비디오 압축 영역에서 MPEG-7 컬러 기술자 추출 연구

배빛나라[1], 이재욱[2], 노용만[3], 강경옥[4]

한국정보통신대학원대학교 멀티미디어 그룹 [1, 2, 3]
한국전자통신연구원 방송미디어연구부 방송컨텐츠응용연구팀[4]

Research for MPEG-7 Color Descriptor Extraction from Compressed Video

Beetnara Bae, Jae Wook Lee, Yong Man Ro and Kyeongok Kang
Correspondence[3]: (e-mail) yro@icu.ac.kr, (phone) 042-866-6129

요 약

디지털 TV 서비스의 시작으로 시청자의 효율적인 방송 컨텐츠 이용의 필요성이 제시되었고, 이를 위한 사용자 맞춤형 서비스와 방송 컨텐츠의 특징 추출 및 표현 방법 등, 지능형 TV에 대한 연구가 수행되고 있다. 지능형 TV의 방송 컨텐츠의 비디오 요약과 비디오 하이라이트 등의 제작을 위해 MPEG-7 비주얼 기술자를 이용하여 비주얼 특성을 추출할 수 있다. 특히 비디오의 컬러 특징을 추출하는 MPEG-7 컬러 기술자를 이용한 특징 추출 위해서는 압축된 비디오를 복호하고 비디오의 프레임을 획득하여 특징을 추출하게 되지만, 이 과정은 비주얼 특징 추출 시간을 증가시켜 압축된 방송용 컨텐츠에 MPEG-7 컬러 기술자의 적용을 어렵게 한다. 본 논문은 압축된 비디오의 컬러 특성 추출 속도의 향상을 위해 비디오 압축 영역에서 MPEG-7 컬러 기술자를 적용할 수 있는 방법을 제안하였다. 압축된 비디오에서 DC 프레임을 추출하여 컬러 특징 추출 시간이 감소하였고 검색 율의 변화는 크지 않았다. 본 논문에서 제안한 방식은 방송용으로 압축된 비디오의 비주얼 특징을 고속으로 추출할 수 있는 방법으로의 활용이 가능할 것이다.

1. 서 론

디지털 TV 서비스가 시작되면서, 시청자의 효율적인 방송 컨텐츠의 이용의 필요성이 제시되었고, 사용자 맞춤형 서비스와 디지털 방송 컨텐츠의 특징 추출 및 표현 등과 같은 지능형 TV 서비스를 위한 SmarTV와 같은 연구가 활발히 진행되고 있다. 디지털 방송 컨텐츠의 특징 추출 및 표현은 사용자의 방송 컨텐츠의 내용 기반 이용을 위하여 필수적인 기능으로서 MPEG-7이나 TVAF(TV Anytime Forum)와 같은 표준에서 기술자 및 기술 방식을 정의하고 있다. 그 중 MPEG-7 비주얼 기술자는 이미지 및 비디오의 비주얼 특성을 분석하는데, 디지털 TV 컨텐츠의 내용 기반 표현을 위해 MPEG-7 비주얼 기술자를 응용할 수 있다.

MPEG-7 비주얼 기술자는 컬러, 질감, 모양, 움직임 등의 이미지 및 압축된 비디오로부터 비주얼 특징을 추출하는 알고리즘과 추출된 특징의 표현을 위한 방법을 정의하고 있다. 이러한 MPEG-7 비주얼 기술자를 이용하여 디지털 방송 컨텐츠 대한 컬러 정보를 추출 할 수 있지만 추출 시간 및 계산의 복잡성 때문에 방송 컨텐츠의 적용에는 많은 문제점이 있다. MPEG-2 방식으로 압축되어 있는 디지털 방송용 비디오로 컬러 기술자를 추출하기 위해서는 압축된 비디오를 복호하여 비디오의 프레임을 획득해야 하고, 이런 과정은 디지털 방송용 비디오의 컬러 특징을 추출하기 위한 시간을 증가시켜 압축된 방송 컨텐츠에 적용을 어렵게 한다.

본 논문은 방송용 컨텐츠의 압축영역에서 컬러 정보에 대한 추출 시간 증가의 문제점을 해결하기 위하여 압축 비디오 영역에서 컬러 기술자를 추출하는 방법에 대해 제안하였다. 제안된 방법은 블록 DCT를 채택하고 있는 MPEG-1,2 형식으로 압축된 비디오의 블록 DCT 계수 영역에서 컬러 특징을 추출하는 방식이다.[1]

2. MPEG-7 컬러 기술자

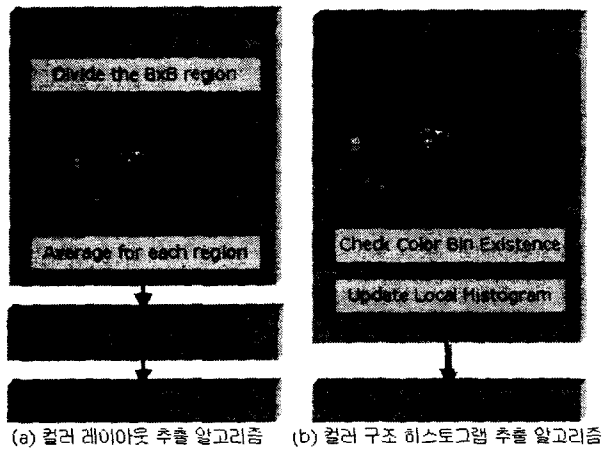
MPEG-7 컬러 기술자는 컬러 공간 기술자(Color Space), 컬러 균등 양자화 기술자(Color Quantization),

컬러 레이아웃 기술자(Color Layout), 컬러 구조 히스토그램 기술자(Color Structure), 주요 컬러 기술자(Dominant Color), 계층적 컬러 히스토그램 기술자(Scalable Color) 그리고 GoF 컬러 기술자(GoFGoP Color)로 총 7개가 있고, 크게 압축된 이미지 혹은 비디오의 헤더 정보로부터 추출할 있는 기술자와 공간 화소 값을 이용하여 특징을 추출하는 기술자로 분류된다.[2,5,6]

i) 컬러 공간 기술자 : 이미지 혹은 비디오의 컬러 공간을 정의한다. RGB, YCbCr, YUV, HMMD 등의 컬러 공간을 정의할 수 있고, 이미지, 비디오 파일의 헤더 정보로부터 컬러 공간 정의를 획득할 수 있다.

ii) 컬러 균등 양자화 기술자 : 각 이미지 및 비디오는의 헤더에서 정의하고 있는 양자화 테이블을 획득하여 컬러 균등 양자화 특징을 정의한다.

iii) 컬러 레이아웃 기술자 : 한 이미지의 컬러 정보를 그림 1의 (a)와 같이 개략적으로 추출한다. 입력된 이미지를 YCbCr의 컬러 공간으로 변환하고, 8x8 화소 크기의 작은 이미지를 각 영역마다 평균 값을 구하여 생성한다. 생성된 작은 이미지의 Y, Cb, Cr의 각 성분에 대해 8x8 DCT를 수행하고 DCT 계수의 개수를 선택함으로써 컬러 특징을 추출한다.

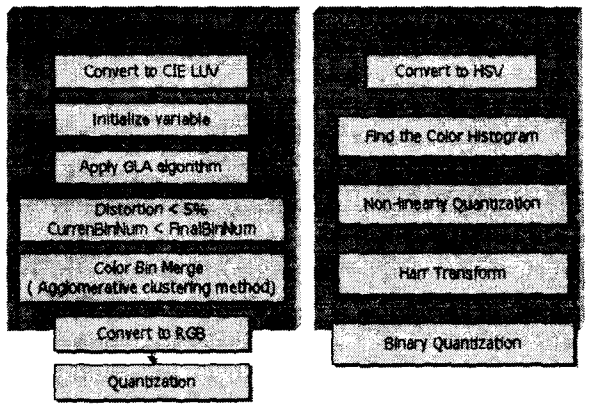


(a) 컬러 레이아웃 추출 알고리즘 (b) 컬러 구조 히스토그램 추출 알고리즘

그림 1. 컬러 레이아웃과 컬러 구조 히스토그램 기술자의 추출 알고리즘

vi) 컬러 구조 히스토그램 기술자 : 한 이미지의 컬러 빈 값의 공간적인 분포를 그림 1의 (b)와 같이 추출한다. CIF 크기(가로크기 352개 화소, 세로크기 288개 화소)의 이미지를 기준으로 하여 8x8 크기의 윈도우 마스크를 이용하여 로컬 히스토그램을 추출한다. 로컬 히스토그램의 컬러 빈 값들이 존재하는 경우에 최종 히스토그램을 갱신하게 되어 컬러 빈 값의 공간적인 구조를 추출하게 된다.

v) 주요 컬러 기술자 : 한 이미지의 대표적인 화소 값을 그림 2의 (a)와 같은 순서에 의해 추출한다. GLA(Generalized Lloyd Algorithm)에 컬러 값의 클러스터의 중심 값을 구하고, 분산 값과 클러스터의 수를 측정하여 반복한다. 클러스터에 의해 획득된 컬러 빈 값을 양자화함으로써 주요 컬러의 특징을 표현한다.



(a) 주요 컬러 기술자 추출 알고리즘 (b) 계층적 컬러 히스토그램 기술자 추출 알고리즘

그림 2. 주요 컬러 기술자와 계층적 컬러 히스토그램 기술자 추출 알고리즘

vi) 계층적 컬러 히스토그램 기술자 : HSV 컬러 공간에서 그림2의 (b)와 같이 컬러 히스토그램을 구하고, 컬러 히스토그램의 결과값을 Harr 변환을 취하여 계층성을 획득한다.

vii) GoF 컬러 기술자 : 이미지 그룹을 대표할 수 있는 컬러 히스토그램을 그림 3과 같이 추출하는 기술자로서 각 이미지의 계층적 컬러 히스토그램 기술자의 특징 값을 추출하고, 그룹의 대표 값을 중간 값 또는 평균 연산을 이용하여 정의한다.

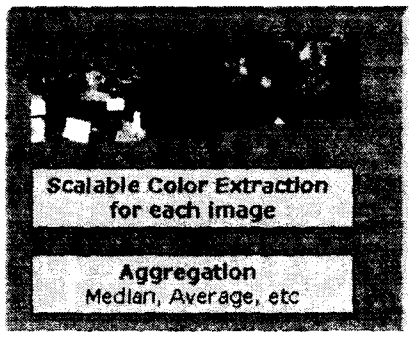


그림 3. GoF 컬러 기술자의 추출 알고리즘

3. 비디오 압축 영역에서의 프레임 추출 방법

본 논문에서 제안한 방식인 비디오의 압축 영역에서 DC 프레임을 추출하여 컬러 특징을 추출하는 방법은 그림 4와 같은 과정으로 수행된다. 입력된 압축된 방송

용 비디오에서 DC 프레임을 추출하기 위하여 I 프레임의 DCT 계수를 이용하여 DC 프레임을 획득한다. I 프레임으로 획득된 DC 프레임과 움직임 벡터를 이용하여 B, P 프레임으로부터 DC 프레임을 추출한다. 원 비디오의 영상의 가로 크기와 세로 크기가 각각 1/8로 축소되고 DC 프레임의 크기는 비디오 한 프레임 크기의 1/64이 된다.[1]

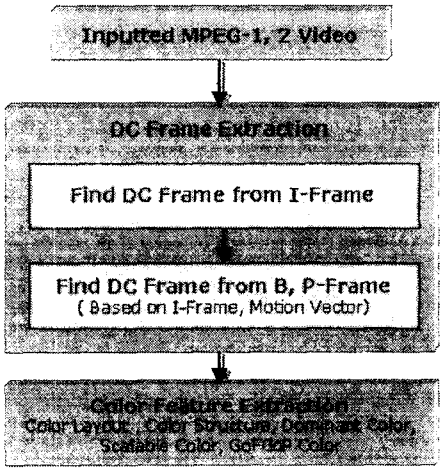


그림 4. 비디오로부터 컬러 특징 추출 방법

3.1. DC 프레임에서의 화소 값의 의미

MPEG-1,2 표준은 압축을 위한 변환으로 블록 DCT(Discrete Cosine Transform)을 채택하고 있고, 블록은 가로 크기 8화소와 세로 크기 8화소로 구성된다. 블록 DCT를 수행하여 획득되는 DCT 계수는 1개의 DC와 63개의 AC 성분으로 구성되고 식 1과 같이 DC 성분은 64개의 화소의 컬러 및 밝기 값의 평균이 된다. [1]

$$DCT\ (0,0) = \frac{1}{64}\sum_{k=0}^{7}\sum_{j=0}^{7} block\ (k,j) \qquad (1)$$

$DCT(0,0)$ 하나의 DCT 블록의 DC 값을, $block(k,j)$ 는 하나의 블록 내의 k, j 번째 화소 값을, k 와 j 는 블록 내의 인덱스를 의미하고 각각 0 부터 7까지의 값을 가진다.

3.2. I 프레임에서의 DC 프레임 추출

압축된 비디오의 I 프레임의 DC 프레임은 식(2)와 같은 방식으로 획득되고 JPEG으로 압축된 이미지의 DC 이미지를 추출하는 방식과 동일하다.

$$dc(i,j) = \frac{1}{8}\sum_{u=0}^{7}\sum_{v=0}^{7} F(i\times 8+u, j\times 8+v)\delta(u)\delta(v) \quad \delta(k) = \begin{cases} 1 & k=0 \\ 0 & k\neq 0 \end{cases} \qquad (2)$$

i 는 DC 이미지의 가로 위치, j 는 DC 이미지의 세로 위치를 의미하고, i, j 크기는 원 비디오의 화소 크기의 1/8 이다. $dc(i,j)$ 는 DCT 계수 영역에서 추출한 DC 이미지 i, j 위치의 화소 값, F 는 DCT 계수 값이다.

3.3. MPEG-1,2로 부호화된 비디오 시퀀스로부터 DC 시퀀스의 획득

P 프레임의 DC 계수 값은 I 프레임에서 획득한 DCT 계수와 움직임 벡터의 값을 이용하여 그림 5와 같은 방식으로 획득하고, 정의식은 식(3)과 같다.

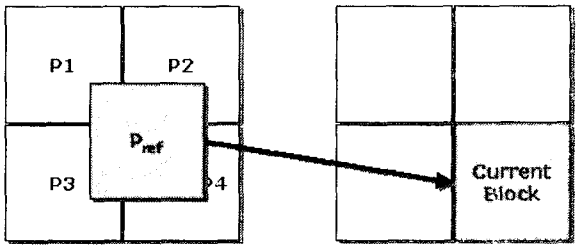


그림 5. 참조 블록과 움직임 벡터와 원 블록

$$dc(P_{REF}) = \sum_{i=1}^{4}\left\{\sum_{m=0}^{7}\sum_{l=0}^{7} w'_{ml}[DCT(P_i)]_{ml}\right\} \qquad (3)$$

이 식에서는 P_{REF} 현재 블록의 참조 블록을, w'_{ml} 는 가중치 행렬을, $[P]_{ij}$ 는 (i,j) 의 P 값을 의미한다.

식(3)에 의한 P 프레임에서의 DC-이미지의 추출은 최대 256 번의 곱셈연산을 요구한다.

$$w'_{ml} = [DCT(S_{i1})]_{0m} \times [DCT(S_{i2})]_{l0} \qquad (4)$$

S_{i1} 는 $\begin{pmatrix} 0 & 0 \\ I_{hi} & 0 \end{pmatrix}$ 혹은 $\begin{pmatrix} 0 & I_{hi} \\ 0 & 0 \end{pmatrix}$, S_{i2} 는 $\begin{pmatrix} 0 & 0 \\ I_{wi} & 0 \end{pmatrix}$ 혹은 $\begin{pmatrix} 0 & I_{wi} \\ 0 & 0 \end{pmatrix}$ 로 정의된다. 여기서 I_n 은 n 크기의 정행렬이다.

$$[DCT(P_i)]_{00} = \sum_{i=1}^{4}\frac{h_i w_i}{64}[DCT(P_i)]_{00} + c \qquad (5)$$

여기서 $c = \sum_{i=1}^{4}\left(\sum_{(m,l)\neq(0,0)} w'_{ml}[DCT(P_i)]_{ml}\right)$ 이다.

식 (5)에 DC 값을 개략적으로 추출하게 되고, 계산량은 식 (3)의 방식보다 dc 계수를 구하기 위한 연산량을 4배를 감소할 수 있고, c 는 개략적으로 추출하게 되는 값을 보상하기 위한 값이다. 대부분의 실험에서 c 의 값은 거의 0에 가깝다. [1]

4. 실험 결과 및 검토

4.1. 실험 환경 및 방식

MPEG-7 표준 비디오 데이터 베이스를 이용하여 DC 프레임의 추출 속도를 구하였고, DC 프레임 이용 방식의 유효성 검증을 위하여 CCD(Common Color Database)를 이용하였다.

헤더 정보에서 추출하는 컬러 공간 기술자와 컬러 양자화 기술자를 제외한 다섯 가지 컬러 기술자에 대해 특징 추출 시간 및 검색 율을 구하였다. 데이터 베이스는 크게 3가지로 구성되는데, 각 컬러 기술자의 성능 검증을 위한 MPEG-7 컬러 기술자 표준 데이터 베이스인 CCD(Common Color Database)와 컬러 기술자 표준 질의 이미지 데이터 베이스인 CCQ(Common Color Database)를 이용하여 특징 추출과 검색 율을 구하였다.

4.2. 압축된 비디오의 DC 프레임 추출

표 1. 비디오의 컬러 기술자 데이터 베이스의 구성

비디오 파일 명	CD 번호	총 프레임 수	비디오 압축 방식
News2.mpg	18	57948	MPEG-1 (352x240)
Misc1.mpg	20	57875	MPEG-1 (352x240)
Misc2.mpg	21	57909	MPEG-1 (352x240)
Camiloefilho.mpg	22	57017	MPEG-1 (352x240)

비디오는 연속된 이미지의 형태로서 이미지 그룹의 계층적 컬러 히스토그램을 표현함으로써 압축된 비디오의 DC 프레임의 연속된 형태를 검증할 수 있다.

표 2. DC 프레임의 추출 속도

비디오 파일 명	총 추출시간	추출시간/ 비디오 play 시간	단일 DC 프레임 추출시간
News2.mpg	1299	0.672499	0.02242
Misc1.mpg	1270	0.658315	0.02194
Misc2.mpg	1290	0.66829	0.02227
Camiloefilho.mpg	1220	0.641914	0.02139

[실험환경] CPU : 2.2GHz, RAM:512MB, HDD : 7200rpm
[추출 시간 단위] 초 (Second)

표 2의 MPEG-1 형태의 4개의 파일로부터 DC 프레임을 BMP 형태로 추출하였고 DC 프레임의 화면 비율은 크기는 가로 크기는 44 화소, 36화소로 구성된다.

표 2과 같이 3장에서 제안한 방식으로 DC 프레임의 추출 하였을 때 DC 단일 프레임을 평균 0.022초에 추출하였고 총 추출 시간은 압축 비디오의 총 재생 시간의 0.64~0.67배가 소요되었다. 즉, 압축된 비디오 파일로부터 프레임을 추출하는 시간이 실시간 재생 시간보다 축소되었다.

4.3. 이미지와 DC 이미지를 이용한 기술자의 검증과 DC 이미지 효과 분석

3장의 방식에 의해 추출된 DC 이미지를 이용한 컬러 특성의 유효성 측정을 위해 CCD의 JPEG 압축된 5420개의 이미지에서 DC 이미지를 추출하였다. DC 이미지를 이용한 컬러 특성 추출의 속도는 원 이미지를 사용하였을 때보다 1/9~1/11정도로 감소하였다. 이는 압축 영역에서 직접적으로 기술자를 추출하여 복호 과정이 생략되면서 속도가 향상되었고, DC-이미지의 생성 과정에서 이미지의 크기가 줄어들고 화소 수의 감소는 추출 시간의 향상을 가져오는 다른 원인이다.

표 3. DC 이미지와 원 이미지의 컬러 특징 추출 속도와 평균 검색율

기술자	이미지 형태	추출시간 /이미지	추출 시간 감소율(%)	평균 검색율(%)
컬러 레이아웃	DC 이미지	0.00095	82	64.931
	원 이미지	0.05338	0	66.237
컬러 구조 히스토그램	DC 이미지	0.01602	88	82.293
	원 이미지	0.13474	0	88.251
주요 컬러	DC 이미지	0.02723	99	58.337
	원 이미지	0.48787	0	56.851
계층적 컬러 히스토그램	DC 이미지	0.00967	88	76.473
	원 이미지	0.07998	0	84.056

[실험환경] CPU : 2.2GHz, RAM:512MB, HDD : 7200rpm
[데이터 베이스] CCD Database 5420 개, CCQ Database 48개
[추출 시간 단위] 초 (Second)

$$Av_Rate = \frac{1}{N}\sum_{k=1}^{N} Rate(k) = \frac{1}{N}\sum_{k=1}^{N}\frac{Correct(k)}{n(k)} \qquad (6)$$

여기서, N 은 총 질의 이미지 수, Av_Rate 는 각 기술자의 평균 검색 율, $n(k)$ 는 k 번째 질의 이미지의 검색 이미지 수, $Correct(k)$ 는 k 번째 질의 이미지의 검색 이미지 중 질의 영상과 일치하는 이미지의 수, $Rate(k)$ 는 k 번째 질의 이미지의 검색 율을 의미한다.

DC-이미지를 이용한 검색 율을 원 이미지의 경우와

비교함으로써 본 논문에서 제시한 비디오의 고속 컬러 특징 추출에 대한 결과의 검증 위해 을 표1 과 같이 평균 검색 율을 측정하였고, DC-이미지의 생성과정에서 발생하는 결과 값의 손실로 인해 평균 검색 율이 감소하였다. 식(6)의 방식으로 계산한 평균 검색 율은 주요 컬러 기술자와 컬러 레이아웃의 검색 율의 변화 정도가2%내외인데, 이는 두 기술자의 컬러 특징 추출 과정에서 평균 값을 취하는 연산이 수행되기 때문이다. 반면 컬러 구조 히스토그램과 계층적 컬러 히스토그램은 평균 검색 율이 5% 이상 감소하였지만, 각각 82%, 76%의 평균 검색 율로 컬러 기술자의 적용에 따른 문제점은 발생하지 않았다. 그림 6, 7은 CCQ의 9번째 질의 영상의 질의 결과로서 원 이미지 데이터 베이스를 사용한 경우는 7개의 질의 결과 모두가 정확하였고, DC 이미지를 이용한 데이터 베이스를 사용한 경우는 하나의 질의 결과가 오류를 나타내었다.

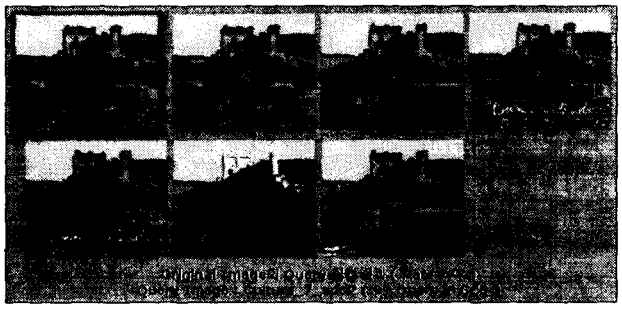


그림 6. 원 이미지 데이터 베이스에서 질의 실행 결과 예 (계층적 컬러 히스토그램 기술자)

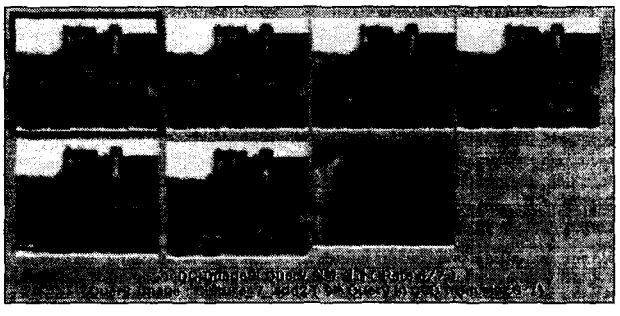


그림 7. DC 이미지 데이터 베이스에서 질의 실행 결과 예(계층적 컬러 히스토그램 기술자)

5. 결 론

MPEG-7 컬러 기술자는 이미지 및 비디오의 컬러 특징의 표현을 위한 추출 알고리즘과 표현 형식을 정의하고 있고, 이미지의 경우 공간 영역 정보의 입력을, 비디오의 경우 프레임을 이미지로 획득하여 프레임의 영역 정보를 입력하여 특징을 추출한다. MPEG-1,2 방식으로 압축된 비디오에서 컬러 특징을 추출은 복호 과정에서 추출된 프레임을 이용하게 되고, 이는 비디오의 컬러 특징 추출 속도 증가의 원인이 된다. 이런 현상을 극복하기 위해 본 논문에서 제시한 디지털 방송 컨텐츠 방식인 MPEG-2로 압축된 비디오 영역에서 컬러 기술자를 추출하는 방법으로 컬러 특징 추출 시간이 향상된 반면, 정확성은 크게 감소하지 않았다. 이러한 압축된 비디오 영역에서 컬러 기술자의 추출 방식은 디지털 방송용 비디오의 색인이나 요약 생성을 위한 연산 시간의 축소는 물론 디지털 방송용 비디오의 압축 영역에서 컬러 이외의 MPEG-7 비주얼 기술자의 적용을 위한 방법으로 활용이 가능할 것이다.

Acknowledgement

본 논문은 한국전자통신연구원 "지능형 에이전트 및 메타데이터 관리 기술연구"에 관한 공동연구과제 수행의 일환으로 얻어진 연구결과 입니다.

References

[1] Liu, Bede, Rapid scene analysis on compressed video, IEEE transactions on circuits and systems for video technology : a publication of the Circuits and Systems Society,v.5 no.6,1995,pp.533-544

[2] Manjunath, B.S.; Ohm, J.-R.; Vasudevan, V.V.; Yamada, A., Color and texture descriptors, Circuits and Systems for Video Technology, IEEE Transactions on , Volume: 11 Issue: 6 , June 2001 pp. 703 -715

[3] MPEG-7, "Text of ISO/IEC 15938-3/FDIS Information technology - Multimedia content description interface - Part 3 Visual", ISO/IEC JTC1/SC29/WG11/N4358, July 2001, Sydney

[4] MPEG-7, "Text of ISO/IEC TR 15938-8 (Extraction and Use of MPEG-7 Descriptions)", *ISO/IEC JTC 1/SC 29/WG 11 N4918, July 2002*, Klagenfurt

[5] MPEG-7, "Common Datasets and Queries in MPEG-7 Color Core Experiments", ISO/IEC JTC1/SC29/WG11, M5060, October 1999, Melbourne.

[6]MPEG-7,"Group-of-Frames/Group-of-Pictures Histogram Descriptor", ISO/IEC JTC 1/SC 29/WG 11 ,M4808,July 1999, Vancouver