

MPEG 압축 비디오로부터 특징 정보의 직접 추출을 통한 빠른 장면 전환 검출 알고리즘

김 영민, 이 성환

고려대학교 컴퓨터학과/인공시각연구센터
E-mail:{ymkim, swlee}@image.korea.ac.kr

A Fast Scene Change Detection Algorithm using Direct Feature Extraction from MPEG Compressed Videos

Young-Min Kim and Seong-Whan Lee

Dept. of Computer Science and Engineering/Center for Artificial Vision Research,
Korea University

요 약

비디오 데이터의 효율적인 저장, 관리를 위해서는 장면 전환 검출을 통한 비디오 분할 기술에 대한 연구가 필요하므로, 최근 들어 압축 비디오상의 특징 정보를 직접 추출하여 장면 전환 검출에 사용하는 방법에 대한 연구가 많이 이루어지고 있다.

본 논문에서는 MPEG 압축 비디오 상의 에지 정보를 복호화 과정을 거치지 않고 직접 추출하여 장면 전환 검출에 사용하는 새로운 방법을 제안하였다. 이산 여현 변환(DCT)된 블록 내 AC 계수의 부호를 통해 에지의 모양을 알아내었으며, AC 계수 간의 상관 관계를 통해 에지의 방향과 세기를 측정하여 프레임을 정합하는 방법을 사용하였다. 실험 결과 사용한 특징 정보가 명도나 색상 변환에 무관하여 잘못 검출하는 비율이 현저히 적었으며, 영상을 완전 복호화한 후 에지를 구하여 장면 전환 검출을 하는 방법에 비해 약 5-6배 속도가 빠름을 확인할 수 있었다.

1. 서론

장면 전환 검출 기술은 효과적인 비디오 처리를 위한 필수 요소 기술로[1][2], 최근 들어 압축 상태의 데이터에서 직접 특징을 추출하여 장면 전환점을 찾는 기술에 대한 연구가 많이 이루어지고 있다.

장면 전환의 검출을 위해 기존에 사용된 명도나 색상 히스토그램 비교법은 명도나 색상의 변화가 큰 비디오 데이터에 대해서는 장면 전환점을 잘못 검출하는 비율이 매우 높다는 단점이 있었다. 사람은 사물을 구별할 때 주로 에지 정보를 사용하여 명도나 색상의 변화가 크더라도 사물을 오인식하지 않는다는 특성이 있으며[3], MPEG 비디오 데이터의 경우 영상의 복호화 과정에서 중앙 처리 장치 수행 시간의 최소 40% 정도를 역 이산 여현 변환(IDCT) 과정에 사용한다고 알려져 있다[4]. 따라서 IDCT 과정을 거치지 않고도 에지 정보를 추출해 낼 수 있다면, 매우 효율적이고 빠른 장면 전환 검출이 가능하다는 것을 알 수 있다.

본 논문에서는 MPEG 압축 비디오 상에서 IDCT 과정을 거치지 않고 에지 정보를 압축 상태에서 직접 추출하여 프레임 정합에 사용하는, 보다 효율적이고 빠른 장면 전환 검출 방법을 제안하였다.

2. 관련연구

Zabih 등[5]은 비압축 비디오를 대상으로 에지 영상의 진입 화소와 퇴장 화소값을 장면 전환 검출에 사용하였는데,

이 방법은 명도나 색상 변화에 민감하지 않으므로 상대적으로 우수한 성능을 보이나, 영상을 완전 복원해야 하고 프레임 내의 에지 영상을 추출해야 하므로 시간과 비용이 많이 든다는 단점이 있다.

Yeo 와 Liu[6]가 제안한 압축 비디오의 DC 영상을 이용한 방법은 전체 영상을 모두 복호화하지 않아도 되므로 속도가 빠르고 성능도 우수하다고 알려져 있다. 그러나 이 방법은 명도의 변화에 민감하여 장면 전환점을 잘못 검출할 수 있다는 단점이 있다.

3. 제안한 장면 전환 검출 방법

본 논문에서 제안한 장면 전환 검출 방법은 크게 에지 정보 직접 추출 단계와 프레임 간 정합을 통한 장면 전환 검출 단계 나눌 수 있다.

3.1 에지 정보 직접 추출 단계

비디오 압축 국제 표준인 MPEG은 블록 단위 DCT를 수행하는데, DCT를 하게 되면 영상이 DC 계수와 AC 계수로 나뉘게 되며, DC 계수는 블록의 평균값을 나타내고 각 AC 계수들은 수직이나 수평 성분의 값에 따라 고유한 의미를 지니게 된다.

Shen 과 Sethi[7]는 이러한 AC 계수의 특징을 사용하여 에지 정보를 추출해 내는데 성공하였다. 본 논문에서는 AC 계수가 에지의 수직과 수평 성분을 나타내며, 에지의 모양은 AC 계수의 부호에 의해 결정된다는 특성을 이용

하여 Shen 과 Sethi의 에지 정보 추출 방법을 수정, 에지를 추출하였다. 그림 1은 본 논문에서 제안한 에지 추출 방법을 나타내며, F_{uv} 는 세로로 u번째, 가로로 v번째 AC계수를 나타낸다.

수직우세 $ F_{01} > F_{10} $	$F_{01} > 0$	$F_{10} > 0$	$F_{01} < 0$	$F_{10} > 0$	$F_{01} < 0$
	$F_{10} < 0$			$F_{10} < 0$	
수평우세 $ F_{01} < F_{10} $	$F_{01} > 0$		$F_{10} < 0$	$F_{01} > 0$	
	$F_{01} < 0$		$F_{10} < 0$	$F_{01} < 0$	
대각선 $ F_{01} = F_{10} $ $F_{10} = F_{01} \neq 0$	$F_{01} > 0, F_{10} > 0$		$F_{01} < 0, F_{10} > 0$		
	$F_{01} > 0, F_{10} < 0$		$F_{01} < 0, F_{10} < 0$		
수평 $F_{01} = 0, F_{10} \neq 0$			수직 $F_{10} = 0, F_{01} \neq 0$		

그림 1. 본 논문에서 제안한 에지 추출 방법

3.1.1 에지의 방향 성분

에지의 방향은 에지의 모양을 결정하는 중요한 성분 중 하나로서, AC 계수의 수직 우세 성분과 수평 우세 성분의 비로 결정되어 진다. 본 논문에서는 아래의 식 1을 사용하여 에지의 방향을 결정하였다.

$$\tan(\theta) = \frac{\sum_{v=1}^7 F_{0v}}{\sum_{u=1}^7 F_{u0}} \quad (1)$$

3.1.2 에지의 세기 성분

에지의 세기는 수평 방향의 세기와 수직 방향의 세기로 나눌 수 있다. Shen 과 Sethi[7]는 AC 계수의 수평 우세 성분과 수직 우세 성분을 사용하여 에지의 세기를 구하였으나, 이 방법은 에지가 어느 한 쪽으로 기울게 되면 에지의 세기에는 변화가 없는데도 값이 작아지는 문제점이 생기게 된다.

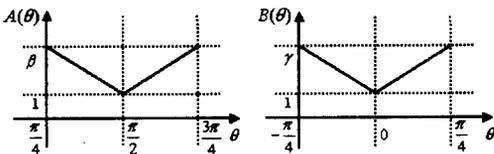


그림 2. 본 논문에서 정의된 함수 A(theta)와 함수 B(theta)

본 논문에서는 이러한 문제점을 해결하기 위하여 함수 A(theta)와 B(theta)를 정의하여 식 2, 3과 같이 에지의 세기를 구하였다. (단, $0 \leq \xi \leq 1$)

$$V_{Strength} = ((1 - \xi)F_{01} + \xi F_{02}) * A(\theta) \quad (2)$$

$$H_{Strength} = ((1 - \xi)F_{10} + \xi F_{20}) * B(\theta) \quad (3)$$

그림 2는 본 논문에서 정의한 함수 A(theta)와 B(theta)를 나타낸다.

3.2 프레임 간 정합 단계

본 논문에서 제안한 방법을 사용하여 추출된 에지 정보를 사용하면 명도 변화나 색상 변화에 무관한 효과적인 프레임 간 정합이 가능하다. 그림 3은 프레임 정합을 위한 전체적인 구조도를 나타낸다.

본 논문에서는 프레임 정합을 위해 특징 정보로 추출된 에지의 방향 성분과 세기 성분을 사용한다.

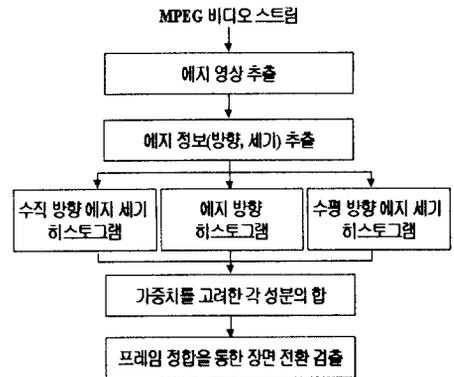


그림 3. 제안한 프레임 정합의 전체 구조도

3.2.1 에지 방향 성분 비교

블록 단위로 추출된 에지는 고유의 방향을 지니게 된다. 따라서 프레임 내 에지의 방향 성분을 사용하여 프레임 정합을 할 수 있다. 본 논문에서는 식 4를 사용하여 프레임 정합을 위한 에지의 방향 성분값의 차(DOAH)를 구했다. 식 4에서 f_n 는 n번째 프레임, Q_A 히스토그램 단계, $AH_x(i)$ 는 프레임 x의 i번째 각도 히스토그램을 나타낸다.

$$DOAH(f_n, f_{n+1}) = \sum_{i=0}^{Q_A-1} |AH_{f_n}(i) - AH_{f_{n+1}}(i)| \quad (4)$$

3.2.2 에지 세기 성분 비교

에지를 통해 추출할 수 있는 정보는 방향 성분 외에 세기 성분이 있다. 방향 성분만 사용하면 프레임 내 사물의 움직임이 빠르거나 회전하는 경우 장면 전환점을 잘못 검출하는 경우가 있었다. 따라서 이를 보완하기 위하여 식 5-12를 사용하여 에지의 수직/수평 방향 세기 성분의 히스토그램을 구하였다. 식 5-12에서 $M * N$ 은 영상의 크기, W, K 는 각각 히스토그램의 단계, $ST(i, j)$ 는 (i, j)번째 블록의 에지 세기를 나타낸다. 또한, ∇_R 와 ∇_C 은 각각 임계치를 나타낸다.

$$Y(h) = \sum_{j=h * \frac{M}{K}}^{\{(h+1) * \frac{M}{K}\} - 1} \sum_{i=0}^{M-1} ST(i, j) \quad (5)$$

$$D_Y = |Y_{f_n}(h) - Y_{f_{n+1}}(h)| \quad (6)$$

$$\varphi_{f_n, f_{n+1}}^Y(h) = \begin{cases} 1 & \text{if } D_Y < \nabla_R \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$DOSH_Y = (f_n, f_{n+1}) = \frac{\sum_{h=0}^{K-1} \varphi_{f_n, f_{n+1}}^Y(h)}{K} \quad (8)$$

$$X(h) = \sum_{i=h * \frac{M}{K}}^{\{(h+1) * \frac{M}{K}\} - 1} \sum_{j=0}^{N-1} ST(i, j) \quad (9)$$

$$D_X = |X_{f_n}(h) - X_{f_{n+1}}(h)| \quad (10)$$

$$\varphi_{f_n, f_{n+1}}^X(h) = \begin{cases} 1 & \text{if } D_X < \nabla_C \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

$$DOSH_X = (f_n, f_{n+1}) = \frac{\sum_{h=0}^{W-1} \varphi_{f_n, f_{n+1}}^X(h)}{W} \quad (12)$$

3.2.3 에지 방향과 세기 성분의 차이를 이용한 정합

위에서 구한 에지의 방향과 세기 성분값의 차를 이용하여 식 14와 같이 프레임 정합을 수행하였다. (단, $0 \leq \kappa \leq 1$)

$$DOSH = DOSH_X + DOSH_Y \quad (13)$$

$$DOF(f_n, f_{n+1}) = (1 - \kappa)DOAH + \kappa DOSH \quad (14)$$

4. 실험 및 성능 분석

4.1 실험 환경

본 논문에서 제안한 에지 정보를 이용한 장면 전환 검출 방법의 성능을 평가하기 위해 Windows NT 환경에서 Pentium II 350MHz PC를 사용하여 실험하였다. 실험에 사용된 실험 데이터의 종류와 크기는 표 1과 같다.

표 1. 실험 비디오 데이터

종류	분량(mm:ss)	용량	N_{total}	N_{cut}
뉴스	06:18	62MB	9,511	106
다큐멘터리	17:13	184MB	35,775	139
뮤직 비디오	16:49	164MB	30,496	113

표 1에서 N_{total} 은 총 프레임의 수, N_{cut} 은 컷의 개수를 나타낸다.

4.2 실험 결과 및 성능 분석

표 2. 장면 전환 검출 방법에 따른 성능 비교 ($N_C : N_{FN} : N_{FP}$)

종류	뉴스	다큐멘터리	뮤직비디오
DC	(98:8:4)	(135:4:9)	(110:3:13)
ED	(101:5:1)	(137:2:0)	(110:3:4)
PM	(99:7:3)	(135:4:1)	(108:5:5)

표 2는 표 1을 대상으로 Zabih 등[5]이 제안한 비압축 환경에서의 에지 정보를 장면 전환 검출에 사용하는 방법(ED), 압축 환경에서의 장면 전환 검출 성능이 뛰어나고 알려진 Yeo 와 Liu[6]가 제안한 DC 영상의 명도차를 이용한 방법(DC)과 본 논문에서 제안한 방법을 성능 평가한 것이다. 표 2에서 N_C 는 정확히 검출된 장면 전환점의 수를 나타내고, N_{FN} 는 놓친 컷의 수, N_{FP} 잘못 검출한 수를 나타낸다. 표 2를 보면 DC 영상을 이용한 방법은 명도 변화가 심한 뮤직 비디오나 다큐멘터리 데이터에서 잘못된 장면 전환점의 수가 많아 ED와 PM이 상대적으로 우수함을 알 수 있다.

표 3. 장면 전환 검출 방법에 따른 속도 비교

종류	뉴스	다큐멘터리	뮤직 비디오
DC	11.4	11.2	9.8
ED	2.4	2.0	2.1
PM	11.6	10.7	10.2

표 3은 위에서 사용한 방법들의 속도를 비교한 결과이다. 표 3에서 보면 DC와 PM은 압축 상태에서 바로 특징 정보를 추출하기 때문에 ED보다 약 5-6배 빠른 것을 알 수 있다. 위 두 실험에서 볼 수 있듯이 본 논문에서 제안한 방법은 정확할 뿐 아니라 속도도 매우 빨라 다른 방법들보다 우수함을 확인할 수 있었다.

5. 결론 및 향후 연구 방향

본 논문에서는 MPEG 압축 비디오 상의 에지 정보를 복호화 과정을 거치지 않고 직접 추출하여 장면 전환 검출에 사용하는 방법을 제안하였다. DCT 블록 내 AC 계수의 부호를 통해 에지의 모양을 알아내었으며, 계수의 상관관계를 통해 에지의 방향이나 세기를 측정하여 프레임을 정합하는 방법을 사용하였는데, 에지 방향과 세기 히스토그램을 사용하여 프레임 내 사물의 빠른 움직임이나 회전 등의 변화에도 무관함을 확인하였다. 실험 결과 사용한 특징 정보가 명도나 색상 변환에 무관하여 비교적 성능이 우수한 Yeo와 Liu의 DC 영상을 사용한 방법보다 잘못 검출하는 비율이 현저히 적었으며, 영상을 완전 복호화한 후 에지를 구하여 장면 전환 검출을 하는 Zabih 등의 방법에 비해 약 5-6배 속도가 빠름을 확인할 수 있었다.

압축 비디오에서 직접 추출 가능한 특징 정보인 움직임 벡터 정보를 사용하면 프레임 내 전체적인 움직임 정보를 알아낼 수 있다. 이러한 움직임 정보는 장면 전환 검출시, 단순한 프레임간 정합에 의해 발생하는 오류를 보완하는데 유용하게 사용될 수 있다. 또한, 제안된 장면 전환 검출 방법은 페인드 인이나 아웃, 디졸브, 와이프 등 점진적 장면 전환에 대해 고려하지 않았으나, 이 부분을 보완한다면 보다 효과적인 장면전환 검출이 가능할 것이다.

감사의 글

본 연구는 부분적으로 과학기술부 창의적연구진흥사업의 연구비 지원을 받았음.

참고 문헌

- [1] 김 영민, 이 성환, "효율적인 비디오 검색 및 브라우징을 위한 대표 프레임 군집화," 한국정보과학회 가을 학술발표논문집, 수원, 제 25권 제 2호, 1998년 10월, pp. 553-555.
- [2] M.-S. Lee and S.-W. Lee, "Automatic Video Parsing Using Shot Boundary Detection and Camera Operation Analysis," *Pattern Recognition*, Vol. 33, 2000. (To appear)
- [3] T. Kato et al., "A Sketch Retrieval Method for Full Color Image Database," *Proc. of 11th Int. Conf. on Pattern Recognition*, The Hague, pp. 530-533, 1992.
- [4] E. Feig and S. Winograd, "Fast Algorithms for the Discrete Cosine Transform," *IEEE Trans. on Signal Processing*, Vol. 40, No. 9, pp. 2174-2193, 1992.
- [5] R. Zabih et al., "A feature-based algorithm for detecting and classifying scene breaks," *Proc. of ACM Multimedia '95*, San Francisco, CA, pp. 189-200, 1995.
- [6] B. L. Yeo and B. Liu, "Rapid Scene Analysis on Compressed Videos," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 5, No. 6, pp. 533-544, 1995.
- [7] B. Shen and I. K. Sethi, "Direct feature extraction from compressed images," *Proc. of SPIE*, San Jose, CA, Vol. 2670, pp. 404-414, 1996.