

동적 컬러와 모션 정보를 이용한 내용기반 동영상 검색 시스템 구현

A Content-Based Video Retrieval System using Dynamic Color and Motion Information

김 영 재* **, 이 철 희**, 권 용 무*

한국과학기술연구원 영상미디어 연구센터*, 연세대학교 전기·컴퓨터 공학과**

요약

본 논문에서는 내용 기반 동영상 검색을 위한 효율적이고 자동적인 특징 추출 알고리즘을 컬러 정보와 모션 정보에 대해 제안하고, 이를 동영상 검색 시스템에 적용한다. 컬러 정보의 경우 기존의 key-frame단위의 컬러 특징 추출의 한계를 극복하고, 동영상의 컬러 히스토그램 정보와 컬러의 공간분포 정보를 반영할 수 있는 컬러 특징 추출 알고리즘을 제안한다. 그리고 모션 특징은 MPEG-1 동영상 내의 모션 벡터와 컬러 정보를 조합한 컬러-모션 특징을 추출하여 사용한다. 최종적으로 추출된 특징을 이용한 검색 시스템을 구현해, 제안된 알고리즘의 성능을 평가하였다.

1. 서 론

최근 들어 멀티미디어 데이터는 관련기술의 발달로 인해 그 양과 수요가 급격히 증가하였다. 그 중에서도 동영상 데이터는 데이터 압축과 국제 표준화의 도출로 인해 그 비중이 점차 커지고 있어, 이에 대한 관심이 증폭되고 있다. 이러한 동영상 데이터의 양적인 증가는 많은 데이터베이스로부터 원하는 데이터를 검색하는, 효율적이면서 자동적인 검색 알고리즘을 요구하게 되었다. 특정한 응용분야 없이 일반적인 동영상을 대상으로 할 때 고려되는 주요 검색 특징은 컬러와 모션 특징을 들 수 있다. 본 논문에서는 표준화 제정과 함께 그 수요가 급격히 증가하고 있는 MPEG-1 동영상을 대상으로, 이들의 검색을 위한 특징 추출 알고리즘을 컬러와 모션에 대해 제안하고 이를 기반으로 한 검색 시스템을 구현한다. 컬러 특징의 경우 동영상 내에 존재할 수 있는 다양한 컬러 정보 반영과 함께, 이들의 공간 분포를 고려할 수 있는 알고리즘을 제안한

다. 또한 컬러 그룹의 전체적인 공간 분포를 반영하기 위해 양자화 된 컬러 그룹에 속하는 화소들의 좌표 분산값을 특징으로 추출한다. 그리고 모션 특징 추출의 경우는 MPEG-1 동영상 내의 모션 벡터와 컬러 정보의 조합을 통한 컬러-모션 특징을 추출하여 검색의 효율을 높이고자 하였다. 본 논문의 구성은 다음과 같다. 2장에서 동영상 특징 추출과 관련된 기존의 연구를 살펴보고, 3장에서 본 논문에서 제안하는 동영상 특징 추출 알고리즘을 논한다. 4장에서는 이 기반한 검색 시스템을 구현과 이를 통한 제안된 방법의 실험과 평가를 행하며, 5장에서 결론과 향후과제를 논한다.

2. 관련 연구

본 논문에서 기술하게 될 동영상 검색 및 특징 추출 알고리즘은 최근 들어 동영상 데이터의 증가와 더불어 많은 연구가 있어왔다 [3][4]. 일반적으로 동영상의 특징 추출은 컬러와 모션 두 가지 특징을 기초로 이루어지는데, 컬러 특징

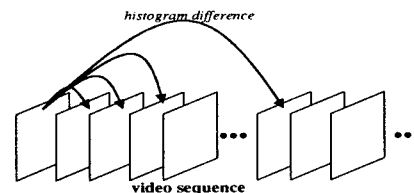
의 경우는 동영상을 shot 단위로 분할하여 각 shot을 대표할 수 있는 key-frame을 추출한 뒤 이에 대한 컬러 특징을 추출하는 방법이 대표적이다. 컬러를 기반으로 한 정지 영상 검색기법은 많은 연구가 있어왔으며, 이에 대한 주요 연구 결과로서는 Swain [6]의 컬러 히스토그램 인터섹션 방법, Mehtre [7]의 Reference color table 방법, 그리고 Stricker [8]의 히스토그램 인터섹션 방법 등이 있다. 모션 특징 추출 알고리즘에는 MPEG-1 동영상을 대상으로 하는 경우, 디코딩 된 이미지 나열을 대상으로 컬러 기반의 계산을 행하는 방법과, 압축된 MPEG-1 동영상 내부에 포함되어 있는 모션 벡터를 사용하는 방법 2가지가 있다. 연속된 이미지 시퀀스로부터 직접 컬러 기반의 계산과정을 거치는 모션 특징 추출 알고리즘에는 여러 알고리즘이 있으며 [10], 이들은 기본적으로 동영상을 이루고 있는 이미지 나열로부터 직접 화소, 혹은 블록의 컬러 값을 기반으로 한 계산과정을 거쳐 모션 특징을 추출하게 된다. 그러나 이러한 방법들의 경우 일반적으로 특징을 추출하는데 드는 계산량과 시간이 상당하다는 단점이 있다. 따라서 이를 해결하기 위해 압축된 MPEG-1 동영상 내에 이미 포함되어 있는 모션 벡터를 이용하는 위치 기반 모션 인덱싱 방법 [9]이 제안되었다. 이 방법에서는 하나의 shot을 대상으로 처음, 중간, 마지막 프레임을 추출하고, 이들 세 프레임의 매크로 블록 내 모션 벡터를 읽어들이어 각 모션 벡터의 방향을 양자화 한다. 그리고 양자화 된 모션 벡터를 이용하여 매크로 블록의 위치에 따라 모션 히스토그램을 작성하여, 기존의 복잡한 계산 없이 MPEG-1 동영상의 특징을 이용한 모션 특징을 추출하고자 하였다. 그러나 이 방법은 시간적인 세 단계 위치의 프레임만을 대상으로 하게 되고, 단순히 위치를 기반으로 한 특정 모션의 존재여부만을 사용하는 문제가 있다.

3. 동영상 특징 추출 알고리즘

3.1 컬러 특징 추출

앞 장에서 밝힌 바와 같은 key-frame 기

반의 컬러 특징 추출에는, 하나의 정지영상으로 각 shot의 컬러 정보를 대표할 수 있다는 가정이 뒤따른다. 그러나 동영상 내의 컬러 정보는 시간의 흐름에 따라 다양하게 변하는 경우가 존재할 수 있으므로, 하나의 shot내부에도 많은 컬러 정보의 차이가 존재할 수 있다. 따라서 본 논문에서는 shot을 컬러 정보를 고려하여 sub-shot으로 재분할한다. 여기서 대상으로 하고 있는 MPEG-1 동영상의 영상 처리 시 많이 쓰이고 있는 DCT계수를 이용한 DC 영상을 추출하여[5] 이용하며, 이렇게 재분할된 sub-shot 단위로 컬러 특징을 추출한다. 아래 그림 1은 이와 같은 sub-shot 분할 과정을 나타낸다.



if histogram diff > threshold, sub-shot is declared.

그림 1. sub-shot 분할 과정

이러한 sub-shot 단위의 특징 추출에 있어, 본 논문에서는 정지영상 단위 특징 추출의 한계를 극복하기 위한 히스토그램 행렬 비교 방법 [11]을 사용한다. 또한 이 방법이 정지 영상을 대상으로 하는 것이므로, 이를 동영상의 컬러 정보 추출에 적합하도록 개선시켜 sub-shot 단위의 컬러 히스토그램 행렬을 생성한다. 우선 각 sub-shot내의 각 프레임에 대해, 정지 영상을 대상으로 한 컬러 히스토그램 행렬을 생성한다. sub-shot의 한 프레임을 대상으로 한 행렬 생성 과정은 다음과 같다.

- ① 각 sub-shot 단위의 양자화 컬러 히스토그램 그룹을 계산하고, 영상을 번호 매김 한다.
- ② 한 프레임 내의 분할된 그룹들의 경계에 해당하는 화소들의 4-neighborhood에 포함되는 이웃 화소들을 선택한다.
- ③ 선택된 그룹 번호를 확인하고, 그 그룹번호의 수를 계산한다.

- ④ 행렬의 대각 성분에는 전체 화소에 대한 자신의 그룹에 해당하는 화소의 비율을 대입한다.
- ⑤ 행렬의 비대각 성분에는 각 그룹들에 대한 이미 계산된, 이웃하는 그룹들의 전체 화소에 대한 비율을 대입한다.
- ⑥ 결국, 주어진 양자화 그룹 수만큼의 정방행렬이 생성된다.

그리고 아래 그림 2는 하나의 프레임에 대해 생성된 히스토그램 행렬의 예를 나타낸다.

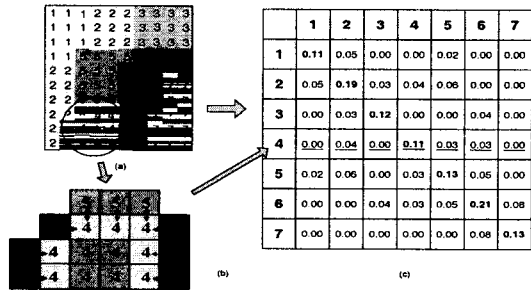


그림 2. 히스토그램 행렬 생성의 예

그림 2(a)는 양자화를 끝낸 후 번호 매김된 영상이고, 이를 바탕으로 생성된 히스토그램 행렬이 그림 2(c)에 나와있다. 그림 2(b)는 그림 2(a)를 일부 확대한 것인데, 그룹 4와 인접한 화소(4-neighborhood)를 카운트하여 행렬의 비대각 성분이 결정되는 것을 보여준다. sub-shot 단위의 히스토그램 행렬은, sub-shot 내에 속하는 프레임들의 히스토그램 행렬 각 성분의 평균을 구함으로써 생성된다. 그러나 동영상에서는 각 컬러 그룹 간의 상관관계가 다양하게 변화할 수 있기 때문에, 평균한 값을 그대로 사용하는 것은 부정확한 결과를 초래할 수 있다. 따라서 검색을 위한 특징을 사용함에 있어 평균값에 대한 신뢰도 문제가 생길 수 있으며, 이를 평가할 수 있는 척도가 필요하다. 이 때 계산된 평균값에 대한 분산이 임계치를 넘는 경우에는 해당하는 컬러 그룹간의 상관관계가 없다고 생각할 수 있으므로, 평균을 통해 구한 행렬의 성분을 $M(i, j)$, 이 성분의 분산을 $V(i, j)$ 라고 할 때 아래 식 (1)과 같이

$V(i, j)$ 가 임계치를 넘게 되면 대응되는 평균 성분을 0으로 보정한다. 그리고 이때의 임계치는 실험적으로 결정되었다.

$$\text{IF } V(i, j) > \text{threshold value, } M(i, j) = 0 \quad (1)$$

이와 같이 생성된 행렬의 대각 성분은 양자화된 그룹의 히스토그램을 나타내고, 행렬의 비대각 성분은 번호 매김된 후의 영상에서 각각의 행과 열의 그룹이 이웃하는 경우의 히스토그램을 나타내며, 이것으로 인해 영상내의 공간상에서 각 컬러 그룹간의 연관성을 알 수가 있다. 따라서 이러한 히스토그램 행렬은 대각 성분을 통해 기존의 히스토그램 기법을 살리면서, 비대각 성분을 통한 컬러 공간상의 각 그룹간의 관계를 고려할 수 있도록 한다.

이웃 양자화 컬러 그룹간의 상관관계를 반영하기 위한 컬러 히스토그램 행렬과 함께, 각 컬러 그룹의 전체적인 공간분포를 고려하기 위한 좌표 분산값을 계산한다. 여기서 좌표 분산 값은 양자화 컬러 그룹에 해당하는 화소들의 분포 정도를 의미하며, 동영상의 프레임 단위로 연산을 행한 뒤 그룹별 분산의 평균을 구하게 된다. 아래 식 (2), (3)이 이를 나타낸다

$$M_i = \frac{1}{N_i} \sum_{(k, l) \in R_i} f(k, l) \quad (2)$$

$$\sigma_i^2 = \frac{1}{N_i} \sum_{(k, l) \in R_i} (f(k, l) - M_i)^2 \quad (3)$$

식 (2)는 한 프레임에서 각 컬러 그룹 화소의 좌표 평균을 계산하는 식이고, 식 (3)은 좌표 평균을 통한 컬러 그룹 화소의 분산을 계산하는 식이다. 여기서 M_i 는 i 번째 그룹에 해당하는 화소들의 좌표값에 대한 평균값이고, σ_i^2 는 i 번째 그룹에 해당하는 화소들의 좌표값에 대한 분산값이다. R_i 는 i 번째 그룹에 속하는 화소들의 집합, N_i 는 R_i 에 속하는 화소들의 수, $f(k, l)$ 은 해당 화소들의 x, y 좌표

값을 나타낸다. 따라서 각 프레임 별로 계산된 좌표 분산의 평균값을 특징으로 추출하여, 앞 절의 컬러 히스토그램의 이웃 컬러 그룹간의 상관관계와 함께, 각 컬러 그룹의 전체적인 공간분포를 고려할 수 있도록 한다.

3.2 컬러-모션 특징 추출

MPEG-1 동영상의 프레임 중 P 프레임에는 앞단의 I나 P 프레임과의 움직임 정보가 매크로 블록 단위의 벡터 형태로 저장되어 있다. 따라서 이들 순방향 모션 벡터를 이용하면 연속적인 움직임 정보를 추출할 수 있으며, 앞서 밝힌 바와 같이 본 논문에서는 이를 양자화 컬러 그룹별로 연산한다. 양자화 컬러 그룹은 앞 절의 컬러 특징 추출에서 사용되었던 양자화 컬러 히스토그램을 사용한다.

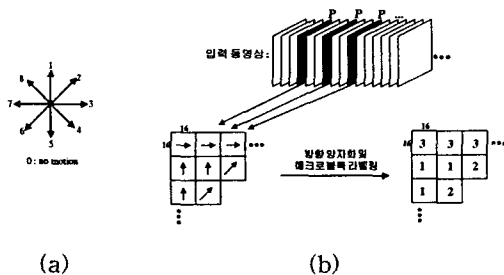


그림 3. 모션 벡터의 양자화 및 매크로 블록 라벨링

연산의 대상이 되는 순방향 모션 벡터는 시간 순으로 입력되는 P 프레임내에 저장되어 있다. 각각의 P 프레임에서 추출된 모션 벡터는, 우선 위의 그림 3(a)와 같이 8 방향으로 양자화 된다. 방향의 양자화는 읽어들이는 모션 벡터의 각(angle)을 구하고 이에 대한 삼각함수 값을 이용하여 구하게 되며, 각 방향은 번호로써 표현된다. 모션이 없는 경우는 '0'으로 나타내고, 입력된 P 프레임 내 각각의 매크로 블록을 저장된 모션 벡터가 갖는 방향에 따라 해당 방향의 번호로 번호 매김한다. 그림 3(b)가 이를 나타내고 있다. 이 때 입력된 프레임에 대한 DC 영상과, 이들에 대한 양자화 컬러 그룹이 컬러 특징 추출을 위해 이미 생성되어 있으므로, 이들 컬러 정보와 앞서 행한 모션 정보와의 정합을 행하여 양자화 컬러 그룹 단

위의 모션 정보를 구한다. 이 때 모션 특징 추출은 한 shot을 대상으로 이루어지므로, 전체 shot을 대상으로 각 컬러 그룹의 모션 정보를 구하는 것임을 고려해야 한다. 따라서 시간 순서대로 입력되는 P 프레임에 대해, 양자화 컬러 그룹들이 차지하는 매크로 블록의 방향 번호와 화소의 수를 읽어들인다. 입력된 동영상의 k 번째 P 프레임에 대해, j 번째 양자화 컬러 그룹에 대한 모션 벡터의 연산이 그림 4에 나와있다.

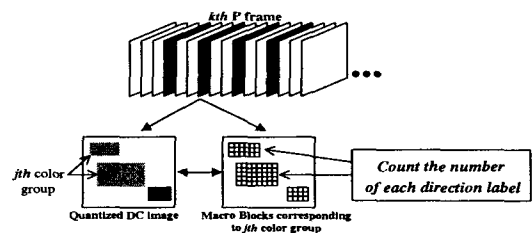


그림 4. 컬러 그룹별 모션 벡터 연산

이 과정을 입력된 동영상의 모든 P 프레임에 대해 행하면, 결국 하나의 양자화 컬러 그룹에 대해, 이 컬러 그룹이 전체 shot에서 차지하는 화소의 수와 각 화소에 대응되는 모션 벡터의 수가 방향별로 카운트된다. 그리고 컬러 그룹별 모션의 번호수를 컬러 그룹의 전체 화소수로 나누어 정규화하면, 그림 5와 같은 모션 행렬이 생성된다.

	color group number						
	0	1	2	3	...	j	...
direction label	0						
1							
2							
3							
4							
5							
6							
7							
8							

그림 5. 컬러 모션 행렬의 생성

생성된 행렬의 각 성분 $m(i, j)$ 는 입력된 동영상에서 j 번째 양자화 컬러 그룹이 i 방향을 갖는 비율을 나타낸다.

4. 시스템 구현 및 결과 고찰

본 논문에서 제안된 컬러와 컬러-모션 특징을 이용한 동영상 검색 시스템을 구현하였다. 컬러 특징 정합은 아래 식 (4)과 같이 컬러 히스토그램 행렬의 차이 D_{mat}^c 과 좌표 분산값의 차이 D_{var}^c 의 합으로 이루어진다.

$$D_{Q,I}^c = (W_1 \times D_{mat}^c) + (W_2 \times D_{var}^c) \quad (4)$$

여기서 Q, I는 각각 질의 동영상과 데이터베이스 내의 동영상을 나타내고, 행렬의 차이값과 좌표 분산의 차이값은 각각의 최대값으로 나누어진 정규화 값이다. 그리고 본 실험에서는 둘의 가중치 W_1 , W_2 가 각각 0.5로 동일한 상태이다. 모션 특징 정합은 제안된 컬러-모션 행렬의 차이값으로 나타나며, 컬러 특징으로 추출된 좌표 분산값을 이용해 연산을 행한다. 즉 모션의 차이를 구하는 데 있어 화면에 존재하지 않는 컬러 그룹까지 연산하는 것은 비효율적이므로, 질의 동영상 내 각 컬러 그룹의 좌표 분산값이 0이 아닌 열만을 대상으로 차이값을 연산한다. 결국 질의 동영상과 데이터베이스의 동영상과의 유사도는 아래 식 (5)와 같이 나타낼 수 있으며, 각각의 가중치 W_c , W_m 은 사용자가 선택할 수 있게 하였다. 여기서 $D_{Q,I}^m$ 는 계산된 컬러-모션 행렬 차이의 정규화 값이다.

$$D_{Q,I} = W_c \times D_{Q,I}^c + W_m \times D_{Q,I}^m \quad (5)$$

본 논문에서 제안한 컬러 특징 추출 알고리즘의 목적이 기존의 key-frame 기반 검색의 한계를 극복하고자 하는 것이었으므로, 구현된 시스템 하에서 두 방법에 대한 실험과 평가를 실시하였다. 실험에 사용한 동영상은 해상도 352×240의 MPEG-1 동영상이며, 그림 6은 이들을 대상으로 구현된 검색 프로그램을 나타낸다. 구현된 프로그램은 Pentium II Processor 상에서 Visual C++ 6.0으로 제작되었으며, 사용자가 원하는 동영상을 불러서 이와 유사한 동영상을 보여주는 구조로 되어있다.

검색된 결과는 그 유사도에 따라 정지 영상 형태로 나타나고, 동영상은 재생과 정지 등이 가능하다. 제안된 알고리즘의 평가를 위해 실험 동영상 데이터베이스에 대해 아래 표 1과 같이 같은 내용을 담고있는 동영상을 그룹으로 생성하였다. 따라서 동영상 1로 질의했을 경우, 검색 결과에는 동영상 2가 반드시 나와야 한다. 실험의 객관성을 위해 반영하는 내용이 확연히 구분되는 것들끼리 그룹으로 생성하였으며, 전체 데이터베이스에 대한 질의 및 검색을 통해 원하는 결과가 나오는 경우와 그렇지 못한 경우를 확인하여 성능을 평가하였다.

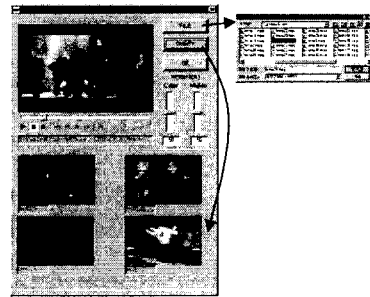


그림 6. 구현된 검색 프로그램

group 1	1, 2	group 2	4, 5
group 3	10, 11	group 4	17, 18
group 5	20, 21, 32, 33	group 6	30, 31

표 1. 그룹으로 생성된 데이터베이스의 예

실험결과 짝으로 생성된 동영상이 결과에 나오지 않는 경우를 검사하여, 아래 식 (6)과 같은 'recall' 항목으로서 제안된 알고리즘과, 기존의 key-frame 기반 검색의 정확도를 평가하였다. 제안된 방법은 가중치를 가변적으로 실험하여 그 중 두 가지를 제시하였고, key-frame 기반 방법은 제안한 방법과 동일한 양자화 과정을 거친 히스토그램 정보로써 검색을 행하였다. 식 (6)에서 correct는 제대로 찾은 결과의 수를 나타내고, missed는 찾지 못한 결과의 수를 나타낸다. 그리고 그 결과표 2에 나와있다.

$$recall = \frac{correct}{correct + missed} * 100 (\%) \quad (6)$$

검색 방법	recall (%)
제안된 방법($W_c=0.8, W_m=0.2$)	96.6
제안된 방법($W_c=1.0, W_m=0.0$)	94.2
key-frame 기반 방법	90.0

표 2. 제안된 방법의 검색 결과

5. 결론 및 향후계획

본 논문에서는 내용기반 동영상 검색을 위한 효과적이고 자동적인 특징 추출 알고리즘을 컬러와 모션 두 가지로 나누어 제안하고, 이를 이용한 검색 시스템을 구현하였다. 컬러 특징의 경우는 동영상의 다양한 컬러 정보의 반영을 위한 sub-shot 단위의 특징 추출을 행하고, 또한 컬러의 공간적인 분포를 고려하기 위한 히스토그램 행렬을 동영상에 맞게 개선하여 이용하였다. 또한 컬러의 전체적인 공간 분포를 고려하기 위한 컬러 그룹 화소의 좌표 분산값을 계산하여 특징으로 사용하였다. 모션 특징은 MPEG-1 동영상 내의 모션 벡터와 컬러 정보를 정합한 컬러-모션 특징을 추출하여 검색 효과를 높이고자 하였다. 앞으로 다양한 동영상 검색을 위한 특징 추가와 함께, 보다 객관적인 정확도 평가를 위한 충분한 데이터 베이스의 구축이 향후과제로 남아있다.

참 고 문 헌

[1] Chang SF, Chen W, Meng HJ, Sundaram H, Zhong D, "A Fully Automated Content - Based Video Search Engine Supporting Spatio temporal Qeries," *IEEE Transactions on Circuits & Systems for Video Technology*, vol.8, no.5, pp. 602-615, 1998.

[2] 김우생, 김진웅, 임문철, "MPEG-7 표준화 및 내용기반 정보 검색," *전자공학회지*, vol.25, no.8, pp. 772-785, 1998.

[3] Atsuo Yoshitaka, Tadao Ichikawa, "A

Survey on Content-Based Retrieval for Multimedia Databases," *IEEE Transactions on Knowledge and Data Engineering*, vol.11, no:1, pp. 81-93, 1999.

[4] Y. Aip Asladogan, Clement T. Yu, "Techniques and Systems for Image and Video Retrieval," *IEEE Transactions on Knowledge and Data Engineering*, vol.11, no.1, pp. 56-63, 1999.

[5] B. L. Yeo and B. Liu, "A Unified Approach to Temporal Segmentation o Motion JPEG and MPEG Compressed Video," *Internal Conf. on Multimedia Computing and Systems*, May. 1995.

[6] M.J. Swain and D.H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp.11-32 1991.

[7] B.M. Mehtre, M.S. Kankanhalli, A.D Narsimhalu and G.C. Man, "Color matching for image retrieval," *Pattern Recognition Letters* vol. 16, pp. 325-331, 1995.

[8] X. Wan and C.C. Jay Kuo, "Color Distribution Analysis and Quantization for Image Retrieval," *Proceedings of SPIE* 2670, pp. 8-16, 1996.

[9] Zaher AGHBARI, Kunihiko KANEKO Akifumi MAKINOUCHE, "A Motion-Location Based Indexing Method for Retrieving MPEG Videos," *Proceedings of the Ninth International Workshop on Database and Expert Systems Applications*, pp. 102-107, 1998.

[10] Christoph Stiller, Janusz Konrad "Estimating Motion in Image Sequences," *IEEE SIGNAL PROCESSING MAGAZINE* pp. 70-91, July, 1999.

[11] 김익재, 이제호, 권용무, 박상희, "변형된 질의 영상에 강한 내용 기반 영상 검색 기법," *한국방송공학회지*, vol. 2, no. 1, pp. 74-83, June, 1997.