

MPEG-7 기반의 Digital Image Signature 개발

*오원근 **최경호

*한국전자통신연구원 **목포대학교

*owg@etri.re.kr khchoi@mokpo.ac.kr

Development of A Digital Image Signature Based-on MPEG-7 Descriptors

*Oh, Weon-Geun **Choi, Kyoung-ho

*Electronics and Telecommunications Research Institute **Mokpo National University

요약

본 논문에서는 MPEG-7 비주얼 디스크립터를 기반으로 Digital Image의 효과적인 검색이 가능한 시스템의 개발하였다. MPEG-7에 포함되어 있는 비주얼 디스크립터 틀은 컬러, 텍스처, shape, motion, localization, 얼굴 인식 등을 포함한다. 이들 MPEG-7에서 제공하는 비주얼 디스크립터를 그대로 이용하여 Digital Image의 검색 시스템을 구현하기에는 시스템이 불필요하게 커질 수 있으며 Digital Image의 검색 성능이 그다지 높지 않다는 문제점이 발생한다. 구체적으로는 모든 디스크립터를 이용하여 데이터베이스에 존재하는 모든 Digital Image에 대한 검색을 수행하기에는 많은 처리시간이 요구된다는 것과 어떠한 디스크립터를 이용해야 정확한 검색이 이루어질지 알 수 없기 때문이다. 이를 위해 본 논문에서는, MPEG-7 비주얼 디스크립터의 특성을 저작권위원회에서 제공받은 데이터베이스를 이용하여 분석하고 이들 디스크립터의 효과적인 결합 기술을 개발하였다. 기존의 디스크립터 결합 방식은 각각의 디스크립터에 동일한 가중치를 부여하고 검색을 수행하는 방식이었으나 본 논문에서는 정보이론을 기반으로 디스크립터의 가중치를 자동으로 부여하는 방식으로 검색 시스템을 구성하였다. 개발된 시스템은 기존의 동일한 가중치를 부여한 시스템에 비해서 데이터베이스에 대한 각 디스크립터의 특성을 반영하여 가중치를 결정하도록 구성하였다.

1. 서론

디지털 콘텐츠에 대한 수요가 증가하면서 우리 주변에는 방대한 양의 멀티미디어 콘텐츠(비디오, 음악, 영상 등)가 끊임 없이 생성, 제작, 유통 및 서비스 되고 있다. 그 중에서도 디지털 Image는, 휴대용 디지털 카메라의 보급과 대용량 저장장치나 휴대용 저장매체의 가격 하락과 함께 가히 폭발적으로 그 수가 증가하고 있다. 또, 새로운 저작물이나 기존의 저작물을 이용하여 제작된 저작물도 점차 그 수가 증가하고 있다. 그러나 이러한 저작물은 고의 혹은 비고의적으로 복사(혹은 모사)되거나, 인터넷이나 재생장치의 종류에 따라 본래의 특성(영상의 크기, 형태나 품질 등)이 변하여 유통되는 경우가 발생하고 있는데 이러한 유통과정을 미연에 방지하는 것이 불법 저작물의 유통 방지 기술이고 여기에 필요한 것이 Digital Image Signature 기술이다.

Digital Image Signature 기술은, 기본적으로 Image가 가지고 있는 고유 영상 특징정보(색상, 텍스처, Image내 객체의 형상정보 등)를 표현하는 기술인데 MPEG-7 Visual Descriptor는 그중 가장 대표적인 기술이고 많은 분야에서 활용되어지고 있다. 본 논문에서는, 10,000여 종의 Digital Image UCC 저작물에 대해 상기 MPEG-7 Visual Descriptor를 적용, 통계 및 분석 작업을 하여 저작물간의 상호 식별 및 유사성을 측정하기 위한 최적의 Descriptor(혹은 Descriptor군)를

개발하였으며, 이를 기반으로 기존에 등록되어 있는 저작물의 검색 및 새로운 저작물의 등록을 효과적으로 가능하게 하였다.

2. 비주얼 디스크립터 결합 기법

MPEG-7에서 제공되는 컬러 디스크립터는 Color Space, Color Quantization, Dominant Colors, Scalable Color, Color Layout, Color-Structure, and GoF/GoP Color 등 7개로 구성되어 있으며, 텍스처 디스크립터로 Homogeneous texture, Edge histogram, Texture browsing 등 3개의 디스크립터, Shape descriptor로 Region shape, Contour shape, Shape 3D가 있다. 이외에도 Motion 디스크립터로 Camera motion, Motion trajectory, Parametric motion, Motion activity가 있으며 Localization 디스크립터로 Region locator, Spatio-temporal locator 가 있다. 이 절에서는, 기존의 연구내용 분석을 통해 보다 우수한 디스크립터 결합 전략을 세우고자 한다.

가장 간단한 방법은 각각의 디스크립터에 동일한 가중치를 주거나 사용자의 피드백을 통해 디스크립터에 사용자가 지정한 가중치를 부여하는 방법일 것이다. 이 논문에서는 MPEG-7 기술자의 결합을 위한 기법으로 두가지 기법을 제안되었다: 1) 가중치를 확률로 변환해서 반영하는 방법, 2) 가중치를 각 비교 연산 결과값의 중요도로 간주하는

방법. 제안된 방법의 개략적인 구성도는 그림 2-1과 같다. 각각의 디스크립터에서 계산된 값은 정규화 과정을 거치게 되며 정규화된 각각의 디스크립터 결과 값은 지정된 웨이트와 곱해지고 최종적으로 결합되게 된다.

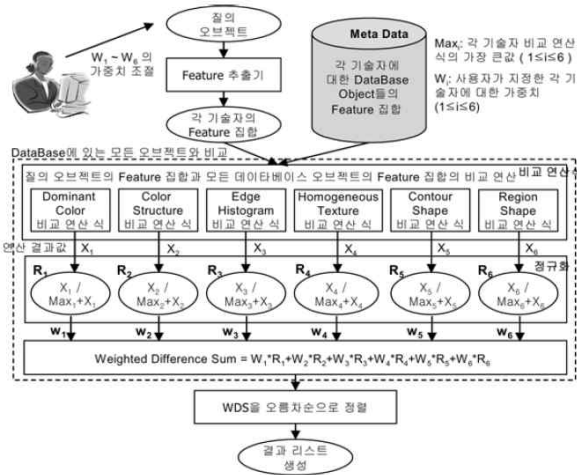


그림 2-1 디스크립터 가중치 결합을 통한 검색 알고리즘 구성도

3. MPEG-7 기반의 Digital Image Signature

3.1 시스템의 구성

MPEG-7의 기본클래스는 크게 2개로 나눌 수가 있다. Feature Data들이 저장소인 Descriptor, 그리고 Descriptor들을 이용해서 작업하는 Tools로 나눌 수 있다. Tools는 입력영상에 대해 feature를 추출하는 extract와 query와 reference의 distance를 구하는 search, 그리고 파일저장 및 불러오기를 위한 CodeSheme 등 3개로 다시 나누어진 다. Descriptor와 Tools 사이에는 Interface를 통해서 데이터 접근을 이루어지도록 되어있다.

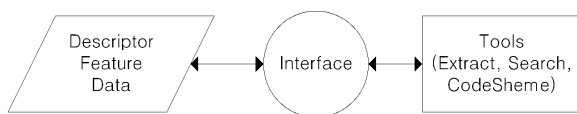


그림 3-1 클래스 내부 인터페이스 구조

MPEG-7의 기본적인 클래스 상속관계는 다음과 같다.

- 디스크립터
DscrtBase -> Descriptor -> 각각Descriptor
- 디스크립터 툴
DscrtBase -> CExtractor -> 각각Extractor
DscrtBase -> CSearch -> 각각Search

DscrtBase클래스는 멤버함수 GetDescriptorType()과 멤버변수 myType을 가지고 있는데, 이는 최종 디스크립터 클래스의 생성자에서 초기화해주고, Extractor, Search 그리고 데이터 저장 및 읽기에 있어서 디스크립터 데이터가 자신에게 맞는 데이터인지 비교할 때 사용한다.

3.2 영상 관련 구조체 MomVop

MomVop은 MPEG-7에서 사용하는 영상 읽기용 구조체이다. BGR Array를 각각의 컬러채널 R채널 G채널 B채널로 나눠서 데이터를 담고 있는 구조체이다. 먼저 Descriptor의 사용을 위해서 영상을 MomVop라는 구조체를 불러들여서 사용한다. 여기서는 OpenCV라이브러리를 사용하였다.

3.3. sort의 2가지 방식 (순위방식과 distance방식)

Query 영상에 대해 디스크립터들을 이용해서 특징을 추출한 후에 비슷한 영상을 찾기 위해서 각각의 DB와 비교한 후에 비교 결과를 합산하여 정렬하여 비슷한 영상을 찾는다. 정렬을 위한 값을 순위방식과 distance방식으로 나누어 사용한다. distance는 search tool을 이용해서 얻을 수 있는 float형 값이고, 순위는 각 디스크립터들의 distacne를 이용해서 sort한 결과이다. 즉 순위 방식은 distance의 합산한 후에 sort를 하는 것이 아니고, 디스크립터들 마다 distance에 따라 순위를 정하고 정해진 순위를 더해서 다시 정렬하는 방식이다.

- 순위 방식 : Descriptor -> Search -> Distance에 따른 정렬 -> 각 DB 파일마다 순위합 -> 최종순위로 유사도 판별
- Distance 방식 : Descriptor -> Search -> Distance합 -> 정렬하여 유사도 판별.

이러한 방식의 차이는 정규화 방식의 차이에 있다고 할 수 있다. 각 디스크립터들의 distance는 같은 파일이어도 다르기 마련이기 때문에, distance를 구하고 0~1.0사이 값으로 정규화를 하게 되고, 이 값을 합하여 distance방식을 취하게 된다. 하지만 순위방식은 구해진 distance값을 정규화를 통하지 않고 디스크립터 자체에서 sort를 하면서 1~DB개수의 값으로 정규화가 된다. distance와 순위 방식의 가장 큰 차이점은 군집화에서 나타나게 된다. 추가로 ContourShape 디스크립터는 query와 reference의 distance 계산에 있어 예외 처리가 되도록 되어있는데, 이 때 반환 값은 FLT_MAX 즉 float형이 지원하지는 최댓값이 된다. 그렇기 때문에 ContourShape의 반환값이 FLT_MAX인 reference는 영상의 정렬 목록에서 제외시키도록 했다.

3.4. I-Checker_cmd

I-Checker_cmd의 경우는 실제로 영상을 비교하는 프로그램은 아니다. 영상이 담겨져 있는 폴더 내에서 실행을 하게 되면 Feature폴더를 생성하고 폴더 안에 각 디스크립터들의 DB를 생성하는 프로그램이다. 단순히 영상이 담겨져 있는 폴더에서 실행만 하면 되는 프로그램이다.

4. 실험영상의 분석

본 논문에서는 실험을 하기 전에 대상 영상들의 구조를 분석하여, 분석결과를 개발된 기술의 적용가능성과 세부 Descriptor를 조합, 적용하는 기준으로 활용하였다. 총 분석영상은, 2,340개로 아래의 세 가지로 분류할 수 있다.

4.1.1 실험영상의 종류

실험영상은 크게 다음과 같이 세 종류로 분류할 수 있다.



a) 배경과 객체를 쉽게 분리할 수 있는 경우



b) 객체 판별이 쉬운 도면, 도안 DB들



c) 객체 판별이 쉬우면서 도면 도안이 아닌 DB들
그림 4.1 실험영상의 종류

4.1.2 배경이 복잡한 경우 vs 배경이 단순한 경우

가장 중요한 선별요소라고 할 수 있을 듯하다. Color디스크립터의 경우 색상에 따라 그룹화를 하거나 카운트를 하는데, 배경이 객체에 대해 영향을 미치게 되는 경우가 되어 선별을 해야 하며, Shape관련 Descriptor들도 세그먼트가 되어있는 객체에 대해서 뛰어난 Description을 보여주기 때문이다.

4.1.3 도면, 도안 vs 그 외의 객체

건축물 등의 도면의 경우 판별해야할 객체 외에 주변 환경적인 요인들도 같이 그려진 경우가 대부분이기 때문에 descriptor들의 오류를 높이는 요인을 갖추고 있고, letter등의 도면들은 실제 판별해야하는 객체 부분이 글씨와 함께 있는 경우가 대부분이었고 이는 descriptor의 성능 저하를 가져오는 요인이라 판별되어 일괄적으로 분리를 해주었다.

4.1.4 문제가 될 수 있는 DB 영상들

- 영상 하나에 여러 개의 객체가 존재 : 아래와 같이 2가지 타입의 입력 영상이 있다. 왼쪽의 경우에는 1개의 객체에 대해서 여러 가지 모

습이 담겨 있는 것이고, 오른쪽 영상은 단하나의 경우에 대해서만 descriptor에 대해서 description이 될 것이다. 하지만 왼쪽의 영상은 자체로써 descriptor에 대해 description될 것이고, 이는 한 장씩의 영상에 대해 올바른 결과를 가져오기 힘들게 된다. 이 경우에는 기존의 DB에 대해 개개의 분할된 영상을 실험영상으로 추가하였다.



그림 4.2 여러 개의 객체를 가지고 있는 그림

- 대상 객체 배치에 문제가 있는 영상 : 아래와 같은 영상은 실제 객체 부분보다 입력된 배경이 넓은 경우이다. 특히 Color 디스크립터의 경우에는 흰색에 영향이 가장 크게 미치게 된다. 판별해야할 객체 부분을 따로 추출해서 추가하였다.

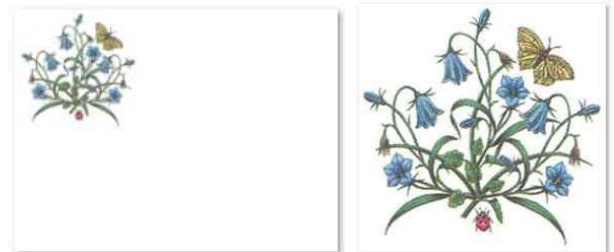


그림 4.3 객체의 위치가 중심이 아닌 경우와 처리 결과

4.1.5 실험 영상 DB 혹은 Query 영상에 대한 분석 결과

각 디스크립터들의 성능을 높이기 위해서는 객체만 있는 영상이 필요하다. 또한 각 디스크립터들은 alpha channel을 지원을 하는데, Color 디스크립터들의 경우 알파채널(투명채널)의 부분은 배제를 하고 특징을 추출한다. Shape 디스크립터들의 경우에는 알파채널 자체를 이용해서 description을 수행하는데 이는 알파 채널 외의 부분을 객체로 판단하고 outline의 shape을 이용하기 때문이다.

5. 실험결과

실험에 사용된 컴퓨터 사양은 다음과 같다.

- OS : Windows7 Professional 64Bit
- CPU : Intel i5 750 2.67Ghz
- RAM : 4GB

5.1 디스크립터 추출시간 측정

특정 추출시간 측정은 5가지 영상(RGB모델)에 대해 30회씩 진행되어 평균을 구하였다.

- 영상 로딩시간

	테스트영상 1	테스트영상 2	테스트영상 3	테스트영상 4	테스트영상 5
해상도	517 x 656	1441 x 1338	2332 x 2671	3000 x 3468	4508 x 4488
로딩 시간(sec)	0.0255	0.0527	0.1364	0.2737	1.1722

5번 테스트 영상의 경우 영상의 해상도도 높으며 32bit의 CMYK 영상이었기 때문에 RGB모델로의 변환과정이 추가되기 때문에 영상 로딩시간이 상당히 소요된다.

- 디스크립터 추출 시간

	테스트 영상 1	테스트 영상 2	테스트 영상 3	테스트 영상 4	테스트 영상 5	
Color	Layout	0.0088	0.0081	0.0082	0.0084	0.0080
	Dominant	0.3427	0.3068	0.3288	0.3318	0.3225
	Structure	0.0462	0.0427	0.0418	0.0417	0.0561
Texture	Homogeneous	0.0245	0.0249	0.0244	0.0246	0.0243
	Edge	0.0032	0.0029	0.0030	0.0030	0.0027
Shape	Region	0.0274	0.0206	0.0086	0.0129	0.0171
	Contour	0.0032	0.0029	0.0029	0.0030	0.0026

디스크립터의 추출시간은 영상의 크기와 크게 비례하지는 않음을 알 수 있다. 두드러진 특징은 Dominant Color 디스크립터의 추출시간이 다른 디스크립터에 비해서 굉장히 길다는 점이라고 할 수 있다.

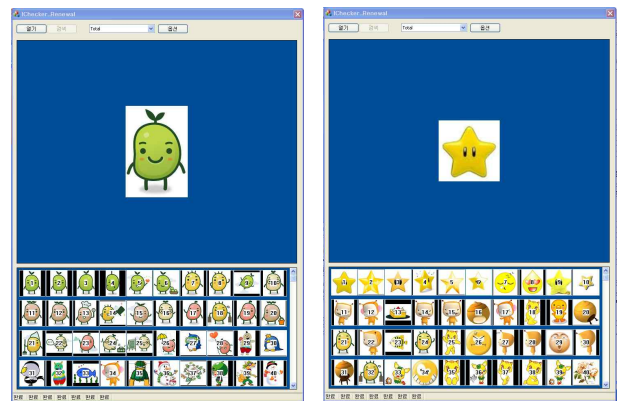
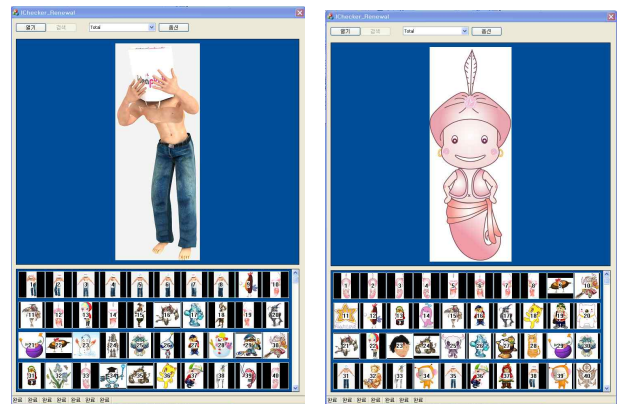
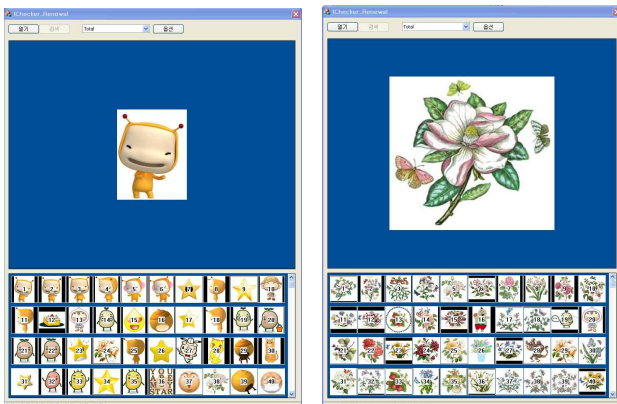
- 영상 검색 시간

		avg	min	max
Color	Layout	0.0049	0.0047	0.0055
	Dominant	0.0216	0.0170	0.0253
	Structure	0.0077	0.0072	0.0085
Texture	Homogeneous	0.0069	0.0066	0.0077
	Edge	0.0052	0.0050	0.0058
Shape	Region	0.0045	0.0043	0.0051
	Contour	5.7415	5.6443	5.9503

여기서는 객체 판별이 쉬우면서 도면 도안이 아닌 DB들로 선별된 1024개 DB에 대한 30개의 Query 영상의 검색 시간을 측정하였다. Contour의 경우 검색시간에서 많은 시간이 소요됨을 알 수 있다. 테스트 PC의 사양이 높고, 커맨드 모드의 간략화된 모듈 상태, DB개수가 증가되는 상황을 고려할 때, Indexing 기법을 적용하지 않고 순차적인 검색 방법으로는 한계가 있음을 알 수 있다.

5.2 검색결과

아래 그림은 선정된 질의영상에 대한 검색결과 예시 보여준다.



6. 결론

MPEG-7 디스크립터는 Digital Image 에 포함되어 있는 컬러, 텍스트, 형태 정보를 기술하기 위해 최적화 되어 있는 국제 표준 디스크립터이다. 따라서 MPEG-7에서 제공된 디스크립터를 활용하되 최대의 검색 효과를 제공할 수 있는 디스크립터의 효과적인 활용에 초점을 맞추어 시스템을 개발할 필요가 있다.

이를 위해, 본 논문에서는 MPEG-7 비주얼 디스크립터의 특성을 저작권위원회에서 제공받은 데이터베이스를 이용하여 분석하고 이들 디스크립터의 효과적인 결합 기술을 개발하고, 기존의 디스크립터 결합 방식은 각각의 디스크립터에 동일한 가중치를 부여하고 검색을 수행하는 방식이나 유전자 이론을 이용하여 가중치를 결정하는 방식을 소개 하였다. 또, 정보이론을 기반으로 디스크립터의 가중치를 자동으로 부여하는 방식으로 검색 시스템을 구성하였다. 개발된 시스템은 기존의 동일한 가중치를 부여한 시스템에 비해서 데이터베이스에 대한 각 디스크립터의 해당 데이터베이스에 대한 특성을 반영하여 가중치를 결정하도록 구성하였다.

참고문헌

- 1) The MPEG Home Page(<http://mpeg.chiariglione.org/>)

본 연구는 지식경제부, 문화체육관광부 및 정보통신연구진흥원의 IT산업원천 기술개발사업의 일환으로 수행하였음. [2011-KI001797, Rich UCC 기술 개발]