

효과적인 샷 경계 검출을 위한 플래시라이트 처리 기법에 대한 비교 및 분석

주현진[○] 남종호

서강대학교 컴퓨터공학과

d20150455@sogang.ac.kr, jhnang@sogang.ac.kr

Survey on Anti-flashlight Techniques for Effective Shot Boundary Detection

Hyunjin Joo[○] Jongho Nang

Department of Computer Science and Engineering, Sogang University

요 약

효과적인 영상의 샷 경계 검출은 Semantic Video Analysis의 첫 단계로 이를 위한 많은 알고리즘들이 연구되어 왔다. 동일한 샷 내에 삽입된 Flashlight은 연속된 프레임 사이의 유사성을 떨어뜨려 샷 경계로 잘못 검출될 수 있으며, 이 문제를 해결하기 위한 다양한 기법들도 제시되었다. 본 논문에서는 다양한 샷 경계 검출 모델에서 Flashlight을 검출하거나 그로 인한 영향을 제거하기 위해 어떤 방법들을 사용했는지 샷 경계 검출의 각 단계에 맞춰 비교 분석해본다. 또한 샷 경계 검출 외에 Flashlight 검출이 응용될 수 있는 분야에 대해서도 간단히 살펴본다.

1. 서 론

영상의 샷 경계 검출은 효율적인 멀티미디어 데이터베이스 구축을 통한 Automatic Indexing, Browsing 및 Retrieval의 첫 번째 단계로서, 여기서 샷은 편집 없이 한 번에 촬영된 일련의 프레임들을 말한다[1]. 결국 샷 경계 검출이란 인접한 두 샷 또는 장면의 전환지점을 찾는 것으로, 비디오는 시각적인 관점에서 2차원 이미지와 그 이미지의 시간 축에 따른 변화로 설명할 수 있기 때문에[1], 대부분의 검출 알고리즘은 프레임 간의 시간적 특성을 이용하여 샷 경계를 찾게 된다. 하나의 샷에 속한 프레임들은 시간 축에 대해서도 높은 유사성을 보이므로, 인접한 프레임들이 서로 얼마나 유사한 지가 일반적인 판단의 기준이 된다.

그 동안 영상의 샷 경계 검출을 위해 많은 연구들이 진행되었으며, 샷 간의 경계가 비교적 명확한 Cut뿐만아니라 Fade-in/Fade-out, Dissolve, Wipe 등과 같이 점진적으로 변화하는 경계를 찾기 위한 알고리즘도 다양하게 제시되었다. 여기서 몇 가지 공통적으로 발견된 어려움이 있는데, 급격한 밝기 변화나 커다란 Object 및 카메라의 움직임 등으로 인해 샷 경계가 잘못 검출되는 것이 그 중 하나이다.¹

본 논문에서는 이 중에서도 통상 Flashlight으로 표현되는 급격한 밝기 변화를 여러 알고리즘에서 어떻게 처리하는 지에 대해 얘기해 보려 한다. Flashlight을 Cut으로 오인하는

경우는 샷 경계 검출에서 가장 흔하게 볼 수 있는 오류의 하나로, 대부분 Luminance 또는 이를 기본 요소로 하는 Color를 기반으로 Visual Content를 표현하기 때문인데, Visual Content의 변화를 가장 효과적으로 특징지을 수 있는 Feature 또한 Luminance이기도 하다[1].

본 논문의 구성은 다음과 같다. 2장에서는 일반적인 샷 경계 검출의 각 단계를 Flashlight에 초점을 맞춰 살펴본다. 3장에서는 알고리즘 별 종합적인 비교 및 분석을 통해 한계점을 알아보고, 마지막으로 4장에서 결론과 함께 샷 경계 검출 외에 Flashlight 검출이 응용될 수 있는 분야에 대해서 간단히 언급한다.

2. 샷 경계 검출의 단계별 분석

샷 경계 검출 과정을 단계적으로 살펴보면 크게 다음과 같다. 먼저 각각의 이미지를 잘 표현할 수 있는 특징(Feature)을 선택하고(Representation of Visual Content), 그 특징을 이용하여 장면의 전환을 수학적으로 모델링한다(Representation of Shot Transitions). 그 후 각 모델에 맞는 방법으로 프레임 간의 유사도를 측정하여 전환지점을 찾는 것이다(Similarity Measure and Decision)[1].

2.1. Representation of Visual Content (Features Used)

샷 경계 검출을 위해서는 영상의 이미지 정보를 담은 수 있는 Feature가 필요하다. 좋은 Feature란 서로 다른 두 샷의 Visual Content를 구분할 수 있을 만큼 Sensitive하면서도 여러 다양한 Disturbance에 영향을 받지 않을 정도로 충분히

본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발 사업의 연구결과로 수행되었음 [R0126-15-1112, 퍼스널 미디어가 연결 공유 결합하여 재구성 가능케 하는 복합 모달리티 기반 미디어 응용 프레임워크 개발].

Invariant한 Feature이다. 특히 Flashlight은 급격한 Luminance 또는 Chrominance의 변화가 주요 특징이기 때문에 이를 포함하는 Feature를 이용하여 검출할 수도 있다.

2.2. Representation of Shot Transitions

대부분의 알고리즘은 선택한 Feature의 Temporal Discontinuity로 장면의 전환을 모델링 하게 된다.

Flashlight을 고려한 샷 경계 검출 모델의 경우, 1)2차 미분이나 Window를 통한 Smoothing 기법 등으로 급격한 Luminance의 변화에 의한 영향을 감쇠시킨 후 샷 경계 검출을 수행하거나, 2)별도로 Flashlight을 검출한 후 샷 경계 후보에서 걸러내는 방법을 사용한다. Flashlight을 검출하는 방법은 알고리즘의 초점을 어디에 두었는가에 따라 크게 ①짧은 구간 동안의 High Luminance로 인해 생기는 특징을 찾는 방법과 ②High Luminance 전후 프레임 사이의 높은 연관성을 이용하는 방법으로 나눌 수 있다.

2.2.1. Similarity Measures

영상의 Transition Model을 만들고 Continuity Signal을 구성하기 위해서는 프레임 간의 연관성 즉 유사도를 측정해야 한다. 수학적으로 두 Feature가 얼마나 다른지를 나타내는 방법 중 잘 알려진 것으로는 Euclidean Distance(L²

Distance)와 L¹ Distance가 있다.

두 Histogram 값 중 최소값을 취하는 Histogram Intersection도 많이 쓰이는 방법이며, 확률분포 사이의 거리를 표현하기 위한 Bhattacharyya Distance나 Chi-square Test 등 다양한 방법이 존재한다.

Edge나 Object 등을 매칭하여 유사도를 측정할 경우에는 Matching Ratio가 많이 쓰이며, 특정 Correlation 계산법을 이용하거나 알고리즘에 따라 적절한 Kernel Correlation을 수행하기도 한다.

2.3. Decision

Shot Transition Model에 따라 측정된 유사도를 기반으로 샷 경계를 판단하기 위해 많은 알고리즘에서 Threshold를 사용한다. 특히 Flashlight은 일반적으로 짧은 시간 동안 지속된다는 특징이 있기 때문에, 최대 프레임 수에 제한을 둘 수 있다. 그러나 일반적으로 Threshold는 영상의 특성에 따라 다르고 그 값에 따라 전체 성능에 차이가 나기도 한다. 이를 보완하기 위해 여러 개의 Threshold를 사용하거나 Local Window 기반의 분석을 통한 Adaptive Threshold를 적용하는 기법이 소개되었다.

또는 샷 경계 검출을 일종의 장면전환 패턴에 대한 패턴 인식 문제로 보는 Machine Learning 기법을 사용할 수도

표 1. Flashlight 검출 알고리즘의 비교

1st Author	Heng	Qian [2]	Yeo	Pei	Truong	W. Zhang
Publication	2003 EURASIP SPIC	2006 IEEE TCSVT	1995 IEEE TCSVT	1999 IEEE TMM	2001 IEEE ICME	2011 IEEE CSCWD
Detection Scope	Flashlight	Flashlight	Flashlight	Cut, Flashlight, GT, Caption	Flashlight	Cut, Flashlight, Fade, Other short/ long GTs
Main Feature	Edge pixel (w/ Direction)	Grayscale histogram of DC images	Pixel of DC images	Macroblock	Avg. luminance and Pixel	Gray histogram
Shot Transition Model	Pick out Flashlight from Abrupt changes detected based on MCB ²⁾	AHD (Accumulated Histogram Difference)	Pixel differencing	Pattern of MB types	High luminance increase → Constant high luminance → Decrease in luminance	Finite state machine
Similarity Measure	# of matched Edge pixels btw 2 abruptly changed frames	L1 distance	L1 distance	# of corresponding MBs	L1 distance	Histogram difference
Decision Method	Threshold	Classification (based on the behavior of the sign of AHD)	Threshold	Thresholds	Thresholds	Thresholds
How to deal w/ Flashlights	상대적으로 Flashlight에 의한 변화가 적은 Edge를 비교	AHD와 Energy variation의 부호 특성에 따라 Flashlight의 Start/ End frame 판별	Two consecutive sharp peaks within a sliding window	Flashlight에 의해 B-frame에서 Intra-coded MB의 수가 급격히 증가하는 특성 이용	Flashlight에 대한 상세한 모델링으로 Start/ End 및 Peak frame 검출	Window 내의 Flashlight 전/ 후 Frame들의 연관성 이용
Pros	Non-ordinary flashlight 검출	Insensitive to # of histogram bins or bit rates	Possible tradeoff for btw speed (Avg. luminance) and accuracy (Pixel)	i) Frame level detection ii) Insensitive to bit rates iii) Modifiable for different GOP structures	Flashlight에 대한 상세한 모델링	직관적 알고리즘
Cons	Pixel level tracking could be expensive	Shot 경계에 있는 Unpaired flashlight은 검출 못함	같은 Sliding window 내에 여러 개의 Flashlight이 있을 경우 검출 못함		Shot 경계에 있는 Unpaired flashlight은 검출 못함	①Quick object/ Camera motion, ②Fire, Wind, Sea wave 등에 의한 False detection

있다. 대표적으로 K-means, KNN, SVM, Neural Network, Decision Tree 등이 있다. Qian 등은 사전에 미리 파악된 Accumulated Histogram Difference(AHD)의 특성을 이용하여 학습 알고리즘 없이 Flashlight을 분류하기도 하였다[2].

3. 분석 및 실험

이번 장에서는 앞서 언급한 여러 알고리즘을 종합적으로 분석하고, 그 중 몇 가지에 대해 실험을 통한 검증을 해보려 한다.

본 논문에서 참조된 많은 알고리즘들이 이미지의 Intensity를 이용하여 영상의 Visual Content 및 장면의 변화를 표현하였으며, L1 Distance로 유사도를 측정하여 적절히 선택된 Threshold를 기준으로 Flashlight이나 샷 경계를 검출하였다. 지난 2001년부터 2007년까지의 TRECVID Activity를 살펴보면, 7년간 좋은 성능을 보였던 알고리즘들의 대부분이 Color histogram과 Machine Learning 기법을 기반으로 한 것을 알 수 있다.

3.1. 실험 결과

다음은 Accumulated Histogram Difference의 특성을 이용한 Qian 등의 모델[2]과, Gray Histogram의 Flashlight 전후 특성을 이용한 W. Zhang 등의 모델[3]을 임의의 뉴스 영상으로 실험한 결과이다.

두 알고리즘 모두 콘텐츠에 적절한 Global Threshold를 통해 Flashlight을 검출하게 되는데, 적절한 Global Threshold는 Feature의 분포와 시행착오를 통해 임의로 선택할 수 밖에 없다는 단점이 있다.

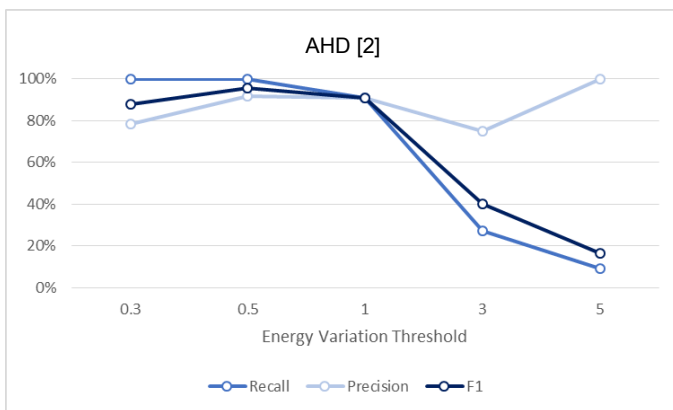


그림 1. Threshold에 따른 검출 성능

특히 W. Zhang의 알고리즘은 Threshold에 민감하여 적절한 Threshold를 찾지 못하면 극단적으로 Recall이 0이 되는 경우가 많았다. 이 문제는 Adaptive Threshold를 적용하여 어느 정도 보완할 수 있었다.

3.2. 한 계

기본적으로 Threshold에 의존한 Decision은 Threshold 값에 따라 성능이 좌우될 수 있는 한계가 있으며, Window나

Adaptive Threshold를 사용한다해도 Window Size, Scaling Factor 등의 파라미터를 설정해 주는 문제가 남아있다.

또한 Histogram의 차이를 이용한 검출은 빠르게 움직이는 Object나 물 또는 불의 미세한 움직임, 그래픽이나 캡션 요소에 영향을 받았다. 이로 인한 False Alarm은 샷 경계 후보에서 Flashlight을 제거하는 목적을 고려했을 때 큰 영향이 없을 수 있으나, Flashlight 검출기를 사용함으로써 짧은 샷(5 프레임 이하)을 Flashlight으로 오인하여 배제하는 경우는 문제가 될 수 있다.

표 2. 실험 결과

비디오	프레임	Flashlight	알고리즘	Recall	Precision
news_1	2,773	11	[2]	100%	92%
			[3]	64%	100%
news_2	2,651	22	[2]	95%	91%
			[3]	48%	91%
news_3	106,178	129	[2]	100%	74%
			[3]	12%	52%

4. 결론 및 응용

효과적인 샷 경계 검출은 Semantic Video Analysis의 첫 단계로 많은 연구에서 다양한 알고리즘을 제시해 왔다. 동일한 샷 내에 삽입된 Flashlight은 연속된 프레임 사이의 유사성을 떨어뜨려 샷 경계로 잘못 검출될 수 있으며, 이 문제를 해결하기 위한 여러 기법들을 본 논문에서 살펴보았다.

특히 Flashlight을 따로 검출하는 경우, Flashlight이 단순히 제거해야 할 Disturbance가 아니라 영상의 주요 장면 검출 등에 응용될 수도 있다.

참 고 문 헌

- [1] J. Yuan, H. Wang, L. Xiao, W. Zheng, J. Li, F. Lin, and B. Zhang, "A Formal Study of Shot Boundary Detection," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 17, No. 2, pp. 168~186, Feb. 2007.
- [2] X. Qian, G. Liu, and R. Su, "Effective Fades and Flashlight Detection Based on Accumulating Histogram Difference," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 16, No. 10, pp. 1245, 2006.
- [3] W. Zhang, S. Teng, and X. Fu, "State Transition-based for Cooperative Shot Boundary Detection," *Proc. IEEE Int. Conf. CSCWD*, pp. 125~130, Jun. 2011.