

멀티미디어 콘텐츠 관련 표준 「MPEG-7」 동화상 검색에 필수적인 국제표준, 넘치는 화상데이터를 정리·저장

저자·堀 修
東芝(도시바)개발연구센터
멀티미디어연구실 주임연구원
역자·정지은
출처·NIKKEI COMMUNICATIONS
2001년 12월

고속·광대역 통신을 이용하여 디지털화된 멀티미디어 콘텐츠를 쉽게 시청할 수 있는 시대가 되고 있다. 그러나 이와 달리 인터넷상의 Web 페이지는 키워드를 이용하여 검색·입수하는 것이 정비되어 있지 않다. broadband 시대의 동화상 검색에 필수적인 「MPEG-7」에 대해 설명한다.

ADSL(asymmetric digital subscriber line)이나 FTTH(fiber to the home)라고 하는 broadband 통신 서비스의 등장에 따라 동화상 등의 멀티미디어 콘텐츠가 급증하고 있다.

ADSL은 최대 1.5Mbit/초 또는 최대 8Mbit/초 대역에서 디지털 영상의 streaming 서비스를 이용할 수 있다. 실질적으로 4Mbit/초 정도의 대역만 있으면 보통 TV 방송정도의 영상송신이 가능하다. 미래 FTTH 서비스가 보급되면 하이비전 및 고화질 영상송신도 불가능하지만은 않다.

방송분야의 디지털화도 급속도로 진행되고 있다. 이미 수백 채널을 이용할 수 있는 CS 방송, 2000년 12월에 본 방송이 시작된 BS 디지털방송, BS 아날로그방송용 위성과 같은 궤도상의 위성을 사용하여 2002년 봄에 시작되는 차기 CS 방송이나 2~3년 후에 시작되는 지상파 디지털방송 등이 그 예다.

콘텐츠 제공사업자가 준비하는 멀티미디어 콘텐츠의 양이 향후 점점 증가할 것은 확실하

다. 더욱이 디지털 카메라나 디지털 비디오가 일반소비자에게 보급됨은 개인이 보유하는 디지털 콘텐츠의 양을 비약적으로 증대시키고 있다. 하드디스크나 DVD 등, 저가의 대용량 디스크 미디어의 등장도 이 디지털 콘텐츠의 증가에 박차를 가하고 있다.

어디에 어떠한 영상이 있는지 알 수 없다.

멀티미디어 콘텐츠의 수가 방대해지면 우선 사용자에게 있어서 콘텐츠를 탐색할 수 없는 문제가 생긴다. 한정된 시청시간 안에 자신이 좋아하는 콘텐츠만을 효율있게 선정하여 보고싶어하는 요구가 높아지고 있다.

한편, 콘텐츠 제작측에 있어서는 제작비용 증가라는 심각한 문제가 있다. 방송서비스가 지상파만 있었던 시대는 고비용이라도 우수한 영상 콘텐츠를 제작하면 멀티미디어에 맞추어 광고수입을 얻는 비즈니스 모델이 성립되었다. 그러나

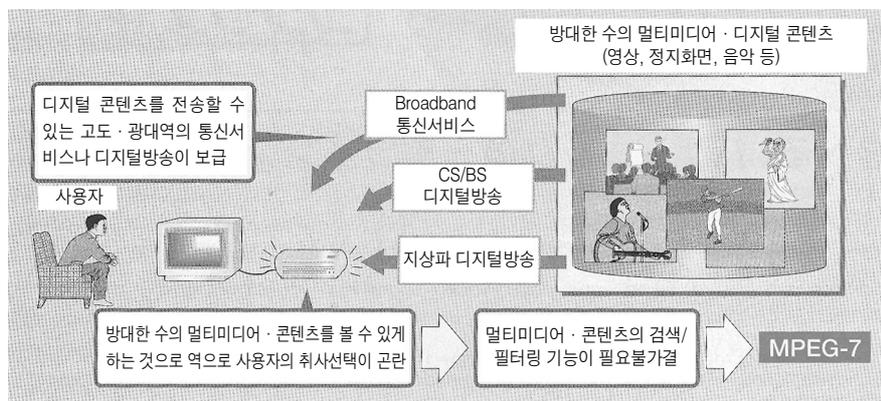


그림 1. 증가한 디지털 콘텐츠의 검색/필터링에 MPEG-7이 필요

향후 고속·광대역 통신서비스나 디지털 방송이 보급됨에 따라 대량의 디지털 콘텐츠의 축적이 진행된다. 이 대량의 콘텐츠 가운데 사용자는 원하는 콘텐츠만을 검색/필터링하고자하는 요구가 높아진다. 그 업무를 실현하는 핵심기술이 MPEG-7이다.

송신미디어가 증가함으로써 방송의 희소가치가 떨어졌다. 대량의 채널을 프로그램으로 채우기 위해서 낮은 비용으로 양질의 콘텐츠를 신속히 작성할 필요가 있다.

이러한 몇 가지 문제를 해결하기 위한 유효한 기술이 멀티미디어 콘텐츠를 기술(記述)하기 위한 표준 「MPEG-7」이다. (그림 1)

콘텐츠 검색에 과거의 소재를 재이용

싼 가격으로 양질의 콘텐츠를 확보하는 하나의 방법이 「legacy」라 하는 과거의 방송 콘텐츠 재이용으로, 「白素材」라고 하는 취재 콘텐츠를 기반으로 한 각종 다양한 프로그램의 효율적인 작성이다. 이것은 콘텐츠 검색기술을 사용하는 것으로 실현된다.

대량의 멀티미디어 콘텐츠에서 자신이 갖고 싶은 것을 찾는 데에는 검색을 위한 인덱스가 필요하게 된다. 검색대상이 문장인 경우는 문장에서 단어를 추출하여 용이하게 검색할 수 있다. 그러

나 화상, 동화상, 오디오, 음성과 같은 멀티미디어 콘텐츠가 대상인 경우는 그 상태 그대로는 검색할 수 없다.

그래서 콘텐츠의 내용을 기술한 데이터를 미리 작성하고, 검색할 수 있게 하는 것이 고안되었다. 이 내용기술 데이터를 「메타데이터」라고 한다. MPEG-7은 멀티미디어를 위한 메타데이터를 어떻게 기술하는가를 정해놓은 국제 표준이다.

영상데이터에 있어서 가장 단순한 메타데이터는 영상의 시작시간과 길이를 지정하여 그

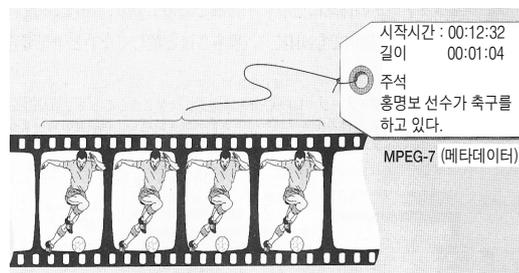


그림 2. MPEG-7의 영상용 메타데이터의 가장 단순한 예 영상 장면의 시작시간, 길이, 주석만을 붙이는 것

부분에 무엇이 반영되어 있는가를 텍스트로 기술한 것이다(그림 2). MPEG-7에서는 이러한 표현방법을 기본으로 하여 다양한 요소를 기술하도록 규격화하고 있다.

MPEG-7의 표준화는 2001년 내에 완료

MPEG-7은 1998년 10월에 요구사항에 의거한 기술제안이 모집되어 1999년 2월에 이들 제안이 테스트·평가되었다. 그 결과를 기초로 약 2년 반에 걸쳐 규격을 책정하여 2001년 10월에 FDIS(국제표준의 최종초안)가 완성되었다.

향후 각 국의 최종승인투표를 거쳐 2001년 말까지는 표준내용이 확정될 전망이다. 빠르면 2002년 여름 경에 국제표준(ISO/IEC)으로 출판된다.

새로운 검색 비즈니스가 탄생된다.

멀티미디어 정보의 범람은 사용자나 콘텐츠 제공사업자에게 있어서 어려운 문제라 할 수 있다. 그러나 그에 따라 새로운 비즈니스 기회가 생기는 것도 확실하다.

현재 인터넷상에는 무수히 존재하는 Web 페이지를 검색하는 검색 엔진이 있다. 서비스 사이트를 방문하는 사용자에게 보여지는 광고를 수입원으로 검색을 비즈니스로 하고 있는 기업도 많다. broadband network가 보급되는 가까운 장래에는 화상데이터까지를 대상으로 한 검색서비스가 등장하는 것은 거의 확실하다. 이때 MPEG-7은 멀티미디어 데이터 검색을 실현하는 중요한 기술이 된다.

MPEG-7의 실용화에는 몇 가지 과제가 있다. 그 하나가 영상데이터에서 MPEG-7의 메타데이터를 부가하는 작업에 시간이 걸리는 것이다. 이것이 영상콘텐츠의 제작비용을 높이는 요인이 된다.

그렇지만 미국의 방송 프로그램에는 closed caption을 부가하는 것이 의무화되어 있어 이 텍스트 데이터를 메타데이터 작성에 활용하는 것은 비교적 용이하다. 일본 내에서도 NHK 등이 독자 포맷의 메타데이터 작성을 진행하고 있어 MPEG-7에의 변환도 그다지 어렵지 않으리라 예상된다. 또, 음성인식기술을 사용하여 메타데이터의 생성을 자동화할 수 있는 가능성도 있다.

이러한 것들을 사용하면 비교적 저렴한 가격에 MPEG-7의 메타데이터를 작성하는 환경이 점차 정비되어 갈 것이다.

좋아하는 장면, 검색가능

그럼 MPEG-7의 메타데이터를 사용하면 실제 어떠한 영상검색을 실현할 수 있는가? (그림 3)

예를 들면 1주간의 뉴스 프로그램 영상을 MPEG-4로 디스크에 저장하는 경우를 생각해 본다. MPEG-7의 메타데이터에서는 뉴스 영상의 주제마다 시작시간, 길이 내용이 기록되어 있다. 그래서 삼성의 이승엽 선수에 관한 영상을 보고 싶은 시청자가 개인 단말에서 「이승엽&홈런」이라는 키워드를 입력한다. 그러면 검색용 애플리케이션이 데이터베이스에 접속하여 메타데이터를 처리하여 입력된 키워드에 관련된 영상데이터의 시작시간과 길이 등의 정보를 얻는다.

더욱이 검색용 애플리케이션은 그 영상데이터의 정보를 표시용 애플리케이션으로 보내어 해당하는 영상 부분만을 시청자의 단말기로 연결 제

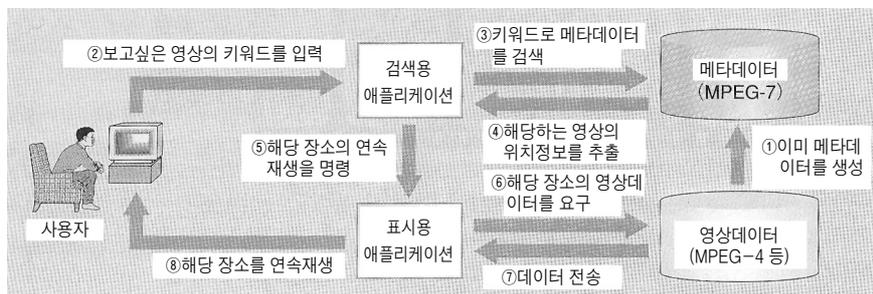


그림 3. MPEG-7의 실제 이용 Image

MPEG-4의 영상 데이터 등으로부터 미리 MPEG-7의 메타데이터를 생성, 입력된 검색 키워드로 메타데이터를 처리하고, 시청자가 희망하는 영상의 위치정보를 얻는다. 이 위치정보를 사용하여 요약된 영상데이터를 재생한다.

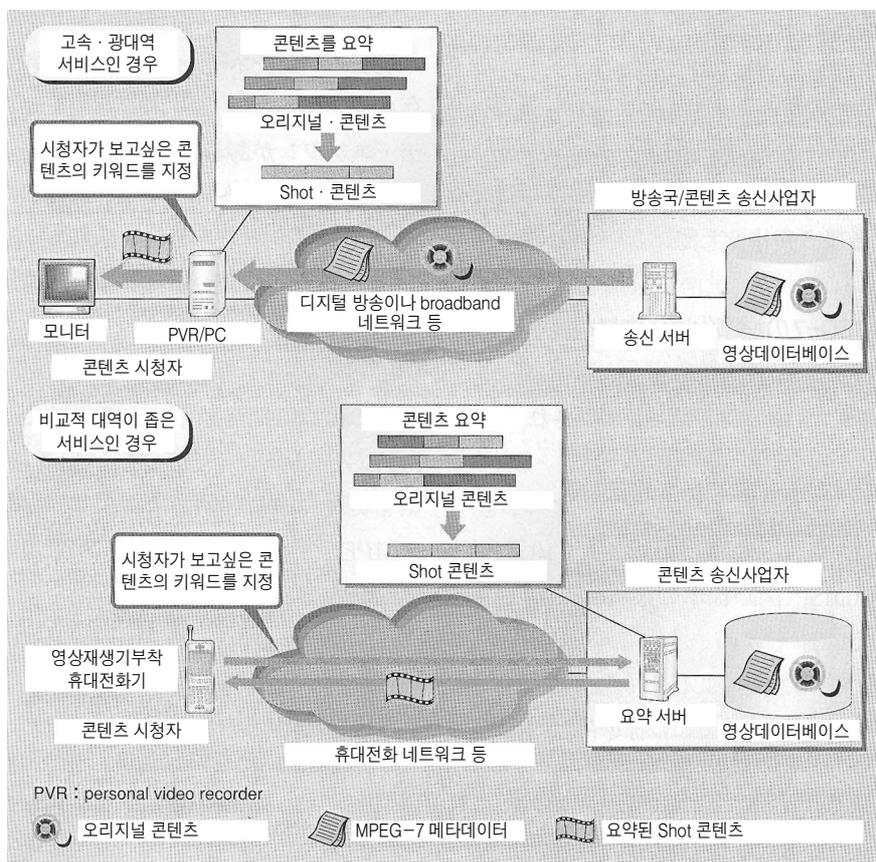


그림 4. MPEG-7의 메타데이터를 사용하여 시청자의 호응에 따른 요약영상을 생성

고속·광역 서비스의 경우는 방송국이나 콘텐츠 송신사업자가 original contents와 MPEG-7 메타데이터 양 쪽을 PVR(personal video recorder)로 송신. 시청자 자신이 요약된 shot contents를 생성한다. 한편 휴대전화 등 비교적 대역이 좁은 서비스에서는 시청자의 반응을 콘텐츠 송신사업자에게 송신하고 요약 서버가 shot contents를 생성한다.

생한다. 이러한 절차에 따라 시청자는 이승엽 선수의 홈런 장면을 포함한 영상만을 1주간 단위로 정리하여 시청할 수 있게 된다.

VTR 후속인 PVR 등을 휴대용으로

구체적인 이용형태로서는 personal video recorder (PVR)를 사용한 영상검색을 들 수 있다(그림 4).

PVR이란, post VTR 이라고도 하는 차세대 디지털 영상재생기이다. 대용량 하드디스크 등을 사용하여 디지털 영상을 장시간에 걸쳐 기록하여 두고, 나중에 좋은 시간에 좋은 프로그램은 보게 하는 새로운 시청 스타일을 실현하는 기기이다.

그러나 기록되는 영상시간이 방대하기 때문에 자신이 흥미있는 부분만을 걸러낼 필요가 있다. 특히 정보 관련 영상콘텐츠인 뉴스 프로그램이나 스포츠 프로그램은 중요한 포인트만을 부분적으로 시청하고 싶어하는 요구가 커지고 있다. 예를 들면 스모 씨름에서 승부를 결정짓는 부분이나 축구의 골인 장면만을 연속하여 보고싶어하는 요구이다. 또 보고싶은 장소를 찾으면서 효율 좋게 임의로 시청하고자 하는 요구도 높아진다. 이와 같은 네비게이션이나 브라우저를 지원하는 tools을 MPEG-7으로 실현할 수 있다.

특히 이와 같은 PVR에 의해 콘텐츠의 메타데이터가 제조업체의 다름을 초월하여 이용가능하다는 것이 중요하며, 따라서 국제표준인 MPEG-7의 의미는 크다. 예를 들면 미국의 TV Anytime이라는 포럼은 MPEG-7 기술을 기본으로 하여 PVR을 위한 메타데이터 framework의 규격화를 진행하고 있다.

또, 영상수신이 가능한 휴대전화로 접속하여 요약된 영상을 온라인으로 제공하는 서비스도 유망할 것이다. 언제 어디서나 정보를 접속할 수 있는 휴대전화는 선도 높은 뉴스나 스포츠와 같은 영상 콘텐츠에 다가가고 있다.

그렇지만 휴대전화로 영상데이터를 계속 수신하는 것은 광대역이나 비교적 높은 통신요금으로 보아 상상하기 어렵다. MPEG-7을 사용하여 이용자가 검색한 영상만을 수신하는 서비스라면 받아들여질 가능성이 높다.

영상압축기술이 MPEG-1/2/4와는 별개

지금부터는 MPEG-7의 사양에 대하여 상세히 살펴보기로 한다.

ISO/IEC의 하부조직인 MPEG이 책정한 표준으로서의 영상압축기술인 MPEG-1, MPEG-2, MPEG-4가 유명하다. 이들과 같은 이름을 사용하고 있기 때문에 MPEG-7도 일종의 영상압축 기술이라는 생각을 갖기 쉽다. 그러나 MPEG-7은 멀티미디어의 내용기술 방법을 표준화하는 것이다.

MPEG도 5번과 6번을 결번으로 하는 것으로 MPEG-7이 지금까지의 표준과 다르다는 것을 강조한 것으로 알려지고 있다.

단, 내용기술이 영상의 에센스를 함축한 데이터라고 생각하면 어떤 종류의 압축이다라고 생각해도 틀리지 않다.

표준화의 대상영역도 MPEG-1, 2, 4와 MPEG-7과는 다르다(그림 5). MPEG-1, 2, 4에서는 압축 부호화된 binary format과 그것을 부호화하는 decoder가 표준화대상이다.

한편 MPEG-7은 XML을 사용하여 메타데이터를 기술하는 표준규격이다. 표현형식을 정의하

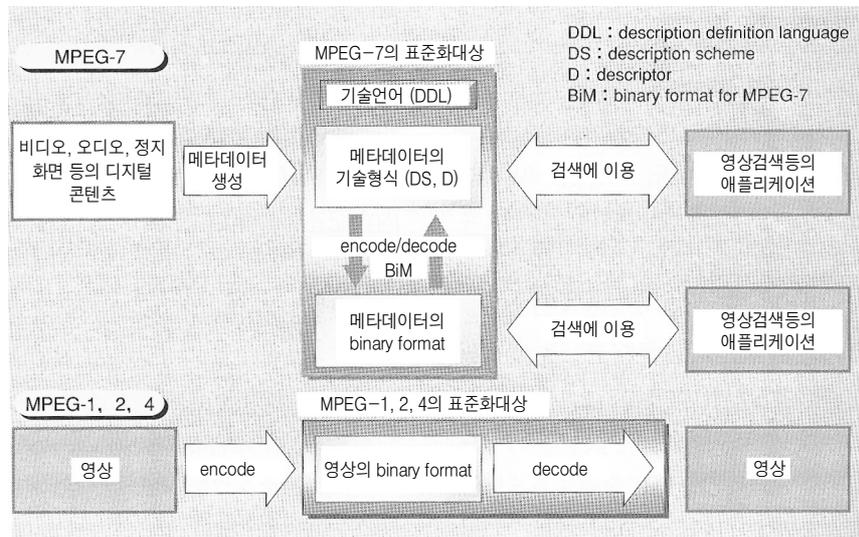


그림 5. MPEG-7과 MPEG-1, 2, 4의 표준화대상 영역의 차이

MPEG-7에서는 기술언어, 메타데이터의 기술형식, 메타데이터를 binary format으로 변환하기 위한 encode/decode 방식이 표준화 대상. 영상을 압축한 binary format과 decode 방식을 규정하는 MPEG-1, 2, 4와는 크게 다르다.

는 언어인 「DDL」(description definition language)과 메타데이터의 표준형식인 「DS」(description scheme), 그 요소인 「D」(descriptor)가 표준화대상이다. DDL의 기술은 텍스트 데이터이지만 그것을 효율 좋게 전송·저장하기 위한 압축방식으로서 「BiM」(binary format for MPEG-7)도 MPEG-7 표준으로 정하고 있다. 메타데이터의 생성과 이용방법은 표준의 대상 외이다.

DDL이란, MPEG-7의 내용기술을 정의하기 위한 언어이다. W3C(WWW 컨소시엄)에서 규격화된 schema 정의언어 「XML schema」를 기본으로 멀티미디어 콘텐츠의 내용기술에 필요한 데이터형을 추가, 확장한 것이다.

또 DS란, 대상을 표현하는 scheme으로 그 속성을 D로 기술한다. 예를 들면 사람을 나타내는 「Person DS」에서는 연령이나 이름 등을 D로 하여 기술한다.

서적을 모델로 콘텐츠 내용을 기술

MPEG-7은 화상·동화상·오디오·음성의 내용을 기술할 수 있지만 그 가운데서도 특히 영상의 내용기술에 관해 깊이 연구되어 설계되어 있다. 영상의 내용기술은 서적과 같은 목차나 색인, 요약 등을 포함하고 있다. 서적의 목차는 장과 절의 계층구조로서 이용자가 흥미 있는 장소를 쉽게 찾을 수 있다. MPEG-7도 같은 형태를 취하여 내용기술의 틀을 정하였다.

구체적으로는 영상의 scene이나 shot의 계층구조가 목차에, 영상 장면을 대표하는 key 프레임(1장의 정지화면)이 색인에 해당한다. 요약이나 하이라이트는 서적의 요약에 해당된다.

영상 콘텐츠를 예로, MPEG-7의 메타데이터를 설명하고자 한다. 우선 영상 콘텐츠는 그

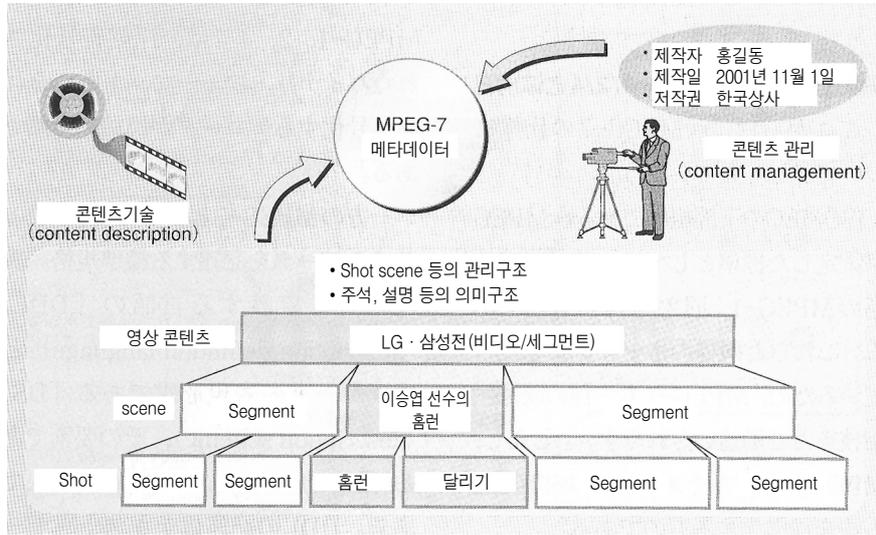


그림 6. 메타데이터의 고레벨 기술

고레벨 기술에는 「콘텐츠 관리」와 「콘텐츠 기술」이 있다. 또한, 콘텐츠 기술은 「관리구조」와 「의미구조」로 나뉜다. 관리구조는 「segment」라 불리는 영상을 분할한 shot이나 scene 등의 계층구조이다. 의미구조는 그 영상내용을 기술한 것이다.

제작일이나 제작자라고 하는 「콘텐츠 관리」 정보를 메타데이터로서 기술한다. 한편 이와는 별도로 영상의 내용을 「콘텐츠 기술」 정보로서 기술한다(그림 6)

콘텐츠 기술은 또한 관리구조와 의미구조로 나뉜다. 관리구조는 카메라로 연속 촬영한 단위의 shot이나 shot을 조합하여 작성한 scene의 계층구조이다. 의미구조는 이들 shot이나 scene의 내용설명이다. shot이나 scene과 같은 영상의 단위를 MPEG-7에서는 「segment」라 한다.

그림 6은 몇 개의 scene/shot으로 구성되어 있는 LG·삼성전의 야구중계 영상의 예이다. 이승엽 선수의 홈런 scene은 홈런의 shot과 이승엽 선수의 달리기 shot으로 구성되어 있다. 이 때문에 scene 또는 shot의 내용설명을 위해 「홈런」이나 「달리기」 등을 텍스트로 기술한다. 구체적으로는 「Text Annotation DS」라

고 하는 의미구조를 기술하기 위한 DS를 사용하여 기술한다.

텍스트 검색을 실현하는 고레벨 기술

이러한 영상 등의 내용을 주로 텍스트 데이터로 기술하는 것을 MPEG-7에서는 「고레벨 기술」이라고 한다. 텍스트에 의한 검색을 가능하게 하며, 그것을 MPEG-7에서는 「MDS」(multimedia description scheme)라 하는 경우도 있다. 고레벨 기술은 6개의 class로 구성하고, 또한 목적별로 다양한 분류로 나뉜다(표 1). 카테고리 안에는 많은 DS가 준비되어 있다.

예를 들면 「기본요소」는 MPEG-7 특유의 기본적인 표현형식을 정한다. 그 가운데 「schema tools」은 XML로 표현된 내용기술이 MPEG-7의 메타데이터인 것을 나타내기 위한 tool이다. 예를

표 1. 고레벨 기술의 분류 기술은 사용목적에 따라 분류되어 있다.

Class	Category	기술 Scheme의 명칭	설명
콘텐츠 조직화 (content organization)	집합체	Collections	복수 콘텐츠의 그룹화 tool
	Model	Models	콘텐츠 특징의 파라미터 표현 tool
콘텐츠 관리 (content management)	제작	Creation&Production	콘텐츠 제작에 관련된 정보표기 tool
	Media	Media	콘텐츠 포맷에 관한 정보표현 tool
	이용	Usage	콘텐츠의 이용에 관한 정보표현 tool
콘텐츠 기술 (content description)	관리구조	Structural aspects	콘텐츠의 물리적인 분할표기 tool
	의미구조	Semantics aspects	콘텐츠의 의미정보 표기 tool
navigation & access	요약	Summaries	영상요약을 위한 tool
	View	Views	동일 콘텐츠의 다양한 표현 tool
	파생물	Variations	동일 콘텐츠의 다양한 형태
User Interaction	사용자 호응	User preference	사용자의 호응에 관한 표기 tool
	이용이력	Usage history	사용자의 이용이력에 관한 표기 tool
기본요소 (basic elements)	Schema Tools	Schema tools	MPEG-7에 관한 XML schema 결정
	기본데이터형	Basic data types	MPEG-7 특유의 데이터형
	링크와 미디어참조	Links&Media Localization	참조 tool
	기본 Tools	Basic tools	기본적인 tool

들면 「<MPEG7>...</MPEG-7>」에서 묶여 있는 부분이 MPEG-7의 메타데이터인 것을 나타낸다 등으로 규정되어 있다.

또 「기본 데이터형」은 XML schema에서 제공되고 있지 않는 MPEG-7 특유의 데이터형을 정하고 있다. 예를 들면 8bit 등에 bit 길이를 규정한 정수 데이터형이 정의되어 있고, 텍스트 표현된 수치데이터를 BiM 으로 압축하기 위해 이용한다.

먼저 이승엽 선수의 홈런 scene을 설명한 「콘텐츠 관리」와 「콘텐츠 기술」도 고레벨 기술의 class로 정하여져 있다.

유사한 영상을 찾기 위한 저레벨 기술

고레벨 기술에서는 영상 콘텐츠의 관리구조나 내용을 기술하지만 원칙적으로 텍스트 정보로 기술하기 때문에 상세한 영상정보를 기록하는 것은 어렵다. 그래서 MPEG-7은 저레벨 기술이라는

별도의 기법을 정하고 있다. 저레벨 기술은 수치 데이터로서 기술되어, 분할된 segment에 대해 부가된다(그림 7).

저레벨 기술은 예시검색을 가능케 한다. 예시검색이란 샘플 데이터를 주어 유사한 것을 검색하는 것이다. 그림 7의 예에서는 이승엽 선수의 얼굴의 특징과 유니폼의 주요색이 「visual 기술」로서 부가되어 있다.

만약 이승엽 선수의 얼굴을 자동으로 검출할 수 있으면 수입력으로 이승엽 선수라 입력하지 않아도 검색할 수 있게 된다.

또 아나운서의 해설이나 응원의 음악 멜로디도 저레벨 기술로 기술할 수 있다. 이렇게 해 두면 아나운서의 발언이나 유사한 멜로디라도 검색이 가능케 된다.

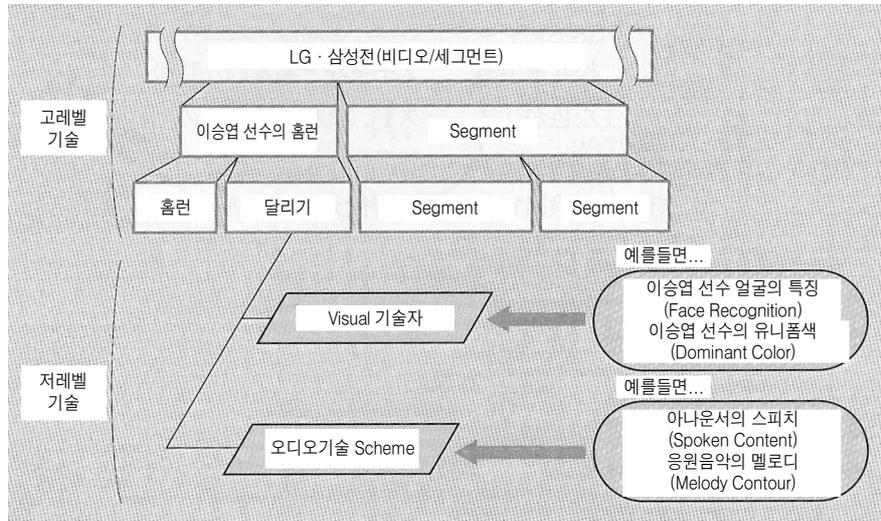


그림 7. 메타데이터의 저레벨 기술

저레벨 기술은 고레벨 기술의 segment마다에 추가하는 option 정보. 화상신호나 음향·음성신호의 특징을 나타낸다. 많은 경우는 컴퓨터 등에 의해 자동적으로 계산되어 추가된다. 「Visual 기술자」와 「Audio 기술 Scheme」 2종류가 있다.

색이나 형태는 Visual 기술로 표준

화상, 동화상, 오디오, 음성과 같은 멀티미디어 콘텐츠는 신호의 특징을 컴퓨터에서 자동 추출할 수 있다. 저레벨 기술은 주로 이와 같은 신호의 특징을 사람에 의한 입력 없이 메타데이터로서 기술한다.

화상 및 동화상의 특징은 「Visual 기술」을 사용하여 기술한다. 구체적으로는 ① 화상이나 동화상의 색, 형태, 텍스처 등 구체적인 특징을 나타내는 「class 1」, ② 얼굴이나 카메라의 움직임 등 고도의 특징을 나타내는 「class 2」, ③ class 1, 2의 특징이 어느 장소에 있는가를 나타내는 「class 3」, ④ 기술자를 적용하는 범위를 지정하는 「class 4」 - 이 4가지로 나뉘어 있다(표 2).

class 3은 예를 들면 축구 시합의 영상에서 어느 인물이 홍명보 선수인지를 위치정보로서

보여주는 것이 가능하다. 또 class 4에서는 화면을 격자형태로 분할하여 개개의 소영역에 기술자를 편집 배정하여 화상을 특징짓는 「Grid Layout」이나 시간계열로서 분할하여 편집배정하는 「Visual Time Series」 등이 규정되어 있다.

음의 특징을 나타내는 오디오 기술 scheme

저레벨 기술에는 visual 기술자 외에 오디오나 음성에 의한 검색을 가능케 하는 「오디오 기술 scheme」이 있다.

visual 기술자가 화상, 동화상의 기본적인 특징을 수치로 기술하는 것에 대해 오디오 기술 scheme은 검색을 위해 수치기술만이 아닌 애플리케이션으로부터의 이용을 가정한 기술 scheme을 정하고 있다.

예를 들면 악기의 음색을 검색하기 위한 「탐바

표 2. 저레벨 기술의 visual 기술자 일람

기술자는 기본적인 카테고리를 표현하는 것으로 구성되어 있다.

Class	Category	기술자 명칭	설명
1. (기본특징)	색구성비	Dominant Color	최고 빈도색의 값과 빈도
		Scalable Color	색 히스토그램의 특징
		Color structure	
	색배치	GoF/FoP Color	복수 프레임의 색 히스토그램
		Color Layout	색의 공간적 배치
	모양	Homogeneous Texture	일정모양의 텍스처
		Texture Browsing	텍스처의 언어표현
	Edge	Edge Histogram	edge의 경사 분포
	형상	Region Shape	2값 화상에 의한 형상
		Contour Shape	윤곽에 의한 형상
		Shape 3D	3차원의 표면형상
	움직임	Motion Activity	동화상 변화의 강약
Prametric Motion		동화상 변화의 파라미터 표현	
궤적	Motion Trajectory	주목점의 움직임의 시공간 궤적	
2. (고도특징)	얼굴	Face Recognition	얼굴의 특징
	촬영조건	Camera Motion	카메라의 움직임
3. (위치지정)	위치지정	Region Locator	정지화면의 영역
		Spatio-Temporal Locator	동화상의 시공간 영역
4. (Container)	Container	Grid Layout	화상의 구형 분할
		Visual Time Series	동화상의 시간구간에 의한 분할
		Multiple View	복수의 방향에서 본 화상

표 3. 저레벨 기술의 오디오기술 Scheme 가운데 악기의 음색을 나타내는 팀바기술 scheme

Class	1	2	3	4
sound의 특성	지속성 있음	지속성 있음	지속성 없음	지속성 있음
	하모니 있음	하모니 없음	-	-
	일관성 있음	일관성 있음	-	일관성 없음
sound의 예	바이올린, 플룻	벨, 트라이앵글	snare drum, 클라베스	심벌, 백색잡음

기술 scheme」이 있다(표 3). 악기의 음색은 인간의 3개의 지각 특성인 지속성, 하모니(화성), 일관성을 조합시킴으로써 4개로 분류할 수 있다. 팀바기술 scheme은 이들 4개 가운데 어느 분류에 속할지를 판별하기 위해 이용된다.

이외에 ① 일반적인 음향효과의 class 분류에 이용되는 「sound recognition 기술 scheme」, ② 발생된 언어를 음색으로 분할하여 그 개개의 음

색이 어느 순번에 연결되는가를 기술하는 「spoken contents 기술 scheme」, ③ 음레벨의 정확한 기술로서 보다 정확한 멜로디 검색을 가능케 하는 「melody 기술 scheme」, ④ 허밍 등에 의한 유사 멜로디 검색을 가능케 하는 「melody contour 기술 scheme」 등이 있다.

XML 데이터는 BiM에서 binary 압축

고레벨 기술에서도 저레벨 기술에서도 표현 형식은 XML schema로 정의되고 XML로 기술된다. 그러나 특히 저레벨 기술과 같이 영상이나 오디오의 수치 데이터를 다량으로 포함하는 메타데이터를 XML에서 텍스트 표현하면 데이터 양이 방대해진다.

그래서 MPEG-7은 XML 표현을 대등한 binary 표현으로 변환하는 BiM을 준비하였다. BiM을 사용하여 XML schema로 정의된 표현 형식을 일정 규칙에 의거하여 binary로 변환한다.

더욱이 visual 기술자에 대하여는 압축효율을 추척한 binary 표현도 특별히 정하고 있다. 화상의 특징을 나타내는 수치데이터가, 가끔 BiM을 이용하여 압축해도 방대한 데이터양이 되기 때문이다.

이 때문에 BiM에서는 binary 데이터의 일부가 visual 전용 binary code인 경우 visual 전용 decoder를 사용하여 복호하는 역할을 갖는다. 고레벨 기술이나 오디오 기술 scheme에서는 이와 같은 예외규정은 없다.

각지의 영상 모음을 횡적으로 사용한다.

멀티미디어 콘텐츠의 부족을 해소하는 한가지 수단으로서 대량의 영상데이터를 도서관과 같이 관리하는 영상모음시스템이 고안되어 있다. 영상모음에서는 제작회사가 희망하는 영상소재를 검색하여 그 소재를 싼 가격에 편집할 수 있는 것이 중요하다. 같은 영상소재를 사용해도 다양한 콘텐츠를 싼 가격에 제작할 수 있게 하기 위해서이다.

이러한 영상검색 실현의 길을 연 것이 영상의 디지털화와 메타데이터에 의한 내용기술이다.

영상모음시스템에 저장되는 영상의 메타데이터에서는 미국의 방송업계단체인 SMPTE (society of Motion Picture and Television Engineers)나, 유럽중심의 방송사단체인 EBU (European Broadcasting Union)가 규정하는 것도 포함될 가능성이 있다.

그러나 국제표준인 MPEG-7은 이들 규격을 포함하도록 검토되어 있고, 그위에 상세한 내용 기술을 가능케 하고 있다. 이미 각 TV국이 보유하고 있는 영상 콘텐츠나 콘텐츠 관련 정보를 MPEG-7을 축으로 하여 상호 교환하는 것도 그리 어렵지 않다고 예상된다.

또 MPEG-7을 이용하여 영상 모음을 분산시켜 관리하고, 하나의 검색 엔진으로 처리하는 것도 가능하게 된다. 분산한 영상 콘텐츠를 관리하고 원격지에서 바라는 영상 콘텐츠를 검색하는 서비스도 실현 가능하다.

새로이 작성하는 영상 콘텐츠의 제작단계부터 메타데이터를 생성해 가는 것도 향후 점차 중요해 질 것이다. 영상모음시스템으로 이용을 쉽게 할 뿐만 아니라 상류공정에서 편집효율을 높이는 것에도 효과가 있기 때문이다.

메타데이터를 붙인 영상 소재를 이용하면 다양한 용도에 따라 내용이나 영상시간을 변화시킨 프로그램의 제작을 자동화할 수 있다. 메타데이터를 첨가하는 시간을 들여도 제작비용의 감소효과는 충분히 기대할 수 있다.

남은 과제는 콘텐츠 보호와 과금

단 MPEG-7에서도 부족한 분야는 있다. 디지털 콘텐츠의 보안 대책이나 과금이다.

디지털 콘텐츠는 시청자 스스로가 품질을 떨어뜨리지 않고 복사할 수 있다. 이 때문에 콘텐츠 송신 사업자에게 있어서는 부정복사 대책이 필수적이다. 더욱이 콘텐츠를 송신할 때의 과금시스템도 보안대책을 충분히 고려해야만 한다. 콘텐츠만이 아닌 메타데이터도 새로이 창조된 콘텐츠라 생각하면 메타데이터를 포함한 보호기술의 확

립이 필요해진다.

이들에 관해서는 현재 MPEG-7에서의 규정은 없다. 향후 검토과제의 하나가 되고 있다. 단, 구체적인 검토도 이미 시작되고 있다. 현재 MPEG-7의 후속인 「MPEG-21」로서 국제표준 책정이 진행되고 있는 단계이다. 

용어설명

차기 CS 방송	동경110도로 쏘아 올린 통신위성(CS)을 사용하여 제공하는 방송 서비스. 방송위성(BS)과 같이 궤도상에 있기 때문에 수신안테나를 BS와 공용할 수 있는 등의 이점이 있다.
FDIS	final draft of international standard 국제표준의 최종초안으로 기술내용이 확정된 도큐먼트. 오자정정이나 포맷수정을 한 후, IS(international standard)로 출판된다.
ISO/IEC	1987년에 설치된 ISO(국제표준화기구)와 IEC(국제전기표준회의)의 공동기구(joint technical committee)인 「ISO/IEC JTC1」에 의해 작성되는 규격. ISO는 지적, 과학적, 기술적 및 경제적 활동의 협조를 촉진하기 위한 표준화를 담당하는 비정부기관. IEC는 전기 및 전자 기술분야의 표준화 역할을 하는 비정부기관. 설립은 각각 1947년, 1906년.
Closed Caption	영상 소프트에 들어있는 자막신호. decoder를 통하는 것으로 화면에 자막을 내보낸다. 미국에서는 시각장애자를 위해 TV 프로그램에 넣는 것이 의무화되어 있다.
Personal Video Recorder(PVR)	디지털 영상을 하드디스크(HDD)나 DVD-RAM에 녹화·재생할 수 있는 개인용 기기. personal digital recorder(PDR)이라고 한다. 이미 도시바나 소니 등이 PVR을 판매하고 있지만 MPEG-7에 의한 검색기능은 미탑재
TV Anytime	비영리목적의 국제적인 포럼. TV 방송을 수신기에 일단 저장하고, 언제라도 시청할 수 있는 시스템의 표준화를 진행하고 있다. 방송국이나 방송기기 제조업체 등으로 구성.

MPEG-7	더욱 상세한 내용은 http://www.cselt.it/mpeg , http://www.mpeg-industry.com 참조
MPEG	Moving Picture Experts Group. ISO/IEC/JTC1 산하에서 동화상압축 등에 관한 국제표준화활동을 하는 조직. 1988년에 탄생. 표준방식 그 자체의 명칭으로 사용되는 경우도 많다.
XML	extensible markup language. 문서정보 등의 언어사양을 정의하기 위한 언어. W3C가 SGML(standard generalized markup language)을 기본으로 표준화하였다.
W3C	WWW 컨소시엄. WWW 관련 기술의 표준화를 진행하고 있는 업계단체. 인터넷 관련 주요 업체가 참가하고 있다. HTTP(hypertext transport protocol), HTML(hypertext markup language), XML 등 많은 중요한 기술을 표준화해 왔다.
XML Schema	XML 문서의 구조와 내용, 의미를 정의하는 schema 언어. W3C에 의해 표준화되었다.
Scheme	XML, 문서의 구조 및 정합성 규약을 정의하기 위한 정의문
Texture	주기적 패턴을 갖는 모양. 예를 들면 옷감의 결 모양.
지속성, 하모니, 일관성	지속성은, 음량이 시간적으로 지속하는 특성. 하모니는 2개 이상의 음(주파수)이 울리는 특성. 일관성은 음의 시간적 변화가 규칙적인 특성.

cdma2000 1x EVDO

cdma2000 1x EVDO는 미국의 통신업체 퀄컴이 지난 99년 하반기에 시험통화를 통해 선보인 3세대 초고속 무선데이터 통신기술이다. 이 방식은 고속의 패킷 전송에 적합하도록 시스템을 최적화함으로써 문자 영상 음악 등의 데이터를 1.25MHz의 협대역에서 최고 2.45Mbps 속도로 전송, 기존 144Kbps 수준인 cdma2000 1x보다 약 20배 빠른 속도로 데이터를 전송할 수 있다. 특히 이 기술은 데이터 수신에 특화된 것으로 유선상의 초고속 인터넷 서비스인 ADSL 서비스와 흡사해 '모바일 ADSL'로 불리기도 한다. cdma2000 1x EVDO 기술을 이용하면 3세대 특유의 메가급 데이터 전송이 가능해 영상전화나 단말기를 통한 실시간 동영상 구현과 같은 초고속 무선 멀티미디어 서비스를 할 수 있다. 아울러 이 기술은 동기식 기술이어서 IS95A B, cdma2000 1x와 데이터 분야에서 호환이 가능, 기존 시스템의 업그레이드에 적합하며 cdma2000 1x 시스템의 네트워크를 그대로 사용할 수 있어서 초기 설비투자를 최소화할 수 있다.

cdma 1x EVDO는 개발 초기 HDR(High Data Rate)로 불렸으나 지난해 10월 동기식국제표준화기구(3GPP2)에서 'cdma2000 1x EVDO(Evolution Data Only)'를 공식 명칭으로 채택, 사용하고 있다.