

특징값 비유사도 영역의 누적 분포를 이용한 점진적 장면전환 검출

이종명[○], 김형준^{*}, 서병락^{**}, 김희을^{*}

한양대학교 전자통신컴퓨터공학과^{*}, 코난테크놀로지^{**}

[○]:jmlee@vision.hanyang.ac.kr, ^{*}wykim@hanyang.ac.kr

Gradual Scene Transition Detection using Summation of Feature Difference Area

Jong-Myoung Lee[○], Hyoung-Joon Kim^{*}, Byeong-Rak Seo^{**}, Whoi-Yul Kim^{*},

Division of Electrical and Computer Engineering, Hanyang University^{*},

Konan Technology Inc.^{**}

요약

장면전환의 검출은 비디오 브라우징, 검색, 요약 등에 관한 많은 응용에 유용하다. 본 논문에서는 점진적 장면전환 검출을 위해 정의된 N -길이 로컬 윈도우 내에서 비유사도 분포가 갖는 최소값만큼 상승하여 형성되는 분포를 구하고, 분포의 상단이 이루는 비유사도 값을 누적하여 설정된 임계값보다 클 경우 점진적 장면전환으로 판단하는 방법을 제안한다. 장면전환 구간에서 이루는 영역의 누적값은 최소-최대 분포를 이용하여 구할 수 있다. 실험에서 기존의 제안된 방법과 비교를 하였고 그 결과 제안된 방법에서 올바른 장면전환 검출 성능은 낮았으나 잘못 검출되는 장면전환 수는 적은 결과를 보였다. 제안된 방법은 점진적 장면전환 검출을 위한 임계값의 선택이 쉬우며, 장면전환 길이에 크게 의존하지 않는 장점이 있고 수행속도가 높아 실시간으로 처리하는데 적합하다.

1. 서론

최근 디지털 방송과 같이 디지털 멀티미디어의 사용이 생활화 되고 수요가 증가함에 따라 사용자가 쉽고 빠르게 원하는 멀티미디어 정보를 검색하고 관리하는 기술의 필요성이 급증하고 있다. 영상, 음성, 문자 정보 등 다양한 정보를 갖는 동영상 데이터는 종합적인 멀티미디어 정보를 갖고 있으며 장면전환 검출은 대용량의 동영상 정보를 효율적으로 검색, 관리하기 위한 첫 단계로서 매우 중요하다.

기존의 장면전환 검출 방법으로는, 두 프레임과의 비유사도를 비교하기 위해 픽셀의 차이 비교 [1], 히스토그램의 차이 비교 [2], 그리고 모든 프레임을 선으로 표현한 패턴을 비교 [3] 하는 방법이 있고, 연속된 두 프레임을 비교하거나 현재 프레임과 현재로부터 k 번째 떨어진 프레임을 비교하는 방법 등이 있다 [4].

장면전환은 크게 급진적 장면전환과 점진적 장면전환으로 나눌 수 있다. 급진적 장면전환은 연속된 프레임간의 비유사도 값이 임계값 이상일 때 급진적 장면전환으로 비교적 쉽게 검출할 수 있으나, 점진적 장면전환은 장면전환 길이가 일정하지 않고, 비디오 편집자가 장면을 전환하는 방법도 다양해서 정확한 장면 검출이 안되거나 잘못 검출되는 경우가 많다.

본 논문에서는 점진적 장면전환 검출을 위해 정의된 N -길이 로컬 윈도우 내에서 비유사도 분포가 갖는 최소값만큼 상승하여 형성되는 분포를 구하고, 분포의 상단이 이루는 비유사도 값을 누적하여 설정된 임계값보다 클 경우 점진적 장면전환으로 판단하는 방법을 제안한다. 점진적 장면전환에서의 누적된 비유사도 값과 장면전환이 아닌 곳에서의 누적된 비유사도 값의 차이가 크므로 장면전환 검출을 위한 임계값의 선택이 쉬우며, N -길이를 가진 윈도우가 포함되는 구간이 형성하는 영역의 누적값을 이용하므로 장면전환 길이에 크게 의존하지 않는 장점이 있다.

2. 제안된 방법

점진적 장면전환은 급진적 장면전환에 비해 상대적으로 낮은 비유사도 값이 여러 프레임에 걸쳐 지속적으로 발생하며, 고정된 임계값을 사용하여 점진적 장면전환 구간을 찾고 구간에서의 누적된 비유사도 값이 실험으로 얻어진 임계값보다 클 경우 점진적 장면전환이 발생하였다고 판단할 수 있다 [5]. 하지만 고정된 임계값으로 검출된 점진적 장면전환 구간에서의 누적된 비유사도 값은 장면전환의 종류와 길이에 따라 다르고 임계값이 정확하게 선택되지 못했을 경우 점진적 장면전환 구

간을 정확하게 찾아내지 못하여 이는 잘못된 판단으로 이어지게 된다.

본 논문에서는 점진적 장면전환 검출을 위해 정의된 N -길이 로컬 윈도우 내에서 비유사도 분포가 갖는 최소값만큼 상승하여 형성되는 분포를 구한다. N -길이를 가진 윈도우를 이동시키며 최소값의 집합을 구하고, 최소값들의 집합에서 N -길이를 가진 윈도우 내의 최대값을 구해 최소-최대 분포를 구한다. 비유사도 분포에서 최소-최대 분포를 빼주고 남은 영역의 비유사도 값을 누적하여 설정된 임계값보다 클 경우 점진적 장면전환으로 판단하는 방법으로 검출한다.

특징값은 RGB 칼라를 이용하였고 연속한 프레임과의 비유사도 판별을 크게 하기 위해 RGB 각각의 성분에서 상위 3비트 만을 선택하여 사용하였다. 연속한 프레임의 특징값에 대한 비유사도 분포는 히스토그램의 차이값을 사용하였다.

식 (1)과 같이 i 프레임 위치에서의 최소-최대값 $d_N(i)$ 를 구하기 위해 N -길이 로컬 윈도우 내에서 최소값을 찾고, 찾아진 최소값의 집합에 대한 최대값을 구한다.

$$d_N(i) = \max(\{\varepsilon_N(i+j)\}), 0 \leq j < N \quad (1)$$

식 (2)는 $f(t)$ 에 대해 N -길이 로컬 윈도우 구간에서의 최소값을 구하는 식이다. 여기서 $f(t)$ 는 비유사도 분포를 나타낸다.

$$\varepsilon_N(i) = \min[f(i+j)], 0 \leq j < N \quad (2)$$

최소-최대 분포는 전체 프레임 i 에 대해 식 (1), 식 (2)를 수행한 결과이다. 장면전환 후보지의 시작과 끝은 식 (3)과 같이 비유사도 분포 $f(t)$ 에서 최소-최대 분포 $d_N(t)$ 을 빼준 결과 $g(t)$ 가 0이 되는 두 위치를 구함으로써 찾을 수 있다.

$$g(t) = f(t) - d_N(t) \quad (3)$$

식 (4)에서와 같이 0에서부터 다음의 0까지 비유사도 값을 누적하고, 누적된 값이 설정된 임계값보다 클 경우 점진적 장면전환이 발생하였다고 판단한다. 여기서 i 와 j 는 장면전환 후보지의 시작과 끝을 나타낸다.

$$S = \left\{ S_j \mid \sum_{i=j-k}^j g(t) > T \right\} \quad (4)$$

그림 1은 $N=24$ 일 때의 최소-최대 분포이며 구간 내에서 비유사도 분포의 최소값만큼 상승하는 형태로 나타나게 된다.

그림 2에서 장면전환이 아닌 구간에서는 시작과 끝을 나타내는 0과 0사이의 거리가 짧고 누적된 비유사도 값이 작으나, 점진적 장면전환 구간에서는 시작과 끝의 거리가 길며 이때의 누적값은 장면전환이 아닌 구간보다 훨씬 큰 값을 가지게 된다. 따라서 장면전환 영역과 아닌 영역의 누적값의 차이가 크므로 장면전환 검출을 위한 임계값 선택이 쉽다. 또한 점진적 장면전환의 길이가 N -길이 로컬 윈도우보다 짧거나 길 경우에도 언제나 장면전환의 시작과 끝은 장면전환 영역을 포함하며, 이때의 누적된 비유사도 값을 사용하여 장면전환을 판단하므로 장면전환 길이에 크게 의존하지 않는다.

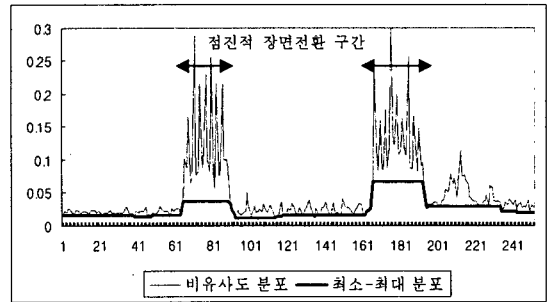


그림 1. 최소-최대 분포

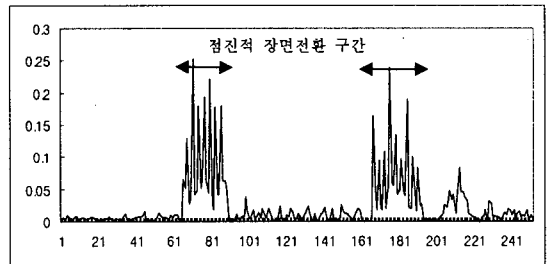


그림 2. 비유사도 분포에 최소-최대 분포를 빼준 결과

영상의 잡음, 플래시 장면과 같이 큰 비유사도를 가진 장면이 존재 할 경우 누적된 비유사도의 값은 설정된 임계값보다 높아져서 잘못된 검출을 할 수 있다. 검출된 점진적 장면전환 구간을 조사하여 짧은 시간내에 큰 비유사도의 장면이 발생하였다면 누적값에서 그 비유사도의 값을 제외시켜 잘못된 검출이 발생하지 않게 한다. 그리고 화면내에 빠른 움직임이 존재할 경우 짧은 시간 간격으로 잘못된 검출 결과가 발생하는데, 정의된 시간내에 장면전환이 연속적으로 발견되었을 경우 점진적 장면검출로 판단하지 않음으로써 잘못 검출되는 것을 막는다.

3. 실험 결과

본 논문에서 제안한 알고리즘의 성능을 비교하기 위해 [4]에서 사용한 MPEG-7 content set의 V3(뉴스), V7(드라마), V16-1(농구), V16-2(축구), V21(퀴즈)를 사용하였다 [6]. 실험은 인텔 펜티엄-4 2.4GHz CPU와 1GB 메모리를 가진 PC에서 진행되었다.

정확한 실험 결과를 얻기 위해 실험 비디오에 대한 Ground-Truth를 제작하였다. 5개의 비디오에 총 104개의 점진적 장면전환을 찾을 수 있었고 [4]에서 점진적 장면전환은 총 105개로 1개의 차이가 존재하였다. 윈도우의 길이 N 의 값은 24로 설정하고 실험한 결과는 표 1 과 같다.

표 1. 실험 결과

Video	제안한 방법		Bescos 방법 [4]	
	Recall	Precision	Recall	Precision
V3 (뉴스)	88%	77%	91%	71%
V7 (드라마)	100%	83%	100%	67%
V16-1 (농구)	89%	71%	97%	77%
V16-2 (축구)	73%	57%	92%	71%
V21 (퀴즈)	80%	100%	83%	83%

실험 결과 제안한 방법은 만족할만한 결과를 보여주었으며 하나의 프레임을 처리하기 위한 시간은 평균 11ms로 실시간 처리에 적합한 성능을 보여주었다. Bescos 방법과 비교했을 때 일부 비디오에서 제안한 방법이 더 높은 성능을 보였고, 검출되지 않은 장면은 그림 3과 같이 연속한 프레임의 특징값이 매우 유사하여 설정된 임계값보다 작을 경우에 주로 발생하였다. 제안된 방법에서 잘못된 검출이 많은 것은 주로 객체 또는 카메라 움직임에 의해 화면에 전체적인 변화가 발생할 경우이다. 특히 실험 동영상 V16-2에서 precision이 낮은 것은 그림 4와 같이 농구, 축구와 같은 스포츠 동영상에서 급격한 움직임과 화면을 가득 채우는 물체가 자주 등장하기 때문이다.

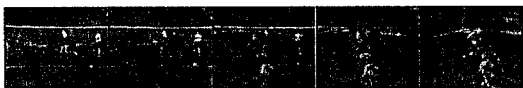


그림 3. 장면이 유사하게 변화하는 경우



그림 4. 화면에 전체적인 변화가 생겼을 경우

4. 결론

본 논문에서는 점진적 장면전환 구간에서 N -길이 로컬 윈도우 내에서 최소값을 찾고, 찾은 최소값 집합에서 최대값을 찾아 최소-최대 분포를 구하며 비유사도 분포에서 최소-최대 분포를 빼줌으로써 점진적 장면전환의 시작과 끝을 찾았다. 시작과 끝까지의 누적된 비유사도 값을 누적하고 누적된 값이 임계값보다 클 경우 점진적 장면전환으로 판단한다. 제안된 방법은 점진적 장면전환 검출을 위한 임계값의 선택이 쉬우며, 장면전환 길이에 크게 의존하지 않는 장점이 있다.

실험에서는 기존의 제안된 방법과 비교를 하였고 그 결과 제안된 방법에서 올바른 장면전환 검출 성능은 낮았으나 잘못된 검출되는 장면전환 수는 적은 결과를 보여주었다. 프레임간의 특징값이 매우 유사한 경우 점진적 장면전환이 검출되지 않았으며, 객체 또는 카메라의 움직임이 크게 발생한 경우 잘못된 검출결과가 발생하였다. 검출률 향상을 위한 특징값의 선택에 관한 연구, 카메라 움직임을 보상하는 방법 등에 관한 연구가 필요하다.

참고문헌

- [1] B. Furht and S. W. Smoliar, Video and image processing in multimedia systems, Lluwer Academic Publishers, 1995.
- [2] Ullas Gargi, Susan H. Strayer, "Performance Characterization of Video-Shot-Change Detection Methods," CirSysVideo, Vol. 10, No 1, pp. 1-13, February 2000.
- [3] Silvia Jamil Ferzoli Guimarães, Michel Couprie, "A method for cut detection based on visual rhythm," SIBGRABI, pp. 297-304, October 2001.
- [4] J. Bescos, J. M. Menendez, G. Cisneros, J. Cabrera, and J. M. Martinez, "A Unified Model for Techniques on Video-Shot Transition Detection," IEEE Multimedia, Vol. 7, No.2, pp. 293-307, April 2005.
- [5] S. W. Smoliar, H. J. Zhang, "Content Based Video Indexing and Retrieval," IEEE Multimedia, Vol. 1, pp. 66-72, 1994.
- [6] "MPEG Atlantic City Meeting," ISO/IEC JTC1/SC29/WG 11, N2466, October 1998.