

Wavelet Transform 을 이용한 Key-Frame 검색

정세은, 김규현, 전병태, 이재연, 배영래

한국전자통신연구원 컴퓨터소프트웨어 연구소

Retrieval of Key-frames using Wavelet Transform

Seyoon JEONG, Kyuheon KIM, Byungtae CHUN, Jaeyeon LEE and Younglae BAE

Electronics and Telecommunications Research Institute

Compture-Software Technology Laboratory

Abstract

본 논문에서는 동영상 데이터베이스에서 key-frame 을 검색하는 방법을 제안한다. 본 논문에서는 key-frame 을 검색하기 위해 컬러 피처를 공간영역에서 추출하지 않고 wavelet transform 영역에서 컬러 피처를 추출하는 방법을 제안한다. wavelet transform 의 저주파 밴드는 영상전체의 특징을 잘 나타내고 고주파 밴드는 texture 와 국부적인 컬러 특성을 잘 나타낸다. 색인과정 알고리즘은 영상의 크기를 정규화하고 RGB 컬러공간에서 HSV 컬러 공간으로 변환을 하여, H,S,V 각 채널에 대해 Daubechies' wavelet transform 을 수행한 후 변환 영역에서 피처를 추출하게 된다. 색인을 위한 피처로 wavelet 계수와 lowest 밴드의 평균과 표준편차를 추출하였다. 효율적인 검색을 위해 검색은 2 단계로 수행된다. 먼저 평균과 표준편차만을 이용한 1 차 검색을 통해 2 차 검색의 후보 영상들을 추출하고 2 차 검색에서는 이 1 차 검색 통과 영상들에 대해서만 wavelet 계수들을 비교하여 최종 검색 결과를 얻게 된다. 검색결과 기존의 컬러 피처를 이용한 방법보다 우수한 검색결과를 얻을 수 있었다.

1. 서론

동영상 데이터 베이스에서 동영상은 계층적인 구조로 저장 된다. 계층적 구조는 동영상, sequence, scene, shot 4개의 레벨로 구성되며 최하위 레벨인 shot에서 key-frame을 추출하게 된다.

동영상 데이터베이스에서 검색을 하게 될 때 key-frame 검색을 주로하게 되는데, 일반적으로 내용 기반 영상 검색 기법을 사용한다. 내용기반 영상 검색에서 주로 사용하는 피처는 컬러 피처이다. 이는 컬러 피처가 추출 하기가 용이하고 검색시 검색 속도가 빠르기 때문이다. 컬러 피처로는 컬러 히스토그램이나 컬러 분포 특성이 사용되고 있다.

그러나, 이러한 컬러 피처만으로는 정확히 원하는 내용을 가진 영상을 검색하기 어렵다. 이는 컬러 피처가 전역 피처이기 때문이다. 전역 피처는 영상의 회전이나 약간의 위치변화에는 영향을 받지않는다는 장점이 있으나, 공간 정보를 전혀 포함하지 못하는 것이 가장 큰 단점이다. 즉, 이러한 공간정보를 포함하지 못하는 전역 피처의 특성 때문에 컬러를 이용하여 검색할 경우 false positive error를 유발할 수 있다. 영상에서 공간정보를 추출하는 것은 매우 어려운 작업이기 때문에 이러한 공간정보의 부족은, 영상을 부 영역으로(Sub-region) 나누어 각 부 영역에 대해 전역 피처를 구하는 방법들이 제안되고 있다. 이러한 방식을 color layout 색인 방법이라고 한다.

본 논문에서는 새로운 color layout 색인 방법으로 wavelet transform을 사용한 방법을 제안한다. wavelet 계수는 영상의 객체 구성과 국부 컬러 변화 정보를 잘 나타내는 특성이 있다. 이러한 특성을 활용하여 보다 영상의 내용을 잘 표현 할 수 있는 색인 방식을 구현할 수 있다.

2. Key-frame 색인과정

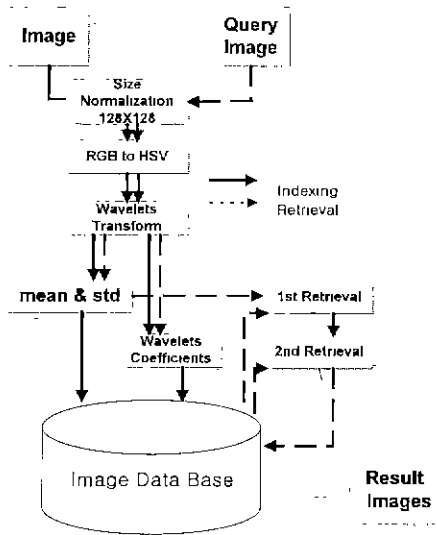


그림 1. Wavelet Indexing/Retrieval System

그림 1은 전체 시스템 블록도이다. 본 절에서는 key-frame 색인 과정을 설명하고 검색 과정은 3 절에서 설명하겠다.

key-frame 영상이 입력되면 먼저 전처리 과정을 수행하고, 전 처리된 영상에 대해 wavelet transform을 하여 wavelet 계수와 평균값들을 색인하게 된다. 전처리로는 입력 영상의 크기를 128X128로 크기 정규화를 하고, 보다 컬러 정보를 효과적으로 추출하기 위해 HSV 공간으로 컬러 공간 변환을 한다. 전처리가 끝난 영상의 각 컬러 채널에 대해 4-layer 2-D Daubechies' wavelet transform을 수행한다. 일반적으로 저주파 밴드에는 영상의 구성 정보가 고주파 밴드에는

texture와 국부 컬러 변화특성정보를 포함하고 있다. 특히, 가장 낮은 저주파 밴드는 영상의 주된 컬러 성분들을 나타낸다. 즉 이 밴드의 wavelet 계수는 영상의 대표 칼라 정보를 나타낸다. 그리고, 이 밴드의 인접한 주변의 3 블록에는 원 영상의 대부분의 구성 정보가 포함되어있다. 고주파 밴드에는 fluctuation 정보가 포함되어있으나 이 정보는 그다지 많은 정보를 포함하고 있지 않다. 영상압축에서는 이러한 특성을 활용하여 저주파 영역 정보는 보존하면서 중요하지 않은 고주파 밴드의 정보를 줄여서 압축을 한다.

제안된 방법에서는 색인데이터로 가장 낮은 밴드에서는 평균과 표준편차를 구하여 저장하였고, 그 다음 인접한 3개의 저주파수 밴드에서는 wavelet 계수를 저장하였다.

3. Key-frame 검색과정

Key-frame 검색은 효율적인 검색을 하기위해 2 단계로 구성된다. 검색 알고리즘은 다음과 같다. 먼저 가장 낮은 저주파밴드의 평균과 표준편차 값을 이용하여 전체 영상에서 2차 검색 후보 영상들을 구한다. 이때 2차 검색의 후보영상은 식(1)의 조건에 의해 결정된다.

$$\left(\sigma'_{ci} > \frac{100}{100 - pstd} * \sigma_{ci} \right) \parallel \left(pstd < \frac{100 * |\sigma_{ci} - \sigma'_{ci}|}{\sigma_{ci}} \right) \\ \parallel \left(m'_{ci} > \frac{100}{100 - pm} * m_{ci} \right) \parallel \left(pm < \frac{100 * |m_{ci} - m'_{ci}|}{m_{ci}} \right) \quad (1)$$

식(1)에서 i 는 1, 2, 3 값을 가지며 컬러 채널을 의미하고, m_{ci} , σ_{ci} 는 각각 i 번째 컬러 채널의 평균과 표준편차이고 pm 과 $pstd$ 은 평균과 표준 편차 허용 오차이다. 조건식 (1)을 만족한 영상들은 허용 오차를 벗어난 영상이므로 2차 검색의 대상에서 제외된다. 즉, 가장 낮은 저주파밴드의 평균과 표준편차가 일정한 오차 범위 '안'의 영상들에 대해서만 좀더 자세한 검색을 수행하게 되므로 검색시간을 줄일 수 있다. 2차 검색은 3개의 밴드의 wavelet 계수들을 비교하여 최

중 유사 영상들을 구한다. 2차 검색에서 사용되는 유사도 함수는 식(2)와 같다

$$D = w_{1,2} \sum_{i=1}^3 w_{c_i} |W'_{c_i,1,2} - W_{c_i,1,2}| + w_{2,1} \sum_{i=1}^3 w_{c_i} |W'_{c_i,2,1} - W_{c_i,2,1}| + w_{2,2} \sum_{i=1}^3 w_{c_i} |W'_{c_i,2,2} - W_{c_i,2,2}| \quad (2)$$

식(2)에서 i 는 컬러 채널을 의미하고 x,y 는 밴드의 위치를 나타내고 $w_{x,y}$ 는 사용자가 정의할 수 있는 각 밴드에 대한 가중치이고 w_c 는 컬러 채널에 대한 가중치이다. $W_{c_i,x,y}$ 는 각 밴드의 Wavelet 계수이다.

4. 실험결과 및 고찰

실험에 사용된 영상 데이터베이스는 영화, 뉴스, 스포츠, 뮤직 비디오, 광고, 다큐멘터리, 애니메이션 등 여러 장르, 총 802장으로 구성되어있다. 그리고, 데이터베이스에서 30장을 선택하여 질의(query) 영상으로 사용하였다. 실험은 제안된 알고리즘과 다른 컬러 layout 색인 방법인 Picture Information Measure(이하 PIM)를 이용한 컬러분포 특성 검색방식과 비교를 하였다 비교 방법으로는 precision과 recall을 구하였다.

표 1. 실험 결과표

	Precision	Recall
PIM	0.797	0.695
제안된 방법	0.839	0.725

표 1에서도 알 수 있듯이 제안된 방법이 PIM 방법에 비해 precision과 recall 모두 우수하다. 특히, wavelet 방법은 texture적인 특성을 갖는 영상에 대해서 PIM 보다 훨씬 좋은 검색 결과를 보였다. 이는 wavelet이 갖는 방향성 정보때문 일 것이다. 그림 2는 제안한 검색 방법으로 검색한 결과 예이다

5. 결론

본 논문에서는 동영상 데이터베이스에서 key-

frame을 검색하기 위한 wavelet transform 영역에서 피처를 추출하여 검색하는 방법을 제안하였고, 검색 과정을 2단계로 하여 검색 시간을 줄여 보다 효율적인 검색을 할 수 있었다. 실험 결과에서도 PIM방법에 비해 우수한 검색 결과를 나타냈다 그러나, 현재의 wavelet을 이용한 검색 방법은 피처의 개수가 너무 많다는 것이 단점이다. 현재 3개의 평균과 3개의 표준편차 그리고 576개의 wavelet 계수를 피처 벡터로 사용하고 있다. 추후 연구 과제는 정보의 손실을 최대한 줄이면서 피처 벡터의 개수를 줄이는 것이다.

참고 문헌

- [1] James Ze Wang and Sha Xin Wei, "Wavelet-Based image indexing techniques with Partial sketch Retrieval capability," *Proceeding of the 4th Forum on Research and Technology Advances in Digital Libraries* pp 13-24, Washington D C May 1997
- [2] Shi-Kuo Chang. "Principles of Pictorial Information Systems Design", Prentice-Hall, pp 61- 81 1989.
- [3] 정세윤, 베잉레 "칼라 분포 특성을 이용한 칼라 인덱싱," *정보과학회 추계 학술 대회*, pp 479-482, Oct 1997
- [4] Martin Vetterli and Jelena Kovacevic. *Wavelets and Subband Coding* Prentice Hall, 1995
- [5] W Niblack and R Barber "The QBIC Project: Querying Images By Content Using Color, Texture, and Shape," *Storage and Retrieval for Image and Video Databases*. PP 173-187, San Jose. SPIE 1993

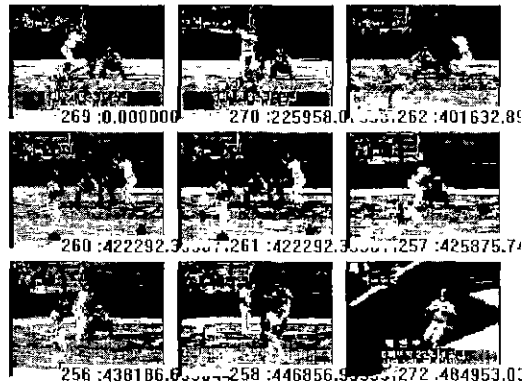


그림 2. Wavelet 검색 방법으로 검색한 결과