

칼라 히스토그램과 엔트로피를 이용한 동영상 장면전환 검출

송현석¹⁾, 안강식¹⁾, 안명석¹⁾, 조석제¹⁾

Scene Change Detection of Video Data Using Color Histogram and Entropy

Hyun-Seok Song¹⁾, Kang-Sik Ahn¹⁾, Myung-Seok An¹⁾, Seok-Je Cho¹⁾

Abstract

In content-based video data retrieval, the detection for scene change is needed. Color histogram approaches are generally used. Color histogram approaches are simple and more robust to object and camera movements but they ignore the spatial information and fail when two different frames have similar histograms. We propose a new algorithm to use color histogram and entropy to enhance the false-positive occurred by brightness variation and the miss by similar color histogram. At the experiments, it is more robust to brightness variation and have spatial information.

1. 서론

디지털 영상 기술이 발달하고, VOD(Video On Demand)등과 같은 서비스가 증가함에 따라 동영상 데이터가 급속히 증가하고 있다. 이런 동영상 데이터를 효과적으로 관리, 검색하기 위해 최근 많은 연구가 진행되고 있다[1-3]. 동영상을 관리하고 검색하기 위해서는 먼저 동영상 데이터를 연관성이 있는 장면별로 분리하는 것이 효과적이며 이를 위해 동영상 데이터의 장면전환 기준이 되는 컷(cut)을 검출하는 일이 중요하다[4].

컷 검출을 위해 일반적으로 칼라 히스토그램을 많이 이용한다[5,6]. 칼라 히스토그램은 물체의 움직임과 약간의 가려짐(occlusion)에 강인하게 컷을 검출할 수 있다. 그러나 공간정보를 포함하지 않고 밝기변화에 민감하여, 컷이 발생하였지만 비슷한 칼라 히스토그램을 가질 경우 놓칠 수 있으며 밝기변화가 일어난 부분에서 컷으로 잘못 검출할 수 있는 단점이 있다.

본 논문에서는 기존의 칼라 히스토그램을 이용한 컷 검출의 문제점을 해결하기 위해 칼라 히스토그램과 블록화된 영상에서의 엔트로피(entropy)를 함께 이용한 방법을 제안한다. 칼라 히스토그램은 적은 계산량에 비해 뛰어난 검출능력이 있으나 밝기변화에 민감하며 공간정보를 가지지 못한다. 엔트로피는 정보의 복잡도를 나타내는 방법으로[7], 영상에 적용할 수 있다. 엔트로피는 영상에 밝기변화가 생길 경우 복잡도의 변화는 크지 않으므로 밝기변화에

1) 한국해양대학교 공과대학 기계정보공학부

강인한 장점이 있다. 또한 영상을 작은 여러 개의 블록으로 나누어 얻은 엔트로피 값은 공간정보를 포함한다. 제안한 방법은 칼라 히스토그램과 엔트로피를 함께 사용함으로써 칼라 히스토그램이 가지지 못하는 공간정보를 포함할 수 있고 밝기변화에 강인하다. 실험을 통해 제안한 방법이 칼라 히스토그램을 사용하는 것보다 결과가 우수함을 확인한다.

2. 동영상의 장면전환 검출에 대한 연구

2.1 동영상 검색과 키 프레임

동영상 검색은 동영상에서 칼라, 움직임, 물체의 형태 등의 특징을 자동으로 추출한 다음 이 특징을 이용해서 원하는 장면을 검색하는 것이다. 동영상 데이터를 저장하고 검색하기 위해 동영상 데이터를 시간적으로 분할할 필요가 있다. 이에 대한 첫 번째 단계로 동영상 데이터의 가장 세부단위의 기준이 되는 컷을 검출하는 과정이 필요하다.

동영상 데이터의 시간적인 구조를 살펴보면 그림 1과 같이 동영상 데이터의 구조를 계층적으로 표현할 수 있다[4,8]. 전체 동영상 데이터를 큰 주제별로 몇 개의 시퀀스(sequence)로 나누며, 시퀀스는 독립적인 주제를 가지는 장면(scene)으로 나누어진다. 또한 장면은 여러 개의 샷(shot)으로 이루어진다. 하나의 카메라에 의해 끊기지 않은 연속적인 장면을 샷이라 하며 샷과 샷 사이의 경계를 컷이라 한다. 일련의 샷을 표현할 수 있는 하나의 프레임을 키 프레임(key-frame)이라 한다.

2.2 칼라 히스토그램을 이용한 동영상 장면전환 검출

칼라 히스토그램은 영상의 칼라분포를 각 칼라별 빈도수로 나타낸 것이다. 칼라 히스토그램은 영상별로 그 모양이 달라 영상의 구별이 쉬우며, 영상 내의 물체를 바라보는 관점, 방향의 변화 등에 영향을 작게 받기 때문에 영상검색과 동영상의 장면전환 검출에서 널리 사용된다. 연속되는 두 프레임 간의 칼라분포는 장면전환이 일어나지 않는 한 유사하며, 반대의 경우 두 프레임의 칼라 히스토그램 모양의 변화가 커진다. 그러므로 동영상 장면전환 검출에 칼라 히스토그램 차이를 사용한다. 연속되는 두 프레임의 칼라 히스토그램 차이를 나타낸 식은 식 (1)과 같다.

$$Diff_H = \sum_{i=0}^{L-1} |H_m(i) - H_{m+1}(i)| \quad (1)$$

여기서 $Diff_H$ 는 칼라 히스토그램 차이, $H_m(i)$ 는 m 번째 프레임의 칼라 히스토그램 그리고 L 은 화소값의 범위를 나타낸다.

칼라 히스토그램 비교방법은 물체의 움직임과 카메라 동작 및 약간의 가려짐에도 좋은 컷 검출 성능이 있으나, 영상의 공간정보를 포함하지 못하며, 밝기변화에 민감하다는 문제점이 있다[1,2,4,5]. 즉, 컷이 아님에도 비슷한 칼라 히스토그램을 가지거나 컷이 있는 부분인데도 밝기변화나 조명같은 효과가 생길 경우 컷을 잘못 검출하거나 놓칠 수 있다. 그림 2는 칼라 히스토그램 비교방법을 사용하여 컷을 검출하였을 때 오류가 생긴 예이다. 그림 2-(a)는 연속된 두 프레임이 컷은 아니지만 밝기변화로 인해 컷으로 잘못 검출된 것이며, 그림 2-(b)는 컷이지만 칼라 히스토그램이 유사한 이유로 검출되지 못한 예를 보여주고 있다.

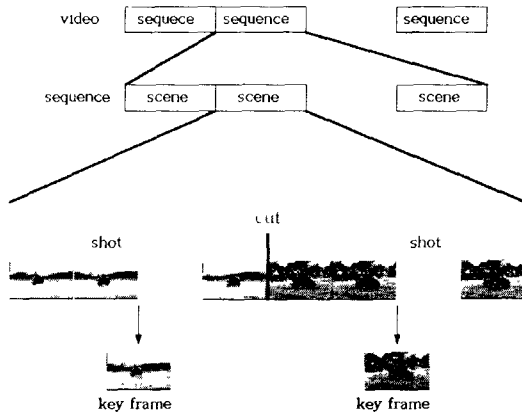
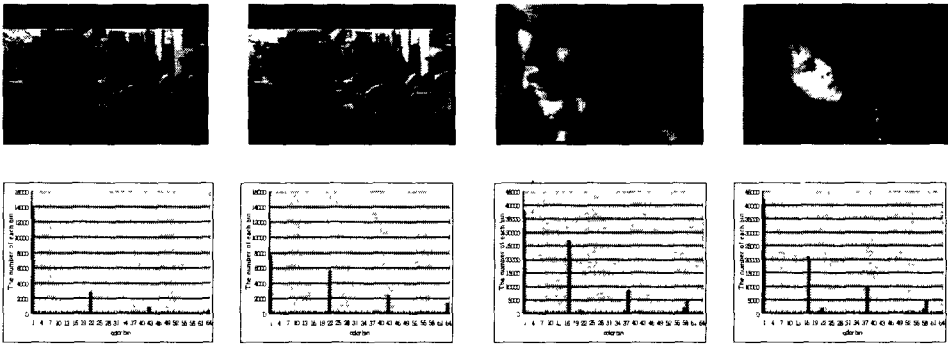


그림 1 동영상 데이터의 구조



(a) 밝기 변화 때문에 잘못 검출된 컷

(b) 히스토그램이 비슷하여 놓친 컷

그림 2 칼라 히스토그램을 사용하였을 때 생기는 문제점

이러한 문제점을 해결하기 위해 본 논문에서는 칼라 히스토그램과 블록화된 영상에서 얻어진 엔트로피를 이용한 방법을 제안한다. 이는 칼라 히스토그램을 이용한 영상의 칼라특징과, 칼라 히스토그램의 단점을 보완하기 위해 영상을 작은 여러 개의 블록으로 나누어 각각의 블록에서 구한 엔트로피의 분산을 이용하여 컷을 검출하는 방법이다. 이 방법은 칼라 히스토그램이 가지지 못하는 공간정보를 포함하며 밝기변화에 민감하지 않다.

3. 제안한 동영상 장면전환 검출기법

본 논문에서는 칼라 히스토그램 비교 방법과, 이 방법의 문제점을 보완하기 위해 블록화된 영상에서의 엔트로피 분산을 함께 사용하여 컷을 검출하는 방법을 제안한다.

3.1 영상에서의 엔트로피 특징

엔트로피는 정보의 복잡도를 측정하는 척도(measure)이다. 일반적인 정보의 복잡도에 대

한 측정 척도로써 새년의 엔트로피(Shannon's entropy)가 쓰이며[7] 식 (2)와 같다 엔트로피는 사건의 확률이 높아질 경우 값이 작으며, 반대의 경우 값이 크다

$$I(a) = -\log P(a) \tag{2}$$

여기서 $P(a)$ 는 사건 a 에 대한 확률이며, $I(a)$ 는 사건 a 에 대한 자기정보(self-information)라고 한다. 여러 사건에 대한 전체 자기정보를 측정할 경우 각 사건의 자기정보에 대한 평균값으로 나타내며 식 (3)과 같이 사용한다.

$$E(a) = -\sum_{j=1}^J P(a_j) \log P(a_j) \tag{3}$$

여기서 $E(a)$ 를 불확실성(uncertainty) 또는 엔트로피라고 한다. J 는 사상의 개수, a_j 는 j 번째 사건을 뜻한다. 식 (3)은 하나의 시스템에서 각각의 사건의 확률에 대한 자기정보의 평균을 뜻한다. 여기서 식 (3)의 값은 여러 사건들의 확률이 고르게 분포되어 있을 경우 커지며, 하나 또는 몇몇 사건들의 확률만 높게 분포되어 있을 경우 그 값이 작아진다. 엔트로피를 이용하여 영상 정보를 표현할 경우 이 식에서 사건의 확률은 각각의 화소값에 대한 빈도수가 된다.

Chang은 엔트로피를 측정하는 방법으로 PIM(Picture Information Measure)을 제안하였으며[9,10] 그 식은 식 (4)와 같다. PIM 값은 히스토그램을 모두 누적한 값에 최고 빈도를 가지는 히스토그램을 뺀 값이다. 화소값이 골고루 분포되어 있을 경우 히스토그램의 최대값은 그리 크지 않아 PIM 값이 커지며, 화소값이 어느 일정한 화소에 높게 분포되어 있을 경우 히스토그램의 최대값이 누적 히스토그램 값에 비해 매우 커지므로 PIM 값이 작아진다.

$$PIM = \sum_{i=0}^{L-1} H(i) - \text{Max}(H(i)) \tag{4}$$

여기서 $H(i)$ 는 영상에서의 히스토그램, $\text{Max}(H(i))$ 는 $H(i)$ 의 최대값, L 은 화소값의 범위를 나타낸다. 본 논문에서는 계산의 편의성과 엔트로피 값의 정량화를 위해, 위의 식 (4)를 사용하였다.

그림 3은 4단계의 명암도를 가진 복잡도가 다른 두 영상이다. 그림 3-(a)는 영상 내의 화소값이 모두 같은 경우로 $H(2) = 16$ 이며 히스토그램의 누적값과 최대값이 같게 된다. 따라서 PIM은 0 이 된다. 그러나 그림 3-(b)와 같이 여러 명도값이 분포하게 되는 경우 $H(0) = 3, H(1) = 3, H(2) = 6, H(3) = 4$ 이다 따라서 $\text{Max}(H(i)) = 6$ 이며 PIM은 10이다. $\text{Max}(H(i))$ 가 히스토그램의 누적값에 비해 작은값을 가지게 되므로 큰 PIM 값이 측정된다. 따라서 복잡한 영상이 엔트로피가 높다

크기가 서로 다른 영상의 PIM 값을 비교하기 위해 식 (4)의 PIM 값을 정규화할 필요가 있다. 식 (5)와 같이 히스토그램 누적값은 1로 표현하고 히스토그램 최대값은 확률밀도함수로써 나타낸다.

$$NPIM = 1 - \text{Max} (p(H(i))) \quad (5)$$

여기서 p 는 확률밀도함수이다.

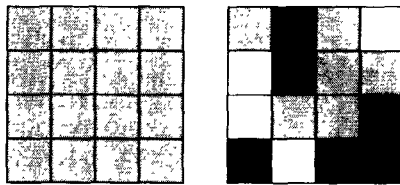
식 (4)와 같이 최대값을 사용할 경우 잡음의 영향을 받거나 너무 밝거나 어두운 영상에서 오류가 발생할 수 있으므로 임의로 k 개의 빈도수가 높은 값으로 표현할 수 있다. 그 식은 식 (6)과 같으며 정규화된 식은 식 (7)과 같다.

$$PIM_k = \sum_{i=0}^{k-1} H(i) - \text{Max}_k H(i) \quad (6)$$

$$NPIM_k = 1 - \text{Max}_k p(H(i)) \quad (7)$$

동영상에서 페이드 인(fade-in), 페이드 아웃(fade-out)이나 조명의 영향 등으로 다음의 그림 4와 같이 밝기변화가 생기는 경우가 있다. 이런 경우 $\text{Max}(H(i))$ 가 나타나는 위치가 변할 뿐 그 값은 크게 변하지 않는다. 이러한 특징은 이와 같은 밝기의 변화에 민감한 일반적인 히스토그램의 차이가 가지는 약점을 보완해준다.

그림 5는 서로 다른 공간정보를 가지지만 엔트로피가 같은 두 영상을 나타낸다. 화소값의 분포가 같기 때문에 같은 엔트로피를 가진다. 여기서 엔트로피가 영상의 공간정보를 포함하지 않는다는 것을 알 수 있다.



(a) PIM=0 (b) PIM=10

그림 3 PIM 값의 비교: 경우 1

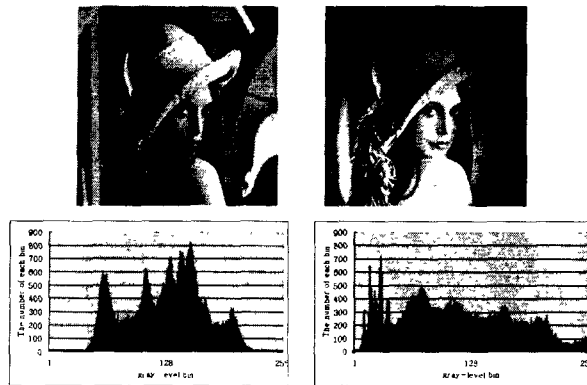
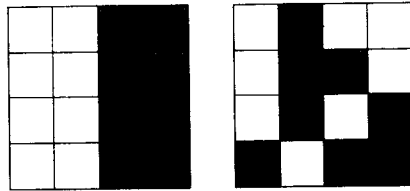


그림 4 밝기변화가 생긴 경우와 명암도 히스토그램



(a) PIM=8 (b) PIM=8
 그림 5 PIM 값의 비교 : 경우 2

3.2 칼라 히스토그램과 블록화된 영상에서의 엔트로피를 이용한 동영상 컷 검출

본 논문에서는 칼라 히스토그램과 블록화된 영상에서의 엔트로피를 함께 사용하는 방법을 제안하였다. 공간정보를 가지는 엔트로피를 위해 영상을 8×8화소 크기의 블록으로 나누어 각각의 블록에 대한 엔트로피를 구하고, 그 엔트로피 값에 대한 분산을 이용해 연속된 두 프레임의 분산 차이를 측정하였다.

식 (8)은 연속되는 두 프레임의 PIM 값들에 대한 분산의 차이를 나타내며, 식 (9)는 식 (1)의 칼라 히스토그램 차이와 식 (8)의 합으로 나타낸다. 즉 칼라 히스토그램의 차이와 엔트로피의 분산의 가중치 합으로 나타낼 수 있다.

$$Diff_E = Var[PIM]_m - Var[PIM]_{m+1} \quad (8)$$

$$Diff_{TOTAL} = Diff_H + wDiff_E \quad (9)$$

여기서 w 는 가중치이다.

그림 6은 영상을 8×8화소 크기의 블록으로 나누어 각각의 블록에 대한 엔트로피를 구한 후, 엔트로피 값의 평균에 대해 높은 엔트로피 영역과 낮은 엔트로피 영역으로 나눈 모습이다. 복잡한 영역의 엔트로피가 높으며, 물체가 배경에 비해 높은 엔트로피 영역을 가진다.

그림 7은 식 (4)를 이용해 두 영상의 엔트로피 평균과 분산을 나타낸 것이다. 이 영상에서 평균값이 큰 것은 영상에 엔트로피가 큰 영역이 많다는 것을 의미하고 분산이 작다는 것은 영상에 전체적으로 블록에 대한 엔트로피의 크기 차이가 작다는 것을 의미한다. 본 논문에서는 이러한 특징을 이용하여 칼라 히스토그램이 가지지 못하는 공간정보를 포함하기 위해 PIM의 분산을 칼라 히스토그램과 함께 사용하였다.



(a) 원영상 (b) 엔트로피가 높은 영역

그림 6 엔트로피에 따른 영역분할

제 4 장 실험 및 결과

실험에 사용한 데이터는 다양한 속도의 움직임과 느낌이 있는 영화, 광고 및 뮤직비디오 등의 여러 종류의 데이터를 사용하였고, 정답 컷의 개수는 사람이 결정하였다. 이 데이터의 특징은 표 1과 같다. 여기서 'Titanic'과 '광고 1'은 다른 데이터에 비해 움직임이 빠르며, 'Shall We Dance'와 '뮤직비디오'는 각각 동영상 흐름상 전체적으로 비슷한 분위기를 가진다. 엔트로피 계산은 256단계 명암도를 사용하였고, 칼라 히스토그램의 계산을 위한 칼라 양자화는 64단계로 하였다.

그림 8은 'Shall We Dance' 데이터에 대해 가중치를 변화시킨 결과를 나타낸 그래프이다. 그림과 같이 칼라 히스토그램에 많은 가중치를 주게 되면 검출개수는 증가하나 잘못검출의 개수와 놓침의 개수까지 증가하게 되며, 엔트로피에 많은 가중치를 주면 잘못검출과 놓침의 개수는 줄일 수 있으나 상당한 칼라정보를 무시하게 되므로 검출의 개수도 같이 감소한다는 것을 알 수 있다. 가중치의 변화는 정답과 오류의 타협점을 찾는 문제이기 때문에 본 논문의 실험에서는 칼라 히스토그램 비교 방법과의 직접적인 비교를 위해 가중치를 1로 고정하였다.

임계치는 모든 데이터를 여러 차례의 실험을 통해 가장 결과가 좋은 15%로 고정하였다. 실험결과는 표 2와 같이, 칼라 히스토그램 비교방법만을 사용하는 것보다 엔트로피 비교방법을 함께 사용하는 것이 좋은 결과가 나오는 것을 알 수 있다. 이는 그림 9, 그림 10과 같이 컷이 발생하였지만 칼라 히스토그램이 비슷하여 놓친 부분이나, 밝기변화에 의해 컷이 아님에도 잘못 검출한 부분을 개선한 효과이다.



(a) 평균: 55
분산: 27



(b) 평균: 33
분산: 280

그림 7 두 영상에 대한 PIM값의 평균과 분산

표 1 실험 데이터의 특성

종류	시간(분:초)	크기	프레임	컷의 개수
Shall We Dance	16:35	160×120	3frame/sec	89
Titanic	04:20	176×120	3frame/sec	75
뮤직비디오	04:27	176×120	3frame/sec	152
광고 1	00:46	176×120	10frame/sec	39
광고 2	00:47	176×120	10frame/sec	27

Shall We Dance

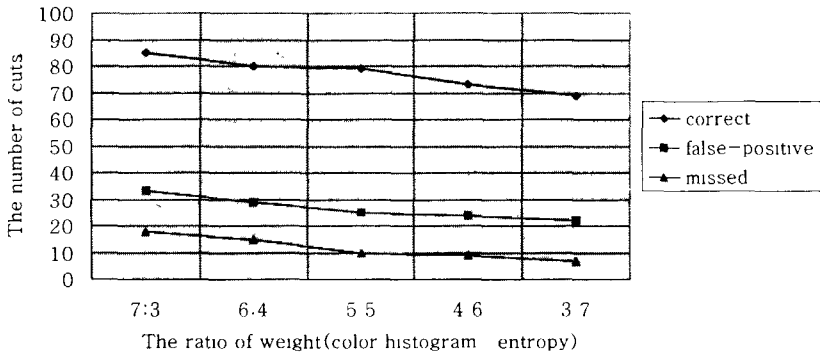


그림 8 가중치 변화에 따른 결과(임계치 15%)

표 2 실험결과

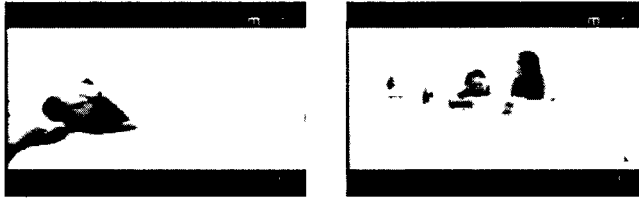
(a) 칼라 히스토그램 비교방법

종 류	검출 (correct)	잘못검출 (false-positive)	놓침 (missed)
Shall We Dance	72	28	17
Titanic	65	22	10
뮤직비디오	116	37	36
광고 1	35	16	4
광고 2	27	6	0

(b) 제안한 방법

종 류	검출 (correct)	잘못검출 (false-positive)	놓침 (missed)
Shall We Dance	79	25	10
Titanic	69	20	6
뮤직비디오	126	33	26
광고 1	37	15	2
광고 2	27	6	0

'광고 2'의 경우 컷의 구분이 확연하고 다른 데이터에 비해 정적인 느낌을 주는 데이터이기 때문에 기존의 방법과 같은 결과를 보인다. 실험결과에서 제안한 방법은 전체적으로 비슷한 분위기가 지속되는 영화나 뮤직비디오 등의 컷 검출에 뛰어난 효과가 있다. 위의 실험 데이터에서 뮤직비디오의 주된 분위기는 밝은 배경에 두 여자가 나오는 것이다. 따라서 칼라 히스토그램이 비슷해서 놓치는 컷이 있으나, 제안한 방법으로 개선하였다.

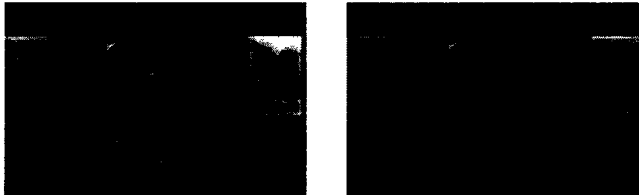


(a) 뮤직비디오

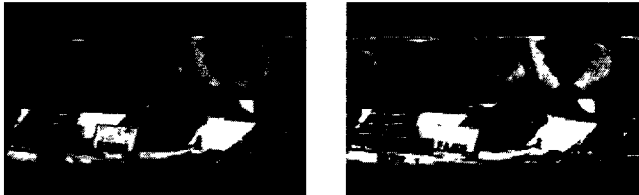


(b) 광고 1

그림 9 실험결과 1 (유사한 칼라 히스토그램)



(a) Shall We Dance



(b) Shall We Dance

그림 10 실험결과 2 (밝기변화가 생긴 부분)

실험결과와 성능을 평가하기 위해 재현률(recall)과 정확도(precision)를 식 (10)과 식 (11)에 의해 구하였다.

$$\text{Recall} = \frac{\text{Correct}}{\text{Correct} + \text{Missed}} \quad (10)$$

$$\text{Precision} = \frac{\text{Correct}}{\text{Correct} + \text{FalsePositive}} \quad (11)$$

재현률이 높다는 것은 올바른 컷을 많이 놓치지 않았다는 것을 뜻하며, 정확도(precision)가 높다는 것은 잘못 검출된 컷의 개수가 적다는 것을 뜻한다. 결과는 아래의 표 3과 같다.

표 3 재현률과 정확도

종 류	칼라 히스토그램 비교 방법		제안한 방법	
	Recall	Precision	Recall	Precision
Shall We Dance	0.81	0.72	0.89	0.76
Titanic	0.87	0.75	0.92	0.78
뮤직비디오	0.76	0.76	0.83	0.79
광고 1	0.90	0.69	0.95	0.71
광고 2	1	0.82	1	0.82

5. 결 론

본 논문에서는 동영상 컷 검출을 위해 칼라 히스토그램 비교방법에 블록화된 영상에서의 엔트로피를 함께 사용하는 방법을 제안하였다. 칼라 히스토그램은 계산이 간단하고 빠르면서도 컷 검출력이 뛰어나지만 밝기변화에 민감하며 공간정보를 가지지 못한다. 칼라 히스토그램의 단점을 보완하기 위해 사용한 엔트로피는 밝기변화에 민감하지 않고 영상을 여러 개의 작은 블록으로 나누어 측정하면 공간정보를 포함한다.

실험을 통하여 칼라 히스토그램 비교방법만을 사용하는 것보다 제안한 방법이 더 우수함을 알 수 있다. 밝기변화는 페이드 인(fade-in), 페이드 아웃(fade-out)이나 급격한 조명변화 등의 영향으로 인해 발생할 수 있으며, 칼라 히스토그램은 동영상 내용이 비슷한 분위기가 지속되는 영상의 경우 그 모양이 유사하다. 제안한 방법은 이런 특징을 가지는 동영상에 대해 강한 컷 검출능력이 있다.

차후 가중치와 임계치의 최적값을 산출하는 법도 연구되어야 하며, 제안한 방법을 확장하여 급격한 장면전환만이 아니라 점진적인 장면전환이 있는 컷을 검출하는 방법도 연구되어야 하겠다.

참고문헌

[1] Irena Koprinska and Sergio Carrato, "Temporal Video Segmentation : A Survey," Signal Processing : Image Communication, v.16, pp. 477-500, 2001.

[2] G. Ahanger and T.D.C Little, "A survey of technologies for parsing and indexing digital video," Journal of Visual Communication and Image Representation 7(1), pp. 28-43, 1996.

[3] H. J. Zhang, A. Kankanhalli, and S.W. Smoliar, "Automatic partitioning of full-motion video," Multimedia Systems 7, pp 119-128, 1993.

[4] R Zabih, J. Miler, and K.Mai, "A feature-based algorithm for detecting and classifying production effects," Multimedia Systems 7, pp. 119-128, 1999.

- [5] A. Nagasaka and Y. Tanaka, "Automatic video indexing and full-video search for object appearances," *Visual Database Systems II*, pp. 113-127, 1995.
- [6] John S. Boreczky and Lawrence A. Rowe, "Comparison of video boundary detection techniques," *IS&T/SPIE*, vol. 2670, pp. 170-179, 1996.
- [7] Rafael Gonzalez and Richard Woods, *Digital Image Processing*, Addison-Wesley, 1992.
- [8] J. Y. Lee, S. Y. Jeong, B. T. Chun, and Y. J. Bae, "A Method for Shot Boundary Detection and R-Frame Selection of Digital Video," *Proc. of Visual Database Systems 4(VDB 4)*, pp. 361-375, 1998.
- [9] 김태희, 정동석, "엔트로피와 색채 특징을 이용한 영상 검색 기법", *정보과학회논문지*, 제26권 제3호, pp. 282 - 290, 1999.
- [10] Seyoon Jeong, Kyuheon Kim, Byungtae Chun, Jaeyeon Lee, and Younglae J. Bae, "Image Indexing using Entropy and Color Correlation features," *ITC-CSCC*, vol. 2, pp. 963-966, 1999.