의 움직임 추정을 통해 움직임 객체를 추출하고, 이를 이용하여 드러난 배경을 제거하고자 한다.

p번째 영역을  $A_p$ 라고 할 때, 현재 프레임  $f_n(i, j)$ 의 변화영역  $A_p$ 과 다음 프레임  $f_{n+1}(i, j)$ 간의 움직임 추정 후 얻어지는 움직임 벡터를  $(mv_{xA_p}, mv_{yA_p})$ 라고 하자. 이 때 움직임 추정 후 얻어지는 예측 오

차의 평균  $ER_i$ 은 다음과 같이 정의된다.

$$ER_p = \frac{\sum_{i,j \in A_p} NFD[f_n(i,j), f_{n+1}(i+mv_{xA_p}, j+mv_{yA_p})]}{\sum A_p} \quad (9)$$

여기서  $\sum_{i,j \in A_p} NFD[f_n(i,j), f_{n+1}(i+mv_{xA_p}, j+mv_{yA_p})]$ 는 영역  $A_p$ 에 대한 프레임  $f_n$ 과 프레임  $f_{n+1}$ 간의 움직임 추정 후 얻어지는 예측오차이며, 두 프레임간 연산은 식 (5)의 정의에 따른다. 따라서, 위 식 (9)는 움직임 추정 결과에 대해 해당 영역의 면적  $\sum A_p$ 로 나눈 값을 의미한다. 움직임 추정 결과에 대해 다음과 같은 경우로 분류할 수 있다. 즉, 가) 움직임 벡터가 영이면서 예측오차가 적은 경우, 나) 움직임 벡터가 영이며, 예측 오차가 큰 경우, 다) 움직임 벡터가 영이 아니며, 예측 오차가 적은 경우, 라) 움직임 벡터가 영이 아니며, 예측 오차가 큰 경우.

여기서 나)의 경우 움직임 추정의 오류로 볼 수 있다. 즉, 움직임이 복잡하거나 고형체의 객체가 아닐 경우 움직임 객체임에도 불구하고 정확한 움직임 벡터의 추정이 불가능하다. 다)의 경우 움직임 벡터가 영이 아니기 때문에 움직임 객체라고 볼 수 있으며, 라)의 경우 움직임 객체이면서 고형체가 아니거나 단순한 움직임이 아닌 복잡한 움직임일 경우도 간주할 수 있다. 한편, 가)의 경우에는 움직임 벡터가 영이면서 예측 오차가 적은 경우 움직임 객체로 보기 어려우며 이 경우를 드러난 배경으로 판정할 수 있다.

### 3) 후처리

위의 단계까지의 처리 결과로서 움직임 객체의 추출은 끝나게 된다. 그러나 일반적으로 카메라의 잡음 또는 부정확한 영역분할 또는 기타 다른 이유로 인하여 경계부의 잡음 등을 제거할 필요가 있다. 이러한 경계부의 잡음제거용으로 널리 사용되는 필터로서 비선형 필터인 수리형태학적 필터링 (morphological filtering)이 자주 사용된다. 폐쇄 (Closing)는 경계부에서 들어간 부분을 매우는 용도로, 또 개방 (Opening)의 경우 튀어나온 잡음 성분을 제거하는데 매우 효과적이다. 본 논문에서는 이러한 폐쇄와 개방을 이용하여 후처리를 행한다.

## IV. 모의실험 및 검토

본 논문에서는 그림 1의 "hall monitor" 영상열 (352×288)에 대해 모의 실험을 행하였다. 변화영역을 검출하기 위한 문턱 값인  $T_{CR}$ 은 20으로 고정시키고 실험을 행하였으며, 그림 2에 변화영역을 검출한 결과를 나타내었다. 그림 2에서 배경영역 및 움직임 객체로 구성된 큰 규모의 변화 영역 외에 미세한 작은 변화 영역들을 볼 수 있다.

이러한 영역들에 대해서는 연결된 요소들의 화소수를 계산하고, 설정된 문턱 값 이하의 화소수로 구성된 요소는 제거된다. 이 때, 적용된 문턱 값은 500이다. 즉, 500개 이하의 화소로 구성된 변화 영역은 제거되었다. 그림 3은 그러한 구성 요소들이 제거된 후의 변화 영역이다. 그림 4는 이러한 변화 영역에 해당하는 원 영상의 부분을 나타낸 것이다. 이 그림에서 알 수 있듯이 움직임 객체의 움직임 방향과 반대 방향에 드러난 배경부분이 포함되어 있음을 알 수 있다.

앞서 본문에서도 제시되었듯이 그림 4를 바탕으로 영역 분할이 행해진다. 여기에 적용된 영역분할기법은 일반적인 영역확장기법이 적용되었다. 즉, 8근방의 화소들을 대상으로 임계값 이하의 차를 가지는 화소들에 대해 동일한 영역으로 처리하고, 그렇지 않을 경우 서로 다른 영역으로 처리된다. 또한 분할된 영역을 대상으로 움직임 추정을 행하고 그 때 적용된 문턱값은 10을 적용하였다. 그림 5는 드러난 배경이 제거된 후의 영상이며, 비선형 필터를 적용한 후의 최종 결과 영상이 그림 6이다. 그림 6에서 볼 수 있듯이 대부분의 영역에서 비교적 정확한 결과를 얻을 수 있으나 일부의 영역에서 부정확한 부분을 보이고 있다. 이러한 결과는 좀더 정확한 영역 분할을 통해 개선할 수 있으리라 기대된다.

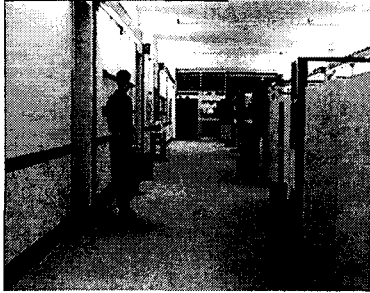


그림 1. 'hall monitor' 영상  
Fig. 1. 'hall monitor' image



그림 4. 변화영역에 대한 영상  
Fig. 4. Image of change region

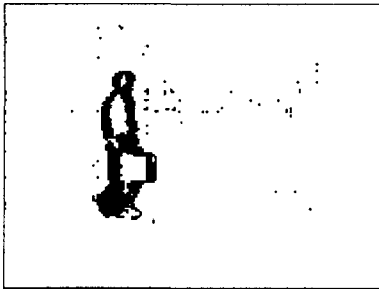


그림 2. 변화영역 검출 1  
Fig. 2. Change region detection 1



그림 5. 드러난 배경영역 제거  
Fig. 5. Deletion of uncovered background

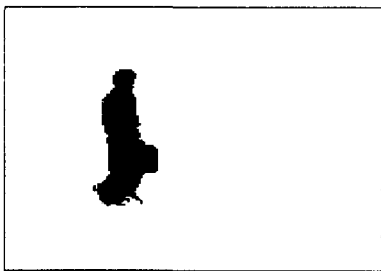


그림 3. 변화영역 검출 2  
Fig. 3. Change region Detection 2

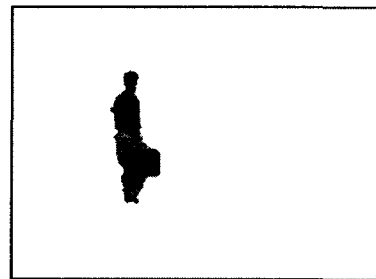


그림 6. 움직임 객체  
Fig. 6. Moving object

## V. 결 론

본 논문에서는 공간영역의 그래디언트 차와 시간 영역의 차, 즉 시공간적 차를 기반으로 움직임 객체를 얻는 새로운 기법을 제안하였다. 움직임 객체는

영역별 움직임 추정 기법에 의해 얻어진다. 제시된 기법은 도의실험을 통해 움직임 객체의 추출이 성공적으로 얻어짐을 확인할 수 있었으며, 추후의 과제로서 카메라의 이동 등 전역 움직임이 존재하는 경우에 대한 추가적인 연구가 필요할 것으로 사료된다.

### 참고문헌

- [1] MPEG-4 Video Verification Model Ver. 15.0 ISO/IEC JTC1/SC29/WG11 N3093, 1999.
- [2] R. T. Collins, A. J. Lipton, and T. K. Kanade et al, "Special section on video surveillance," Pattern Anal. Mach. Intell., vol. 22. Aug. 2000.
- [3] H. H. Nagel, "Recent advances in motion interpretation based on image sequences," Proc. IEEE ICASSP, pp.1179-1186, May 1982.
- [4] M. Tekalp, Digital Video Processing, Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [5] S. Y. Chien, S. Y. Ma, L. G. Chen "Efficient moving object segmentation algorithm using background registration technique," IEEE Trans. Circuits and System for Video tech. vol. 12, no. 7, July 2002.

### 저자소개



김동욱(Dong-wook Kim)

1987년 성균관대학교 전자공학과 졸업

1992년 중앙대학교 전자공학과(공학석사)

1996년 중앙대학교 전자공학과(공학박사)

1998년 3월~현재 전주대학교 정보기술컴퓨터공학부 조교수

※ 관심분야: 영상통신, 컴퓨터 비전