

지역적 χ^2 를 이용한 장면전환검출 기법

신성운*, 백성은*, 표성배**, 이양원*

Scene Change Detection Using Local χ^2

Seong-Yoon Shin*, Seong-Eun Baik*, Seong-Bae Pyo**, Yang-Won Rhee*

요 약

본 논문에서는 비디오의 분할을 위하여 먼저 기존에 제안되었던 차이 값 추출방법들의 단점들을 극복하고 장점을 최대한 활용할 수 있으며 급진적 장면전환부터 점진적 장면전환까지 모두 예측할 수 있는 강건하고 복합적인 차이 값 추출방법에 대해서 제안한다. 이 방법은 지역적 χ^2 -테스트로서 기존의 컬러 히스토그램과 χ^2 -테스트를 결합한 방법이다. 본 논문을 위하여 기존의 히스토그램 기반 알고리즘과 비교하여 좋은 성능을 보여주는 χ^2 -테스트를 변형하였고, 컬러 값의 세분화 작업에 따른 검출효과를 높이기 위하여 명암도 등급에 따른 가중치를 적용한 지역적 χ^2 -테스트를 이용하였다. 이 방법은 복잡하고 다양한 시세계의 영상 변화를 가장 일반적이고 표준화된 방법으로 분석하고 분할하며 표현할 수 있는 방법이다. 기존의 χ^2 -테스트와 제안된 지역적 χ^2 -테스트 방법의 비교는 실험을 통해 입증되었다.

▶ Keyword : 점진적 장면전환(Gradual Scene Change), 급진적 장면전환(Rapid Scene Change), 지역적 χ^2 -테스트(Local χ^2 -test), 컬러 히스토그램(Color Histogram)

1. 서 론

동영상은 일반적으로 카메라의 각도나 위치가 변하거나 의도적으로 편집에 의해 다른 장면이 삽입되지 않는다면 각 프레임들의 성격은 유사하다. 구조화된 동영상 데이터란 장면전환 지점을 기준으로 스트림 형태의 동영상 데이터를 고유의 크기와 특성을 갖는 다수개의 장면으로 나누어 분리된 데이터를 말한다. 장면전환검출이란 비디오 편집이나 카메라 작업 등에서 물리적/논리적으로 현재의 장면과는 다른 장면으로 전

환되는 지점을 검출하는 것을 말한다. 프레임이란 구조화된 동영상 데이터의 최소 단위로써, 이들 프레임들이 물리적/논리적(내용적)으로 특성에 따라 분리되어 내용기반 동영상 응용의 중요한 요소가 된다. 장면전환검출이란 비디오 편집이나 카메라 작업 등에서 물리적/논리적으로 현재의 장면과는 다른 장면으로 전환되는 지점을 검출하는 것을 말한다. 클러스터링이란 유사한 키 프레임을 갖는 장면들을 다시 하나의 논리적 단위로 묶는 것을 말하며 이를 위해 다수의 장면전이 그래프를 만들어 하나의 스토리를 구성하는데, 이 스토리의 묶음은 하나의 비디오를 나타낸다[1].

• 제1저자 : 신성운

* 부산대학교 컴퓨터정보과학과, **인덕대학 컴퓨터소프트웨어과

가본 논문에서는 비디오의 분할을 위하여 먼저 기준에 제안되었던 차이 값 추출방법들의 단점들을 극복하고 장점을 최대한 활용할 수 있으며 급진적 장면전환부터 점진적 장면전환까지 모두 예측할 수 있는 강건하고 복합적인 차이 값 추출방법에 대해서 제안한다. 장면분할을 어렵게 만드는 여러 가지 요인은 특징 값을 고려할 수 없는 영상 변화에 의해 발생하는 미 검출 현상, 프레임간의 특징 비교를 위해 사용되는 적절한 임계값 설정의 문제, 영상의 느린 변화에 의한 점진적인 전환 추출의 문제 등으로 요약될 수 있다. 본 논문에서는 이러한 문제점들을 줄이고 복잡하고 다양한 실세계의 영상변화를 가장 일반적이고 표준화된 방법으로 분석하고 분할하며 표현할 수 있는 기법을 제안하고자 한다. 본 논문에서는 지역적 컬러 χ^2 -테스트를 제한하는데 이는 기존의 χ^2 -테스트와 컬러 히스토그램을 결합한 방법이다.

2. 관련연구

장면전환을 검출하기 위한 가장 빠르고 간단한 방법은 어떤 임계값 이상으로 변하는 픽셀의 수(Pixel comparison)를 계산하여 비교하는 방법이다[2,3,4]. 그러나 이 방법은 작은 영역에서의 큰 변화나 큰 영역에서의 작은 변화를 구별할 수 없는 단점이 있다. 블록기반 비교 방법(Block-based comparison)은 카메라와 객체의 동작에 민감하지 않도록 지역적 특성을 강조한 방법이다. Kast 등[5]은 각 블록들에 대하여 평균과 분산정보를 이용한 유사 비율(likelihood ratio)을 측정하여 장면의 변화를 측정하였다. Shah 등[3]은 프레임을 12개의 겹치지 않는 블록으로 나눈 후에 명암 값을 이용하여 이전의 영상과 가장 잘 일치하는 블록을 찾는 방법을 제안하였다. 히스토그램 비교방법(Histogram comparison)은 장면 전환의 검출을 위하여 사용되는 가장 보편화된 방법이다. Tono 등[2,6,7]은 그레이-레벨의 히스토그램 비교를 통하여 임계 값을 기준으로 장면전환을 추출하는 가장 간단한 방법을 제안하였다. Ueda 등[8]은 장면의 경계를 검출하기 위해서 컬러 히스토그램의 변화 비율을 사용했고, Naga 등[7]은 그레이 레벨과 컬러 히스토그램을 기반으로 한 몇 개의 간단한 통계학적 비교를 수행하였다. Zhan 등[2]은 픽셀 차이, 통계 차이, 그리고 몇 개의 히스토그램 방법을 비교하여 히스토그램 방법이 정확성과 속도사이의 좋은 교환요소를 발견하였다. Naga 등[7]은 두 프레임 사이의 차이 값을 강조할 뿐만 아니라 카메라나 객체의 움직임을 강조

할 수 있는 χ^2 -test를 제안하였다. 그러나 χ^2 -test는 Tono 등[6]이 제안한 선형 히스토그램 비교방법보다 전체적인 성능이 더 좋지 않았으며, 계산량이 증가하는 단점을 갖는다. Garg 등[9]은 6개의 다른 컬러 좌표시스템을 이용하여 3개의 히스토그램 기반방법에 대하여 성능을 평가하였다. Zhan 등[2]은 두 개의 임계치를 이용하여 점진적 장면변화를 검출할 수 있는 twin-comparison 방법을 제안하였다. 또한 비디오 분할을 위한 자동 임계치 결정 알고리즘[10]과 축구 비디오 하이라이트 생성[11]에서도 χ^2 의 변형된 방법들을 이용하였다.

3. 차이값 추출

본 논문에서는 기존의 히스토그램 기반 알고리즘과 비교하여 좋은 성능을 보여주는 χ^2 -테스트를 변형하였고, 컬러 값의 세분화 작업에 따른 검출효과를 높이기 위하여 명암도 등급에 따른 가중치를 적용한 지역적 χ^2 -테스트를 이용하였다.

컬러 히스토그램 비교($d_{r,g,b}(f_i, f_j)$)는 인접한 두 프레임(f_i, f_j)의 각 컬러공간에 대한 히스토그램 비교를 통하여 계산되어지며 식(1)과 같이 정의하여 사용한다.

$$d_{r,g,b}(f_i, f_j) = \sum_{k=0}^{N-1} \left(|H_i^r(k) - H_j^r(k)| + |H_i^g(k) - H_j^g(k)| + |H_i^b(k) - H_j^b(k)| \right) \dots \dots \dots (1)$$

$H_i^r(k), H_i^g(k), H_i^b(k)$ 는 i 번째 프레임(f_i)에서의 각 컬러 공간(r, g, b)에 대한 빈(k)의 수(N)를 나타낸다. 이 방법은 카메라와 객체의 동작과 명암에 매우 민감하며 많은 데이터 유실을 초래하지만, 히스토그램 방법이 정확성과 속도 사이의 좋은 교환 요소를 발견했다.

식(1)로부터 각 컬러공간에 대한 명암도 등급변환을 위한 가중치를 적용하여 식(2)와 같이 재 정의하여 사용할 수 있다.

$$d_{wr,wg,wb}(f_i, f_j) = \sum_{k=0}^{N-1} \left(|H_i^r(k) - H_j^r(k)| * \alpha + |H_i^g(k) - H_j^g(k)| * \beta + |H_i^b(k) - H_j^b(k)| * \gamma \right) \dots \dots \dots (2)$$

α, β, γ 는 NTSC 표준에 따른 명암도 등급변환을 위한 상수를 나타내며 $\alpha=0.299, \beta=0.587, \gamma=0.114$ 로 정의하여 사용하였다.

두 프레임사이의 차이를 강조하기 위한 통계 분석방법 중에서 χ^2 -테스트 비교 ($d_{wx^2}(f_i, f_j)$)는 히스토그램의 비교변화를 통하여 장면전환을 검출할 수 있는 효율적인 방법이며, 식(3)과 같이 정의하여 사용한다.

$$d_{wx^2}(f_i, f_j) = \begin{cases} \sum_{k=0}^{N-1} \frac{(H_i(k) - H_j(k))^2}{\max(H_i(k), H_j(k))} & \text{if } (H_{i,j} \neq 0) \\ 0 & \text{otherwise} \end{cases} \dots\dots\dots (3)$$

이러한 방법은 잡음에 효율적이지만, 통계식의 복잡성 때문에 계산 속도가 매우 느리며, 많은 잘못된 결과들을 만들어내기도 한다.

히스토그램 기반 방법은 공간정보를 적용하지 않아 유사한 컬러분포를 갖는 다른 두 영상인 경우 같은 영상으로 간주하여 검출하는 문제점이 발생할 수 있다. 이러한 문제점은 프레임의 영역을 분할하여 지역적에 대한 히스토그램 분포를 비교하는 방법으로 해결할 수 있다.

식(4)에서는 영역의 분할에 따른 각 영역의 컬러 히스토그램 비교와 누적을 통한 차이 값 생성을 나타내고 있다.

$$d(f_i, f_j) = \sum_{bl=1}^m DR(f_j, f_i, bl)$$

$$DR(f_j, f_i, bl) = \sum_{k=1}^{N-1} |H_i(k, bl) - H_j(k, bl)| \dots\dots\dots (4)$$

$H_i(k, bl)$ 는 프레임(f_i)에 대한 블록(bl)의 k 번째 히스토그램의 분포를 나타내며, m 은 전체 블록의 수를 나타낸다.

위의 식(2)에서와 같이 각 컬러공간에 대한 가중치를 적용한 세분화된 히스토그램 비교와 식(3)의 통계방식을 이용한 차이 값의 확장 및 식(4)의 지역적 히스토그램에 의한 프레임의 공간정보에 대한 활용의 장점을 이용하여, 본 논문에서는 차이 값 추출에 대한 강건함과 신뢰성을 위하여 다음 식(5)와 같이 위의 식들을 결합한 지역적 χ^2 -테스트 차이 값 추출식을 다음과 같이 제안한다.

$$d(f_i, f_j) = \sum_{bl=1}^m d_{x^2}(f_i, f_j, bl)$$

$$d_{x^2}(f_i, f_j, bl) = \sum_{k=1}^{N-1} \left(\frac{(H_i^r(k) - H_j^r(k))^2}{\max(H_i^r(k), H_j^r(k))} \times \alpha \right. \\ \left. + \frac{(H_j^g(k) - H_i^g(k))^2}{\max(H_i^g(k), H_j^g(k))} \times \beta \right. \\ \left. + \frac{(H_i^b(k) - H_j^b(k))^2}{\max(H_i^b(k), H_j^b(k))} \times \gamma \right) \dots\dots\dots (5)$$

위의 식(5)에서 $H_{ri}(k), H_{gi}(k), H_{bi}(k)$ 는 i 번째 프레임(f_i)이 갖는 각 컬러 공간 (r, g, b)에 대한 히스토그램 분포를 나타낸다. N 은 빈(k)의 전체수, m 은 블록(bl)의 전체수를 나타낸다. 또한 본 논문에서도 $\alpha=0.299, \beta=0.587, \gamma=0.114$ 로 정의하여 사용하였다.

본 논문에서는 제안된 식(5)을 이용하여 연속된 프레임사이의 지역적 간 히스토그램 비교를 통하여 차이 값을 생성한다. 생성된 차이 값은 급진적 장면의 변화나 점진적 장면의 변화 모두에 적용할 수 있는 강건한 차이 값의 추출을 보여준다.

4. 실험

본 논문에서는 실험을 위하여 펜티엄4-1.2 GHz(512M) 환경에서 운영체제는 Windows XP를 사용하였으며 편의성 및 확장성을 위하여 Win API와 C++를 병행하여 사용하였고, 윈도우 멀티미디어를 다루기 위한 기본 요소인 DirectX SDK8.1을 이용하였다. 실험 비디오는 광고, 뉴스, 그리고 스포츠 비디오를 이용하였고 비디오 저장도구로는 Sigma TV Capture Board를 이용하였다. 비디오 저장 형식은 MPG(MPEG4-V1)와 AVI(MS-CRAM)을 이용하였고 초당 프레임 수는 30f/s를 320 X 240 사이즈로 이용하였다.

제안된 비디오 차이 값 추출식의 성능평가를 위해서 본 논문에서는 제안된 식(5)의 지역적 χ^2 -테스트 알고리즘과 장면전환 검출에 좋은 성능을 보여주는 전역적 χ^2 -테스트 알고리즘과 비교평가를 하였다.

<표 1>은 실험에서 사용된 비디오의 전체 차이 값들로부터 생성된 최대값, 최소값, 평균과 표준편차를 표로 나타내어 보았다. global은 [7]에서 제안된 전역적 χ^2 -테스트를 나타내며, local은 식(5)에서 제안된 지역적 χ^2 -테스트를 나타낸다.

표 1. 비디오의 전체 차이값들에 대한 비교

	최대 값		최소 값		평균		표준편차	
	glob al	local	glob al	local	glob al	local	glob al	local
광고	467.4	492.6	178.8	284.7	244.8	334.2	51.8	38.2
뉴스	454.9	490.6	161.6	262.2	210.1	308.1	34.6	25.1
스포츠	462.5	491.3	240.4	331.7	314.7	391.6	43.4	33.4

〈표 1〉에서 local의 표준편차가 global의 표준편차 보다 평균적으로 10정도 적은 것을 볼 수 있다. 이는 전체적으로 볼 때는 장면전환 검출에 약점으로 작용될 것이라고 판단되어질지 모른다. 하지만 표준편차가 적으며 〈그림 1〉, 〈그림 2〉, 그리고 〈그림 3〉에서 보여주는 것처럼 그래프가 장면이 전체적으로 차이값이 큰 프레임들은 높고 적은 프레임들은 낮게 나타나므로 차이값 추출에 효율적이다.

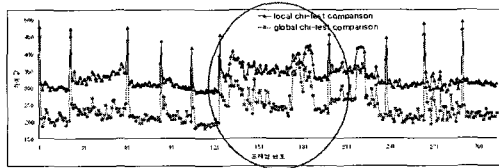
〈그림 1〉에서는 표 1의 실험에 사용되었던 광고 비디오에 대한 실험한 결과를 보여준다.

첫 번째 그래프는 전체 차이 값들에 대한 분포곡선 및 실험을 위하여 선택된 일부영역을 보여주고 있다. 실험을 위하여 선택된 영역은 비교적 넓게 분포되어 있는 점진적 장면의 변화를 갖는 프레임들과 급진적 변화를 갖는 일부 영역도 선택하였다. 이것은 선택된 영역들이 제안된 지역적 χ^2 -테스트에 의하여 시간적으로 밀접한 연관성을 갖는 차이 값들을 형성하고 있음을 보여주기 위한 것이다. 두 번째 그래프는 평균과 표준편차를 계산하여 기존 방법과 제안된 방법의 정규분포 곡선에 대한 차이를 보여준다. 곡선을 통하여 알 수 있는 것처럼 제안된 방법은 기존의 방법에 비하여 보다 좁은 범위의 차이 값들을 형성하며 이것은 유사한 차이 값들이 시간적으로 연속된 범위 안에서 밀접하게 연결되어 나타나는 것을 세 번째 세부 그래프를 통해서 확인할 수 있다.

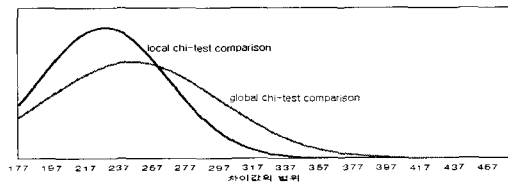
제안된 차이 값 추출식은 급진적 장면전환은 물론 카메라나 객체의 동작에 의한 점진적인 장면변환의 추출도 목표로 하고 있다. 특히 점진적인 장면의 변화와 같은 경우에는 유사한 차이 값들을 구성하는 프레임들이 연속적으로 나타남으로서 실제 군집화 과정을 통하여 점진적 장면구간만을 결정하는 것은 어려운 문제이다.

〈그림 1〉에서도 기존의 전역 방법은 유사한 장면의 변화임에도 연속되는 프레임의 차이 값의 편차가 커지는 단

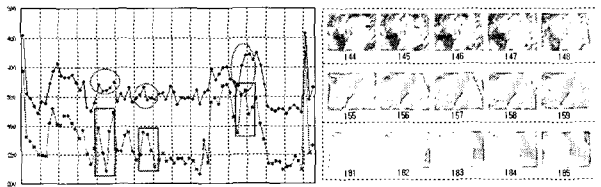
점이 있으나, 제안된 지역적 방법은 이러한 문제점을 어느 정도 제거하고 있다.



1) 전체 차이 값들에 대한 분포곡선 및 선택된 영역



2) 전역적 히스토그램과 지역적 히스토그램의 정규분포 곡선



3) 선택된 영역의 세부 차이 값 분포 곡선 및 프레임 분포
그림 1. 전역적 히스토그램과 지역적 히스토그램의 비교(광고)

따라서 차이 값들에 대한 분포를 보다 완만하게 형성하여 군집화를 통한 점진적인 장면의 변화추출에 보다 높은 효율성을 가질 수 있다.

다음 〈그림 2〉는 장면전환검출을 수행하는 화면으로서 위쪽에는 추출된 키 프레임들이 좌측에는 히스토그램이 우측에는 구동되는 비디오가 온다.

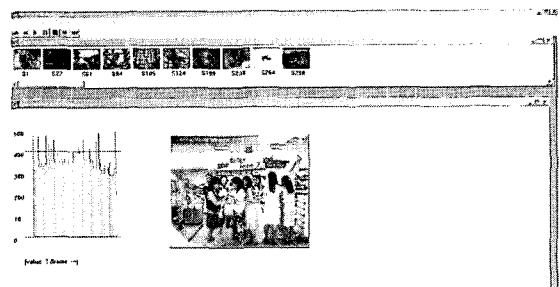


그림 4 장면전환검출 화면

실험에 사용된 임계치와 실험 결과는 다음 <표 2>와 같다. <표 2>에서 총 프레임 수는 비디오의 전체 프레임 수이며 추출 장면수는 지역적 χ^2 -테스트를 이용했을 경우 임계치 이상일 경우 검출된 장면 전환 수이다.

표 2. 실험 결과

구분	임계치	총프레임수	추출 장면수
광고 1	430	364	10
광고 2		405	9
광고 3		432	11
뉴스 1	425	2975	22
뉴스 2		2840	23
뉴스 3		5665	37
스포츠 1	470	665	17
스포츠 2		1376	7
스포츠 3		2295	15

5. 결론

본 논문에서는 지역적 컬러 χ^2 -테스트를 제한하는데 이는 기존의 χ^2 -테스트와 컬러 히스토그램을 결합한 방법이다. 기존의 히스토그램 기반 알고리즘과 비교하여 좋은 성능을 보여주는 χ^2 -테스트를 변형하였고, 컬러 값의 세분화 작업에 따른 검출효과를 높이기 위하여 명암도 등급에 따른 가중치를 적용한 지역적 χ^2 -테스트를 이용하였다. 이 방법은 급진적 장면전환과 점진적 장면전환을 모두 검출할 수 있는 강건한 방법을 제시하여 기존에 장면전환 검출 방법이 안고 있던 문제를 해결하였다. 향후 연구 과제로는 비교하는 프레임 수를 획기적으로 줄이는 방법과 객체와 모션에 정확히 작용하는 모션 벡터를 이용한 방법들이 연구되어야 할 것으로 사료된다.

참고 문헌

[1] Bovic, Al C. "Handbook of Image Video Processing", Communications, Networking and Multimedia Series, Edition 2, 2005.
 [2] Zhang, H. J., Kankanhalli, A., and Smoliar, S. W., "Automatic Partitioning of Full-motion Video," Multimedia Systems, Vol. 1, No. 1,

pp. 10-28, 1993.
 [3] Shahraray, B., "Scene Change Detection and Content-Based Sampling of Video Sequences," in Digital Video Compression: Algorithms and Technologies, Proc. SPIE 2419, pp. 2-13, February 1995.
 [4] Hampapur, A., Jain, R., and Weymouth, T., "Digital Video Segmentation," Proc. ACM Multimedia 94, pp. 357-364, October, 1994.
 [5] Kasturi, R. and Jain R., "Dynamic Vision," In Computer Vision: Principles, IEEE Computer Society Press, Washington, 1991.
 [6] Y. Tonomura, "Video handling based on structured information for hypermedia systems, in: Proc. ACM Int. Conf. Multimedia Information Systems, 1991, pp.333-344.
 [7] Nagasaka, A. and Tanaka, Y., "Automatic Video Indexing and Full-Video Search for Object Appearances," in Visual Database Systems II, Knuth, E., Wegner, L., Editors, Elsevier Science Publishers, pp. 113-127, 1992.
 [8] Ueda, H., Miyatake, T., and Yoshizawa, S., "IMPACT: An Interactive Natural-motion-picture Dedicated Multimedia Authoring System," in proceedings of CHI, 1991 ACM, pp. 343-350, New York, 1991.
 [9] U. Gragi, R. Kasturi, S. Antani, "Evaluation of video sequence indexing and hierarchical video indexing," in: Proc. SPIE Conf. Storage and Retrieval in Image and Video Databases, 1995, pp. 1522-1530.
 [10] 고경철, 이양원, "비디오 분할을 위한 자동 임계치 결정 알고리즘," 한국컴퓨터정보학회논문지, 제10권, 제6호, pp.65-74, 2005.12.
 [11] 신성운, 강일고, 이양원, "자막정보를 이용한 축구 비디오 하이라이트 생성," 한국컴퓨터정보학회논문지, 제10권, 제5호, pp.65-76, 2005.11.