

움직임 벡터를 이용한 히스토그램 장면 전환 검출 기법의 개선

한영욱, 정성일, 김성재, 이시영, 김승호
경북대학교 컴퓨터공학과

Improving Histogram Scene Change Detection Method Using Motion Vector

Young-Wook Han, Seong-Il Cheong, Seong-Jae Kim, Si-Young Lee, Sung-Ho Kim
Dept. of Computer Engineering, Kyungpook National University

요 약

히스토그램 장면 전환 검출(histogram scene change detection) 기법은 입력 영상 내에 카메라 동작(camera operation)이 발생한 부분을 컷(cut)으로 나누는 문제점이 있다. 본 논문에서는 이러한 문제점을 해결하기 위해 프레임 사이의 움직임 벡터를 측정하여 카메라 동작이 일어났는지를 판단하고, 이를 이용하여 잘못된 컷의 인식을 막는다. 카메라 동작이 발생하는 샷의 경계는 컷이 될 수 없으므로, 이외의 샷에 대해 컬러 히스토그램 교집합(color histogram intersection)을 구해서 장면 전환 여부를 판단한다. 제안된 기법은 기존의 히스토그램 장면 전환 검출 기법보다 프리시전(Precision) 면에서 성능 향상을 보였다.

1. 서론

비디오 데이터는 텍스트와 같은 정적인 데이터보다 사람이 쉽게 인지할 수 있기 때문에 정보전달에 있어서 보다 효율적이다. 컴퓨터의 발달에 따라 비디오 데이터의 용량이 증가되었고, 보다 신속히 대용량의 비디오 데이터를 검색하기 위해 인덱싱(indexing)과 브라우징(browsing)하는 기능이 필요하게 되었다. 비디오 데이터를 인덱싱하고 브라우징 하기 위해서는 비디오 스트림(video stream)을 장면(scene) 단위로 나누어주는 자동화된 장면 전환 검출(scene change detection) 기술이 필요하다.

기존의 장면 전환 검출 기법들에는 화소(pixel)단위의 비교를 통한 방법[1], 컬러 히스토그램(color histogram)의 차를 이용한 방법[1-5], DC 영상을 만들어서 컬러 히스토그램과 이미지의 상관관계(correlation)를 이용한 방법[5-6] 등이 있다. 이러한 장면 전환 검출 기법 중에서 컬러 히스토그램의 차를 이용한 방법은 검출시간이 빠르고, 정확도 면에서도 우수하다. 하지만 컬러 히스토그램을 이용한 방법은 카메라 동작이 일어났을 때, 이를 장면 전환으로 판단하는 문제점이 있다.

본 논문에서는 히스토그램을 이용한 장면 전환 검출 이전에 분할된 영역의 움직임 벡터를 이용해 카메라 동작이 발생한 프레임들을 찾고, 이들을 검출 대상에서 제

외시킴으로써 히스토그램 장면 전환 검출 기법이 가지는 문제점을 보완하는 방법에 대해 기술한다.

본 논문의 구성은 다음과 같다. 제 2장에서는 움직임 벡터를 이용한 히스토그램 장면 전환 검출 기법에 대해서 기술한다. 제 3장에서는 실험 결과 및 분석을 기술하고, 마지막으로 제 4장에서는 결론과 향후 연구과제에 대해서 기술한다.

2. 장면 전환 검출

기존의 히스토그램 장면 전환 검출 기법의 문제점을 해결하기 위해 다음과 같은 과정을 따른다. 먼저 이웃한 프레임 사이의 움직임 벡터를 측정하고, 측정된 움직임 벡터를 이용해서 영상 내에 카메라 동작의 유무를 판단한다. 영상 내에 카메라 동작이 일어났거나, 변화가 없는 정지 영상으로 판단되면 다음 프레임에 대해서 같은 동작을 반복한다. 그렇지 않으면 장면 전환이 일어났을 가능성이 높다고 판단하고 컬러 히스토그램 교집합을 구해 실제로 장면 전환이 일어났는지를 판단한다.

2.1 움직임 벡터 측정

각 프레임의 움직임 벡터를 측정하기 위해 프레임을 그림 1과 같이 9개의 영역으로 분할한다. 분할된 영역은 중심에 16x16 크기의 대표 블럭(block)을 갖는다.

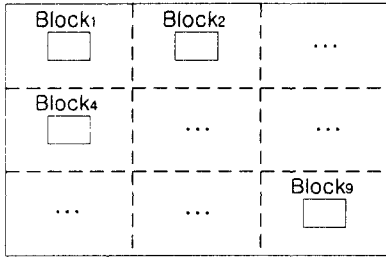


그림 1. 프레임의 영역 분할과 대표 블록 설정

움직임 벡터는 x 축 양의 방향과 이루는 각(θ)과 움직임 벡터의 크기(l)로 정의한다. 식 (1)에서 i 는 9개의 분할된 각각의 영역을 의미하고, G 는 단계 농도치(gray level)를, m 은 프레임 번호를, k 는 화소(pixel)의 위치를, p 는 측정 범위를 의미한다. 실험에서 측정범위는 x 축, y 축 방향으로 각각 ± 8 화소이다.

식 (1)의 D_Sum^i 는 측정 범위(p) 이내에서 측정한다. 측정된 D_Sum^i 중 최소값이 임계치(T_{D_Sum})를 넘지 않을 경우, 이 지점(x, y)에 대한 각(θ)과 거리(l)를 식 (2)를 통해서 구한다. 만약 주어진 범위 이내에서 최소 D_Sum^i 가 임계치(T_{D_Sum})를 넘을 경우, 이 블록에 대한 움직임 벡터는 없다고 판단한다.

$$D_Sum^i = \sum_{k=0}^{p-1} |G_m^i(k+p) - G_{m-1}^i(k)| \quad (1)$$

$$\theta = \begin{cases} \tan^{-1}(y/x) & , y > 0 \text{ or } (y=0 \text{ and } x > 0) \\ \tan^{-1}(y/x) + 180^\circ & , y < 0 \text{ or } (y=0 \text{ and } x < 0) \end{cases} \quad (2)$$

$$l = (x^2 + y^2)^{1/2}$$

2.2 카메라 동작 측정

이 절에서는 2.1 절에서 구한 분할된 영역의 각 대표 블록의 움직임 벡터를 이용해서 영상 정지 및 카메라 동작의 유무를 조사한다.

[단계 1] : 정지영상 조사

식 (3)을 통해 현재 프레임의 움직임 벡터의 크기(l)가 3 미만인 것이 몇 개인지 조사한다. 조사 결과 움직임 벡터의 크기가 3 미만인 것이 6개 이상 존재한다면, 그 프레임은 움직임이 거의 일어나지 않은 정지 영상이라고 판단한다.

$$\text{if } MV^i(l) < 3 \\ \text{then count}++ \quad , 1 \leq i \leq 9 \quad (3)$$

[단계 2] : 카메라 동작 조사

식 (4)에서 $1 \leq i \leq 9, 1 \leq j \leq 9$ 일 때 특정 방향으로 향하는 움직임 벡터의 수(MV_{max})가 6이상이면 프레임 내에 틸팅(tilting)과 패닝(panning)이 존재한다고 판단한다. 식 (5)에서 프레임의 중심을 향해 모여드는 움직임 벡터의 개수($zoom_in$)가 6이상이거나, 프레임의 중심에서 퍼져나가는 움직임 벡터의 개수($zoom_out$)가 6이상이면 프레임 내에 줌(zoom) 동작이 존재한다고 판단한다.

식 (5)의 $MV^i(a)$ 는 프레임의 i 번째 블록의 중심에서 $Block_5$ 의 중심으로 연결한 움직임 벡터의 각(a)을 의미한다.

$$\text{if } 3 \leq MV^i(l) < 12 \\ MV^i(\theta) - 30^\circ \leq MV^i(\theta) \leq MV^i(\theta) + 30^\circ \\ MV^i(l) - 3 \leq MV^i(l) \leq MV^i(l) + 3 \\ \text{then } MV_{group}^i++ \quad (4)$$

$$\text{if } MV_{group}^i > MV_{max} \\ \text{then } MV_{max} = MV_{group}^i$$

$$\text{if } MV^i(a) - 30^\circ \leq MV^i(\theta) \leq MV^i(a) + 30^\circ \\ \text{then } zoom_out++ \\ \text{if } MV^i(a) + 150^\circ \leq MV^i(\theta) \leq MV^i(a) + 210^\circ \\ \text{then } zoom_in++ \quad (5)$$

[단계 1]과 [단계 2]의 연산을 수행해서 정지영상 혹은 카메라 동작으로 판단되면, 다음 프레임의 움직임 벡터를 측정 후 [단계 1]과 [단계 2]를 반복한다. 만일 정지영상 혹은 카메라 동작이 아니라고 판단되면, 이 프레임은 장면 전환이 일어난 프레임일 가능성이 높으므로 컬러 히스토그램 교집합을 구해서 장면 전환 여부를 조사한다.

2.3 히스토그램 장면 전환 검출

이 절에서는 정지 영상 혹은 카메라 동작이 일어난 영상을 필터링 한 후 컬러 히스토그램 교집합을 이용해 장면 전환을 검출하는 방법에 대해서 기술한다. 식 (6)은 두 프레임의 컬러 히스토그램의 교집합을 정규화(normalization) 한 것이다.

$$D(m-1, m) = 1 - \sum_{i=0}^{255} \frac{\min(G_{m-1}^*(i), G_m^*(i))}{3 * G_m^*(i)} \quad (6)$$

식 (6)에서 D 는 이웃한 두 프레임의 히스토그램 차를, m 은 프레임 수를, k 는 RGB색상 중 하나를, i 는 프레임의 화소(pixel) 번호를 의미한다.

히스토그램 교집합 D 가 임계치($T_{histogram}$)를 넘을 경우, 다음 프레임에 대해서는 카메라 동작을 조사하지 않고 식 (7)과 같이 히스토그램 교집합 구하기를 반복한다. 이 과정을 반복한 후 히스토그램 교집합이 임계치를 넘지 못했을 때, 식 (8)을 통해서 장면 전환이 이루어졌는지 조사하고, 다음 프레임에 대해서는 카메라 동작을 조사한다.

$$\text{While } (D(m-1, m) > T_{histogram}) \\ \{ m = m+1, \text{count}++ \} \quad (7)$$

$$\text{if } D(m-1, m) \leq T_{histogram} \\ \text{then Case count} = 1 : \text{abrupt scene change} \\ \text{Case count} > 2 : \text{gradual scene change} \\ \text{Case count} = 2 : \text{flash light detect} \quad (8)$$

3. 실험결과 및 분석

본 논문에서 제시한 방법의 성능을 테스트하기 위하여 기존의 히스토그램 장면 전환 검출 시스템과 동일 영상들에 대해 수행 결과를 비교하였다. 실험을 위해 각각

160x120 크기, 10분 분량의 용량을 가진 드라마, 뉴스, 다큐멘터리의 3가지 비디오 스트림이 사용되었고, 세 가지 영상을 두 시스템에 적용시켰을 때의 결과는 표 1에서 표 3과 같이 나타났다. 비교 시스템들의 성능 평가를 위한 기준으로 식 (9)의 리콜(Recall)과 프리시전(Precision)을 사용하였다.[7]

$$Recall = \frac{Correct}{Correct + Missed} \quad (9)$$

$$Precision = \frac{Correct}{Correct + FalsePositive}$$

식 (9)의 리콜이 크다는 것은 올바른 샷 경계를 잃어 버리지 않았다는 것을 뜻하며, 반대로 프리시전이 높다는 것은 잘못 검출된 샷 경계가 많지 않다는 것을 의미한다. 이 두 가지 값은 임계치에 따라 한쪽이 높아지면 다른 쪽은 낮아지게 되는데, 실험에서는 두 가지의 합이 가장 크게 되도록 임계치 값을 정했다.

실험 결과 제안된 기법은 드라마와 뉴스 영상의 경우 기존의 히스토그램 장면 전환 검출 기법보다 성능 향상을 보였다. 하지만 다큐멘터리 영상에 대해서는 제안된 기법이 별다른 성능 향상을 보이지 못했다. 그 이유는 다큐멘터리 영상의 배경이 단색으로 연속되었거나, 나뭇잎과 같이 움직임이 불규칙적인 영상이 많아서 움직임 벡터 측정이 어려웠기 때문이었다.

표 1. 드라마 영상에 대한 실험 결과

	histogram 기법	제안된 기법
찾은 컷	101	97
잘못 찾은 컷	7	4
놓친 컷	2	3
Recall	97.9%	96.8%
Precision	93%	95.8%

$$T_{histogram} = 0.13, \quad T_{D_Sum} = 5500$$

표 2. 뉴스 영상에 대한 실험 결과

	histogram 기법	제안된 기법
찾은 컷	215	204
잘못 찾은 컷	32	24
놓친 컷	8	11
Recall	95.8%	94.2%
Precision	85.1%	88.2%

$$T_{histogram} = 0.13, \quad T_{D_Sum} = 5500$$

표 3. 다큐멘터리 영상에 대한 실험 결과

	histogram 기법	제안된 기법
찾은 컷	75	61
잘못 찾은 컷	19	12
놓친 컷	11	18
Recall	83.6%	73.1%
Precision	74.7%	80.3%

$$T_{histogram} = 0.14, \quad T_{D_Sum} = 5000$$

4. 결론 및 향후 연구과제

기존의 히스토그램을 이용한 장면 전환 검출 기법은 카메라 동작이 일어난 부분을 장면 전환이라고 판단하는 문제점이 있었다. 본 논문에서는 움직임 벡터를 이용하여 카메라 동작이 일어난 부분을 검출 대상에서 제외시킴으로써 히스토그램 장면 전환 검출 기법이 가지는 문제점을 보완하는 방법을 제안하였다.

본 논문에서 제안한 기법은 먼저 움직임 벡터를 이용하여 카메라 동작의 유무를 판단한 다음 카메라 동작이 일어나지 않은 프레임에 대해서만 기존의 히스토그램 기법을 적용하였다. 그 결과 카메라 동작이 일어나는 부분을 장면 전환으로 판단하는 문제점을 해결하였다.

드라마나 뉴스와 같은 영상의 경우, 움직임 벡터의 예측이 용이해서 카메라 동작의 유무를 정확하게 판단할 수 있었기 때문에 장면 전환 검출 결과가 우수했다. 그러나 움직임 벡터 예측이 어려운 숲, 나뭇잎과 같은 영상들은 검출 정확도가 다소 떨어졌다.

이러한 문제점을 해결하기 위해 영역 분할을 더욱 세분화하고, 움직임 벡터 측정 범위를 더 넓게 하여 움직임 벡터를 더 정확하게 측정하는 것이 필요하다.

참고문헌

- [1] H. Zang, A. Kangkanhalli, and S. Smoliar, "Automatic Partitioning of Full-Motion Video," *Proc. ACM Multimedia System*, ACM Press, New York, Vol. 1, pp. 10-28, 1993.
- [2] A. Nagasaka, and Y. Tanaka, "Automatic Video Indexing and Full-Video Search for Object Appearances," *Visual Database System*, Vol. 11, Elsevier, Amsterdam, pp. 113-127, 1992.
- [3] H. Ueda, et al., "Automatic Structure Visualization for Video Editing," *Proc. InterCHI '93* ACM Press, New York, pp. 137-141, 1993.
- [4] A. Hampapur, R. Jain, and T. E. Weymouth, "Indexing in Video Databases," *SPIE*, Vol. 2420, pp. 292-306, 1995.
- [5] M. M. Yeung, B. B. Yeo, W. Wolf, and B. Liu, "Video browsing using clustering and scene transitions on compressed sequences," *SPIE*, Vol. 2417, pp. 399-413, 1995.
- [6] B. L. Yeo and B. Liu, "Rapid scene analysis on compressed videos," *IEEE Trans. Circuits System Video Technology*, Vol. 5, pp. 533-544, 1995.
- [7] J. S. Boreczky and L. A. Rowe, "Comparison of Video Shot Boundary Detection Techniques," *IS&T/SPIE*, Vol. 2670, pp. 404-414, 1996.