

Sub-Block 투영 영상을 이용한 카메라 동작 검출 방법

한규서⁰ 이재연 정세운 배영래
한국전자통신연구원 영상정보처리연구팀
(kyuseo, leeje, jsy, yljb)@etri.re.kr

A Camera Operation Detection using Projected Image on Sub-Blocks

Kyu Seo Han, Jae Yeon Lee, Se Yoon Jeong, Young Lae Bae
Visual Information Processing Team, Electronics and Telecommunication Research Institute

요 약

멀티미디어 환경의 발전에 따라 동영상에 대한 효과적인 검색 및 관리와 CG와 일반 영상의 합성을 위하여 영상 내의 카메라 동작 요소 검출 기법이 필요하다. 본 논문에서는 sub-block당 투영 영상을 이용한 카메라 동작 요소 검출 방법을 제안한다. 제안한 방법은 sub-block당 평균값을 이용한 투영 영상상에서 각 sub-block내에서의 x,y 방향 이동 성분을 구하여 이를 통한 Optical flow를 얻는다. 제안하는 방법은 기존의 block-matching을 통하여 optical flow를 얻는 방법보다 계산량의 감소와 계산 속도의 증가를 나타낸다. 실험 결과에서는 제안하는 방법에 의하여 얻은 optical flow를 보여주며 예측도의 증가를 보여준다.

1. 서 론

컴퓨터의 대용량화, 고속화 및 인터넷의 발전은 과거의 정지 영상을 이용한 정보 교환에서 동영상을 이용한 정보 교환으로 이동하고 있다. 이러한 동영상 처리에 있어 많은 기술들이 소개되고 있고 또한 동영상의 효과적인 분류 및 관리에 있어 특히 동영상에 포함된 내용을 기반으로 하는 기법으로의 연구가 활발히 진행되고 있는 가운데, 국제적으로는 MPEG-7과 같이 동영상 검색 및 분류를 위한 내용 요소(content descriptors)들에 대한 정의가 표준화 되어지고 있다. 이러한 요소들 중에서 영상내의 카메라 동작에 관한 요소도 지정되어 있다. 또한 특수 효과를 위한 CG와 영상 합성상에서 효과적인 합성을 위하여 일반 영상내에서의 카메라 동작 요소의 추출이 중요한 작업으로 알려져 있다.

카메라 동작 요소로는 Tilt, Zoom, Panning, Dollying등이 있으며 추출 방법으로는 Optical Flow를 이용하는 방법이 주로 제안되어지고 있다[1]. 그러나 Optical Flow를 이용하는 방법은 과도한 계산량을 요구하므로 실시간 처리가 어려운 상태이다. 이는 네트워크를 통한 실시간으로 동영상 검색을 불가능하게 만든다. 또한 각 frame의 line당 평균값을 이용한 방법[2]도 제안되었는데 모든 카메라 동작 요소를 검출하지 못하는 단점이 있다. 이를 개선하기 위하여 본 논문에서는 영상내의 sub-block당 계산되어지는 투영 영상(projected image)를 이용한 카메라 동작 요소 검출 방법을 제안한다. 제안하는 방법은 기존의 방법[2]이 frame 전체에 대한 평균값을 이용하므로써 발생할 수 있는 정보 손실을 최소화하고 또한 영상의 각 sub-region의 정보를 이용함으로써 좀 더 효과적인 카메라 동작 요소 추

출을 가능하게 만든다. 각 frame 영상을 16x16의 sub-block을 분리하여 각 block당 수평, 수직 방향의 평균값을 가지는 2개의 투영 영상을 얻는다. 이렇게 얻은 2개의 투영 영상내에서 egde 방향을 추출하여 얻은 block내의 이동 속도를 카메라 회전, 이동 속도로 표현되는 비선형 방정식에 대입하여 카메라 동작 요소를 검출한다. 제안한 방법에서는 기존의 방법[2]에서 추출이 어려운 zoom-in/out, tilt-up/down이 구별되어 보다 정밀한 카메라 동작 요소를 검출할 수 있다.

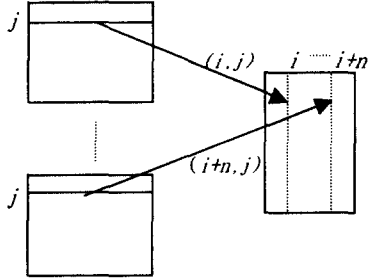
다음 장에서 제안한 방법을 설명하며 3장에서는 제안한 방법에 따른 실험과 결과를 설명한다. 4장에서는 결과에 따른 고찰로써 본 논문은 구성되어 있다.

2. Sub-Block을 이용한 카메라 동작 요소 검출

M x N 크기를 가지는 T개의 frame으로 구성된 동영상상에서의 각 sub-block에 대한 가로 방향과 세로 방향의 투영 영상은 식 (1)과 같다. 각 sub-block은 m x n의 크기를 가지며 총 D개로 구성된다.

$$P_d^h(k,l) = \frac{1}{m} \sum_{i=1}^m B_d^k(i,l) \quad (k=1,2,3,\dots,T)$$
$$P_d^v(k,l) = \frac{1}{n} \sum_{j=1}^n B_d^l(k,j) \quad (l=1,2,3,\dots,T)$$
(1)

여기서, k, l 은 frame 순서 및 투영 영상내의 pixel의 좌표를 나타내며 B_j 는 d 번째 sub_block를 의미하며 P_d^h , P_d^v 는 d 번째 sub-block의 가로, 세로 방향의 투영 영상을 나타낸다. 식(1)에 따르는 sub-block당 가로, 세로 방향의 투영 영상중 P_d^h 의 구조는 그림 1.과 같다.



<그림 1.> 투영 영상의 구조

위의 식(1)로 구해진 각 sub-block에 대한 가로 투영 영상 P_d^h 와 P_d^v 상에 존재하는 edge 들의 방향성분을 구한다. 본 논문에서는 sobel operator[3]를 사용하여 각 edge들의 방향 성분을 구하였다. 이렇게 구해진 방향 성분은 시간에 따른 이동 성분만을 가지고 있으나 실제적으로 원영상내에서 x, y 방향의 이동 성분을 추출하는 것이 필요하다. edge 방향 성분은 투영 영상의 시간을 대표하는 축의 위치에 따라 제한된다. 가로 투영 영상 P_d^h 의 경우 edge의 방향은 $\pm 90^\circ$ 로 제한되며 세로 투영 영상 P_d^v 의 경우 방향은 $180^\circ \sim 360^\circ$ 로 제한된다. 각 sub-block의 x, y 방향의 이동 성분은 식(2)와 같이 구한다.

$$\begin{aligned} x_d &= T / \tan \theta_{v,d} \\ y_d &= T * \tan \theta_{h,d} \end{aligned} \quad (2)$$

여기서 x_d, y_d 는 각 sub-block당 x, y 방향의 이동 성분이며 T 는 총 frame 수, $\theta_{v,d}, \theta_{h,d}$ 는 sub-block 당 투영 영상내의 edge 방향을 나타낸다. 식 (2)에서 얻어진 x, y 방향의 이동 성분과 영상의 총 frame 수 T 를 이용하여 sub-block당 이동 속도를 구할 수 있다. 한편 카메라 동작은 식 (3)과 같은 방정식으로 나타낼 수 있다

$$\begin{aligned} u_x &= \frac{xy}{f} R_x - f(1 + \frac{x^2}{f^2})R_y + f \tan^{-1}(\frac{x}{f}) \cdot (1 + \frac{x^2}{f^2})R_{zoom} \\ u_y &= -\frac{xy}{f} R_y - f(1 + \frac{y^2}{f^2})R_x + f \tan^{-1}(\frac{y}{f}) \cdot (1 + \frac{y^2}{f^2})R_{zoom} \end{aligned} \quad (3)$$

여기서, u_x, u_y 는 (x, y) 위치상에서의 이동 속도, R_x 는 Tilting, R_y 는 Panning, R_{zoom} 은 Zooming, f 는 카메라의 focal length를 나타낸다[4].

식 (3)은 식 (4)와 같이 표현될 수 있으며, 식 (4)의 좌변항은 Pseudo Inverse를 이용하여 구할 수 있다.

$$\begin{bmatrix} u_x \\ u_y \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \cdot [R_x R_y R_{zoom}]^T \quad (4)$$

여기서, a_{ij} 는 식 (3)에서 R_x, R_y, R_{zoom} 에 대응하는 계수를 나타낸다.

식 (4)에 각 sub-block당 이동 속도를 대입하여 카메라 동작 요소를 추출한다. 이 때 x, y 위치값은 각 sub-block의 중앙값으로 한다.

각 sub-block당 추출되어진 카메라 동작 요소값들을 수집하여 가장 많이 추출되어진 카메라 동작 요소값을 영상내의 카메라 동작 요소로 결정한다.

3. 실험 및 결과

본 논문에서 제한한 방법의 검증을 위하여 총 55개의 영상으로 이루어진 Ground-Truth를 작성하였다. 작성된 Ground-Truth의 구성은 표 1.과 같다.

<표 1.> Ground-Truth의 구성

Camera operation	Number
Tilt(up/down)	9
Pan(left/right)	17
Zoom(in/out)	23
No operation	6

실험에 사용된 영상은 352 x 288의 크기를 가지며 각 sub-block의 크기는 16 x 16으로 하였다.

우선 식 (1)에 의하여 각 sub-block당 가로, 세로 방향의 투영 영상을 구한다. 결국 가로, 세로 각 396개의 투영 영상을 만들고 이로부터 각 sub-block당 edge 방향 성분을 구한다. 구해진 edge 방향 성분을 식 (2)에 대입하여 x, y 좌표상의 이동 방향과 크기를 식(4)에 대입함으로써 각 sub-block당 카메라 동작 요소 성분을 검출한다. 검출 카메라 동작 요소는 총 8개이며 카메라 동작 요소가 없는 영상도 포함시켰다. Ground-truth에 따른 실험 결과는 표 2.와 같다.

<표 2.> 제안한 방법에 의한 카메라 동작 추출 결과

Camera operation	Number (correct/total)	Correct(%)	Avg. time(s)
Tilt(up/down)	7/9	77.7	2.22
Pan(left/right)	13/17	76.4	3.48
Zoom(in/out)	21/23	91.3	2.37
No operation	2/6	33.3	1.85

실험 결과에서 알 수 있듯이 zooming, tilting 동작은 결과 정확도가 높으나 panning과 카메라 동작이 없는 경우에는 결과 정확도가 상대적으로 낮은 것을 볼 수 있다. 이는 sub-block당 카메라 동작 요소를 추출할 때 카메라 동작 여부와는 관계없이 영상내의 강한 edge 성분이 이동 성분으로써 작용하는 것을 판별되었다. 즉, 카메라의 이동에 관계없이 영상 내부에 존재하는 edge 성분이 시간의 흐름과 무관하게 나타남으로써 공간적인 방향 성분이 추출되어져 정확한 카메라 이동 정보를 얻을 수 없었다.

그림 3.에서는 제안한 방법을 Zoom-in 요소가 있는 영상에 적용하였을 때 얻어진 Optical flow를 나타내고 있다. 그림 3.(a)는 원본 영상의 일부를 나타내고 있으며, (b), (c)는 투영 영상상에서의 시간에 따른 x, y 이동 방향 성분을 나타내고 있다. 그림 3.(d)에서 (b), (c)를 이용해 얻은 각 sub-block 상의 optical flow를 보여주고 있다.

4. 결론

이상과 같이 각 영상내의 sub-block당 평균값을 이용한 투영 영상을 이용한 이동 방향 성분을 구하여 카메라 동작 요소를 추출하는 방법에 대하여 설명하였다. 결과에서 고찰하였듯이 제안한 방법에 의해서 zooming, tilting 동작의 추출은 정확도가 높은 것으로 판별되었으나 panning 동작 추출은 정확도가 상대적으로 낮은 것으로 판별되었다. 이와 같은 현상은 카메라 동작에 의한 영상내의 변화보다 영상내의 존재하는 사물에 의한 나타나는 edge 성분에 의한 것으로 설명되어진다. 향후 연구에서는 이러한 카메라 동작과는 무관한 edge성분의 제거에 대한 개선이 필요하다.

참고 문헌

[1]J.L.Barron, et.all, " Systems and Experiment:Performance of Optical Flow Techniques". Int'l J. of Computer Vision, Vol. 12, 1994 pp.43-77

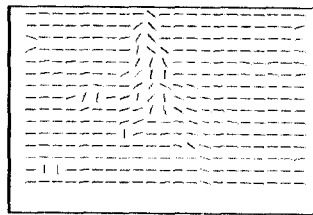
[2]P.Jolly, H.K.Kim, " Efficient automatic analysis of camera work and microsegmentation of video using spatiotemporal images". Image Commu. Vol. 8, 1996, pp.295-307

[3]M. A. Sid-Ahmed, " Image Processing", McGraw-Hill, 1994

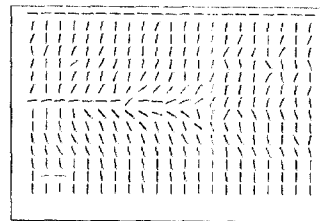
[4]ISO/IEC MPEG-7 Visual Part of XM, Version 8.0 pp 56-58



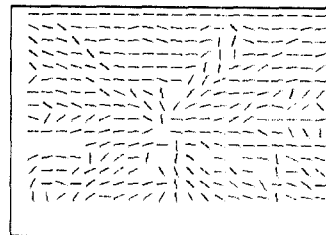
(a) 원본 영상



(b) 세로 투영 영상 P_v^h 에 대한 이동 성분



(c)가로 투영 영상 P_h^v 에 대한 이동 성분



(d)각 sub-block당 optical flow

<그림 3.> Optical Flow in Zoom-in