SIFT를 이용한 영상 변형에 강인한 장면 전환 검출

*문원준 **서영호 ***김동욱

광운대학교 전자재료공학과

*wonjun94@kw.ac.kr

Scene Change Detection Robust to Video Distortion using SIFT

*Moon, Won-Jun **Seo, Young-Ho ***Kim, Dong-Wook

Department of Electronic Materials Engineering, Kwangwoon University

요약

본 논문에서는 비디오 제작 및 유통의 활성화에 따라 필요성이 높아지고 있는 장면 전환을 검출하는 방법을 제안한다. 유통 과정에서 해상도 변환, 자막 삽입, 압축, 영상 반전 등의 다양한 변형이 추가되더라도 동일하게 장면 전환을 검출해야 하므로 전처리 과정과 SIFT를 이용한 특징 추출, 변형을 고려한 매칭 방법을 이용하여 프레임 간의 매칭률을 계산한다. 또한 매칭률의 임계값을 기준으로 장면 전환 여부를 판단한다. 원본 비디오에서의 특징을 가지고 다양한 변형이 가해진 비디오에서의 특징과 매칭률을 계산하여 유효성을 판단한다.

1. 서론

최근 Youtube 등의 비디오 공유 사이트 활성화로 인해 많은 영상들이 제작, 유통되면서 비디오의 압축, 편집, 검색 등을 위해서 장면전환 검출의 필요성이 높아지고 있다[1]. 유통 과정에서 해상도 변환, 자막 삽입, 압축, 영상 반전 등의 추가적인 변형을 가하게 되는데 이러한 변형에도 장면전환이 동일하게 검출되어야 한다.

따라서 본 논문에서는 영상의 특징을 찾아내는 SIFT와 변형을 고려한 처리과정을 수행하여 영상 변형에 강인한 장면전환 검출 알고리즘을 제안한다[2].

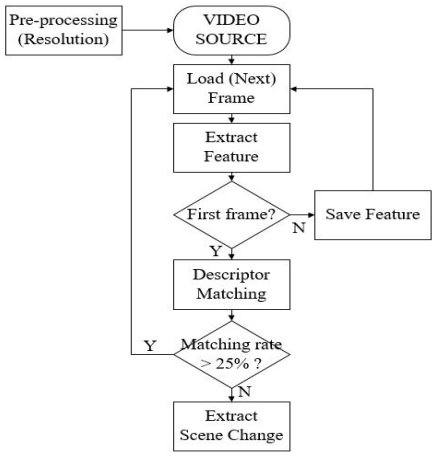

그림 1. 제안하는 화면전환 검출 알고리즘

2. 장면 전환 검출

제안하는 장면 전환 검출 방식은 그림 1과 같다. 전처리 과정에서는 해상도 변형을 고려하여 특정 해상도로 고정을 시킨다. 비디오의 프레임에서 SIFT를 이용하여 특징을 추출하고, 이를 영상의 정보로 저장해 두었다가, 다음 프레임의 SIFT로 추출한 특징과 매칭을 수행한다. 매칭과정은 SIFT에서 추출하는 서술자를 이용하여 매칭을 하는데. 서술자의 계산 방법은 [2]를 따라 각각의 특징마다 128차원의 값을 가진다. 매칭을 할 때에는 모든 특징이 서술자간의 유클리디안 거리를 계산하여 가장 작은 값과 두번째 작은 값을 구하고, 두 값의 비율이 임계값보다 낮을 때 매칭된다고 정의를 하는데, 실험을 통해 임계값을 0.6으로 정의하였다. 하지만 위의 매칭 방법은 영상이 반전되었을 때에는 그림 2처럼 0~7차원에 해당하는 a블록이 24~31차원으로 투영되고, 방향도 바뀌므로 같은 영상이지만 전혀 매칭이 되지 않는다. 따라서 서술자 매칭을 할 때에는 기본 매칭 방법과, 수평반전을 고려한 서술자 변형 후 매칭을 하는 방법 총 두 가지의 방법을 수행한다.

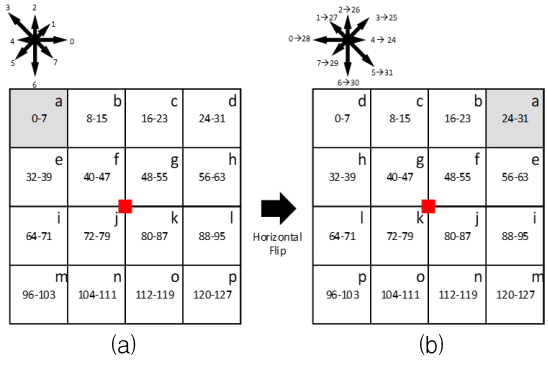

그림 2. (a) SIFT 서술자 구성, (b) 수평반전에 따른 서술자의 변화

3. 실험 및 결과

3.1. 실험환경 및 파라미터

본 실험에서는 1,280x720 비디오에서 10프레임마다 프레임을 추출하여 그림 3과 같이 원본 영상, 원본과 유사한 영상, 장면 전환된 영상을 구분하였다. 변형이 가해지지 않은 비디오에서 매칭률을 계산하여 장면 전환을 확인한다. 또한 비디오에 해상도 변형(0.5배, 2배), 압축(H.264, H.265), 반전(수평), 자막 삽입을 각각 수행하고, 특징을 추출하여 원본 특징과의 매칭률을 계산하고, 임계값에 따라 장면 전환 여부를 판단한다. 배경이 고정된 상태로 객체가 움직이거나 카메라의 이동에 따라 시점이 조금 변화하는 경우 등은 장면전환이라고 볼 수 없기 때문에 중복 추출되는 특징이 존재하여 실험을 통해 매칭률의 임계값을 25%로 정의하였다.

(a)

(b)

(c)

(d)

(e)

그림 3. 실험 영상 (a)원본 영상, (b),(c)원본과 유사한 영상, (d),(e)장면 전환된 영상

3.2. 원본에서의 장면 전환 검출

원본 비디오에 대해 그림 3(a) 영상에 대한 매칭률은 표 1과 같다. 카메라 위치가 이동한 (b)와는 약 65.5%의 높은 매칭률을 보였고, 고정되는 부분과 변하는 부분이 동시에 존재하는 (c)에서도 임계값 이상의 값을 가지는 것을 확인할 수 있다. 하지만, (d)와 (e)의 경우 임계값보다 낮은 매칭률을 가지기 때문에 장면 전환으로 검출한다.

표 1. 원본 비디오에서 그림 3(a)와의 매칭률 비교

	Fig.3(b)	Fig.3(c)	Fig.3(d)	Fig.3(e)
Matching Rate (%)	65.507	28.858	3.920	2.790

3.3. 변형된 비디오에서의 장면 전환 검출

해상도 변환, 비디오 압축, 수평반전, 자막 삽입 변형이 가해진 비디오에 대한 그림3(a) 프레임의 매칭률은 표 2와 같다. 각각의 변형이 가해짐에 따라 새로 생기는 특징과 없어지는 특징이 발생하기 때문에 매칭률은 달라지지만 표 1과 비교하였을 때 매칭률이 크게 차이나지 않고 유지되는 것을 확인할 수 있고 임계값에 따라 모든 변형에서 (b),(c)는 유사 장면으로, (d),(e)는 장면 전환으로 검출한다.

표 2. 원본 비디오의 그림 3(a) 프레임과 변형된 비디오의 그림 3(a)~(e) 프레임간의 매칭률 비교 (단위 : %)

Transformed	Fig.3(a)	Fig.3(b)	Fig.3(c)	Fig.3(d)	Fig.3(e)
Resolution x0.5	85.89	61.69	27.63	4.51	2.79
Resolution x2.0	94.02	65.75	30.33	4.16	2.40
Compression H.264	89.81	64.48	29.40	4.56	2.94
Compression H.265	71.04	54.58	27.65	4.51	2.89
Flip Horizontal	100	65.51	28.86	3.92	2.79
Caption Font size 20	97.94	65.70	29.54	4.46	2.89
Average	89.78	62.95	28.90	4.35	2.78

4. 결론

본 논문에서는 해상도를 고정시키는 전처리와 SIFT를 통한 특징 추출, 서술자를 변화시켜 매칭하는 방법을 통해 장면 전환을 검출하는 알고리즘을 제시하였다. 또한, 유통 과정에서 많이 사용되는 변형들을 원본 비디오에 가한 후에도 원본의 특징만으로 변형 비디오의 장면 전환을 검출하였다. 따라서 비디오의 변형에 강인하기 때문에 저작권 보호, 비디오 압축 등의 높은 활용을 기대한다.

감사의 글

이 논문은 2016 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2016R1D1A1B03930691)

참고문헌

[1] Y. M. Eom, S. I. Park, and C. W. Chung, “An analysis of Scene Change Detection using HEVC coding additional information,” Journal of Broadcast Engineering, Vol. 20, No. 6, pp. 871-879, Nov. 2015.

[2] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, Jan. 2004.