

동영상 정보의 메타데이터 구축에 관한 연구

A Study on the Development of Metadata of Moving Images

김진아(Jin-Ah Kim)*, 김태수(Tae-Soo Kim)**

초 록

동영상 정보에 대한 효과적인 검색도구로 메타데이터를 제안하였다. 메타데이터 체계를 구축하기 위한 기반으로 인터넷 정보자원에 대한 대표적인 메타데이터인 더블린 코어와 MPEG의 새로운 표준이며 국제표준으로 개발되고 있는 MPEG-7을 이용하였다. 메타데이터 기술의 기반으로 RDF를 사용하였으며, 이에 따라 RDF 스키마인 MovingImageCore를 구축하였다. 아울러 ASP를 이용한 웹기반 메타데이터 데이터베이스 시스템을 구축하였으며 이를 통해 RDF/XML형식으로 기술된 동영상 정보의 메타데이터를 웹에서 검색할 수 있다.

ABSTRACT

This paper suggests metadata as an effective search tool suitable for moving images on the web. Dublin Core, a leading metadata scheme for web resources, and MPEG-7, a coming international standard developed by MPEG, are used to construct a hybrid metadata scheme. RDF is selected as a framework for a metadata description and MovingImageCore was developed as a RDF schema. Finally, RDF/XML encoded metadata of moving images are retrieved through web-based database system using ASP technology.

키워드 : 동영상 정보, 메타데이터, 더블린 코어,
MPEG-7, RDF

* 연세대학교 대학원 문헌정보학과(ischoe@hananet.net)

** 연세대학교 문헌정보학과 교수(btsee@yonsei.ac.kr)

■ 논문 접수일 : 2001년 10월 8일

■ 게재 확정일 : 2001년 11월 7일

1 서 론

멀티미디어란 복수의 미디어가 결합되어 일정한 형태의 플랫폼을 통해 공급되는 정보 서비스 혹은 정보 시스템을 일컫는 말이다. 멀티미디어에서는 하나로 통합된 복수의 매체를 통해 정보가 전송되므로 단일 매체를 통한 전송방식보다 더 빠르고 쉽게 정보를 인식하고 이해할 수 있다. 이와 같은 멀티미디어의 장점과 더불어 네트워크와 동영상 압축기술의 발달로 인해, 인터넷에서 이용되는 멀티미디어 정보량이 급속히 증가되고 있으며 이러한 현상은 앞으로 더욱 심화될 것으로 보인다.

웹 자원에 대한 효과적인 검색도구의 개발과 함께 멀티미디어 정보에 대한 검색도구의 필요성도 증대되어 왔다. 지금의 메타데이터는 기본적으로 네트워크 자원을 체계적으로 기술하고 이를 효과적으로 검색하기 위한 것으로, 그 동안 주로 텍스트 중심의 정보를 대상으로 활발히 연구되어 왔다. 그러나 멀티미디어 정보는 그 속성상 텍스트 형식의 정보를 위주로 개발된 검색엔진을 통해서서는 효과적인 검색이 어렵다. 예를 들어 특정 영화에서 주인공이 오토바이를 타고 달리는 장면을 검색하고자 할 때, 이를 현재의 웹 검색엔진을 통해 검색하는 것은 거의 불가능하다. 따라서 멀티미디어 정보의 경우에는 텍스트 정보 검색엔진과는 다른 접근 방식과 검색 도구를 필요로 하고 있다.

이 연구는 복수의 메타데이터를 사용하여 동영상 정보를 통합기술하고 검색하기 위한 기법을 제시하는 데에 그 목적이 있다. 이를

위해 '좋은 나라 소방본부'에서 제작된 소방 관련 동영상을 물리적 구성요소로 분석하고 이를 더블린 코어(Dublin Core: DC)와 MPEG-7의 기술요소로 변환한 다음, 자원기술구조(Resource Description Framework: RDF) 스키마에 따라 통합 기술하였으며, 검색용 메타데이터 데이터베이스를 구축하였다. 다만 동영상 정보 중 청각 정보는 본 연구의 범위에서 제외하였다.

앞으로 다양한 RDF 데이터 유형 개발과 동영상 정보의 시청각 특성에 대한 자동 추출 기술 및 정보통신 기술의 지속적인 개발을 통해, RDF 형식으로 기술된 동영상 정보의 메타데이터는 웹에 분포된 동영상 정보에 대한 검색 도구이자 교환 형식으로 유용하게 이용될 수 있을 것이다.

2 멀티미디어에서 메타데이터의 필요성

멀티미디어 정보에 대한 메타데이터의 필요성은 네트워크 환경과 멀티미디어라는 미디어 특성과 주로 관련되어 있다. 먼저 네트워크 환경과 관련하여 보면, 기본적으로 인터넷 사용의 증가에 따라 양질의 인터넷 정보에 접근하고 이를 검색할 수 있는 도구가 필요하게 된 것이다. 그런데 기존의 검색엔진은 로봇에 의해 자동으로 구축된 색인을 이용하므로 색인의 깊이나 질을 신뢰하기 어렵고, 적합한 정보를 검색하는데 한계를 가지고 있다. 또한 웹 검색엔진은 주로 텍스트 형식의 정보를 대상으로 개발되었기 때문에 이미지,

오디오, 비디오 등 텍스트 이외의 다른 형식의 정보를 검색하는 데에 한계가 있다. 따라서 메타데이터라는 일정한 형식을 통해 인터넷 정보자원을 적절히 기술하고 제어함으로써 검색효율을 높일 수 있으며, 메타데이터를 이용함으로써 정보에 대한 접근이 개선되고, 검색 대상인 객체에 대한 배경 정보를 확보할 수 있으며, 정보에 대한 제어가 가능하다는 점에서 그 필요성이 커지게 된 것이다.

한편, 멀티미디어의 미디어적 특성과 관련하여 메타데이터가 필요한 이유는 질의 패러다임이 텍스트 검색의 그것과 상이하고, 현재로서는 내용 기반 처리 기술이 부족하며, 시스템 성능이나 멀티미디어 정보의 의미 측면과 관련되어 있다(Boll, Klas and Sheth 1998). 질의 패러다임이란 텍스트 형식의 정보검색에서 가능한 완전 일치(exact-match)의 질의 패러다임이 멀티미디어 정보에 대해서는 적절하지 않다는 것을 의미한다. 예를 들어 '사과'라는 과일에 대한 텍스트 형식의 정보를 검색하는 경우 '사과'라는 검색어를 입력함으로써 본문 중에 '사과'라는 단어를 포함하는 정보에 접근할 수 있다. 이것은 검색어와 색인어가 완전 일치되는 질의 패러다임을 사용한 결과이다. 그러나 사과라는 과일이 묘사된 이미지를 검색하기 위해서는 이 방식을 사용할 수 없다. 검색 대상이 되는 이미지에 묘사된 객체와 완전히 일치하는 질의를 작성하는 것이 거의 불가능하기 때문이다.

이밖에 멀티미디어 정보에 대한 내용 기반 처리 기술이 현재로서는 부족한 상태이며, 내용 기반 처리가 가능하다 하더라도 대용량의 멀티미디어 정보에 대해 질의를 수행할 때마

다 내용 기반 처리를 한다는 것은 시스템 성능을 저하시킬 수 있으므로 효율적이지 못하다. 이 경우 메타데이터를 이용함으로써 시스템 응답 시간을 줄일 수 있다. 특히 멀티미디어 정보가 가지는 의미는 보통 원정보의 표면에 드러나지 않는다. 정보의 의미는 분석과 처리를 통해 원정보로부터 추출되어 속성 값이나 개념 등의 메타데이터로 표현될 수 있으며 이용자는 메타데이터를 통해 원정보에 접근할 수 있다. 특히 추출된 메타데이터가 텍스트 형식의 주석이나 속성 값으로 표현되는 경우에는 기존의 텍스트 처리 기법이나 전통적인 정보검색 기법을 사용한 정보처리가 가능하므로 효과적이다.

2.1 더블린 코어

DC는 기술 메타데이터(descriptive metadata)로서, 네트워크 환경에서 웹 자원을 기술하고 접근하는데 필요한 15개의 기본 데이터 요소로 구성되어 있다. 이 DC는 데이터 요소의 단순화를 통해 메타데이터의 유용성과 이용의 용이성을 확보하면서, 필요에 따라 데이터 요소의 확장을 통해 특정한 응용 분야에 적용이 가능하다. 또한 요소의 의미와 내용을 특정한 기술규칙이나 구문과 독립적으로 정의하고 있어 메타데이터의 독립성과 함께 상호 운용성이 확보되어 있다.

그러나 다양한 분야에 단일 형식의 메타데이터만을 적용하는 것이 실제로는 불가능한 경우가 있을 수 있다. 이런 점에서 DC에서도 한정어를 사용한 확장방안이 제시되고 있다. 이 한정어는 상호 운용성을 고려한 최소한의

〈표 1〉 더블링크어 공식 한정어

요소명	요소 한정	인코딩 스킴
Title	Alternative	-
Creator		-
Subject		LCSH, MeSH, DDC, LCC, UDC
Description	Table Of Contents Abstract	-
Publisher	-	-
Contributor	-	-
Date	Created Valid Available Issued Modified	DCMI Period W3C-DTF
Type	-	DCMI Type Vocabulary
Format	Extent	-
	Medium	IMT
Identifier		URI
Source		URI
Language	-	ISO 639-2 RFC 1766
Relation	Is Version Of Has Version Is Replaced By Replaces Is Required By Requires Is Part Of Has Part Is Referenced By References Is Format Of Has Format	URI
Coverage	Spatial	DCMI Point ISO 3166 DCMI Box TGN
	Temporal	DCMI Period Wec-DTF
Right	-	-

한정어로서, 응용 분야에 따라 부가적인 한정어를 사용할 수 있다. 그런데 한정어를 통한 더블린 코어의 확장 대신, 다른 메타데이터와의 협력이나 상호보완을 통해 특정 분야의 요구사항이 해결되어야 한다는 주장도 제기된 바 있다(Lagoze 2000, 2001).

〈표 1〉은 더블린 코어의 데이터 요소와 한정어를 정리한 것이다.

2.2 MPEG-7

MPEG-7은 동영상 전문가 그룹인 MPEG (Moving Picture Experts Group)에서 개발 중인 ISO/IEC(International Organization for Standardization/International Electrotechnical Commission) 표준으로서, 공식 명칭은 '멀티미디어 콘텐츠 기술을 위한 인터페이스' (multimedia content description interface)이다. 기존의 MPEG 표준이 주로 시청각 정보의 코딩 및 데이터 압축을 목적으로 한 것임에 비해, MPEG-7은 멀티미디어 정보를 기술하기 위한 공통의 인터페이스 표준화에 초점을 두고 있다. 다시 말해 MPEG-7은 정지 화상과 그래픽, 입체자료, 소리, 음성, 비디오 등 주로 시청각 정보의 기술(記述)의 표준화를 통해, 멀티미디어 정보의 효율적인 검색 수단을 확보하기 위한 것이다. 이에 따라 MPEG-7에는 콘텐츠의 저작 및 생산에 대한 정보, 콘텐츠의 이용에 관한 정보, 콘텐츠를 구성하는 요소들간의 시공간적 관계에 대한 정보, 콘텐츠에 포함된 시청각적 특성, 콘텐츠에 의해 포착된 개념 등이 포함되어 있다. 이와 같은 내용을 기술하기 위해 MPEG-7에

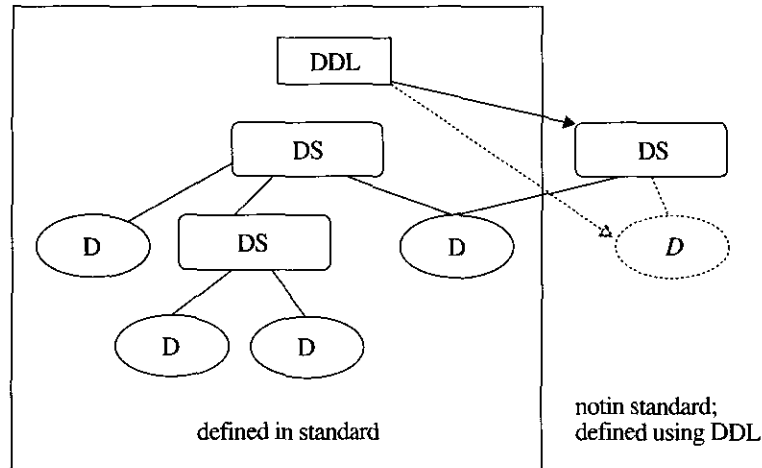
서는 내용 기술을 위한 도구로 디스크립터(descriptor: D), 기술체계(description scheme: DS), 기술정의언어(description definition language: DDL)를 사용한다.

디스크립터는 기술하고자 하는 데이터의 특성을 표현한 것으로 그 구문과 의미가 정의되어 있다. 여기서 데이터의 특성이란 이용자에게 의미 있는 데이터의 성격을 말한다. 예컨대 이미지를 기술하기 위해 '색상'이라는 특성을 사용한다면 이를 표현하기 위한 디스크립터는 색상 히스토그램(color histogram), RGB 백터, 문자열 등이 있을 수 있고, 디스크립터의 값은 RGB = (255,255,255) 혹은 colorstring = "red"와 같이 표현될 수 있다(Hunter and Iannella 1998).

기술체계는 콘텐츠 기술에 사용되는 구성요소들간의 관계에 대한 명세이다. 구성요소란 디스크립터나 기술체계를 의미하는데 디스크립터가 특성을 표현하는 반면, 기술체계는 전반적인 기술구조를 표현한 것이다. 이들은 기술체계를 통해 상호 관계의 의미나 구조가 설명된다.

기술정의언어(DDL)는 기존의 기술체계나 디스크립터를 확장하거나 수정, 혹은 새로운 기술체계나 디스크립터를 생성하기 위해 사용되는 언어로서, 디스크립터와 기술체계를 조직적으로 표현하기 위한 구문 규칙을 정의한다. 따라서 기술정의언어는 시청각 정보를 모형화한 결과를 표현하는 스키마 언어라고 할 수 있다. 〈그림 1〉은 이상에서 설명한 디스크립터와 기술체계, 기술정의언어 간의 관계를 제시한 것이다.

MPEG-7의 멀티미디어 기술체계에는 오



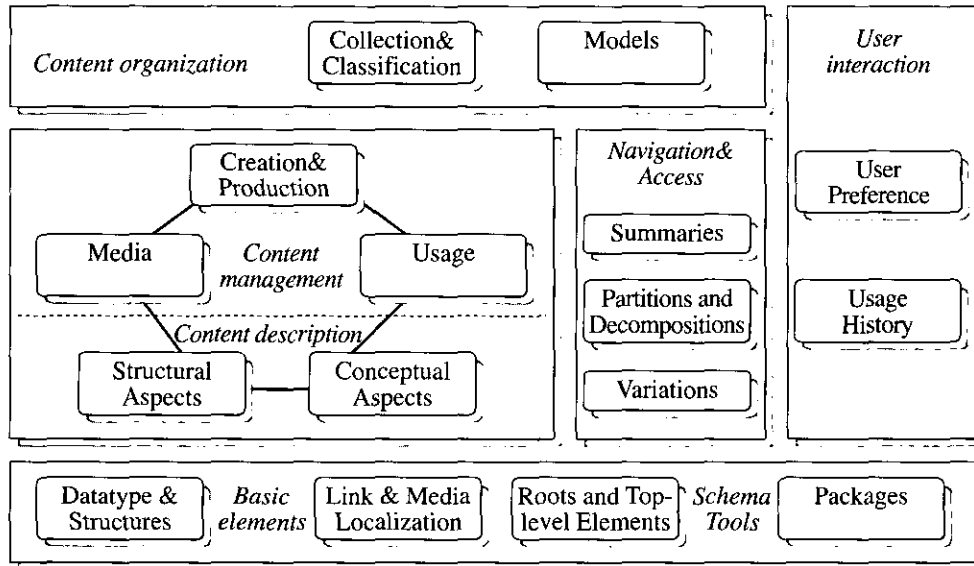
〈그림 1〉 MPEG-7 디스크립터, 기술체계, 기술정의언어간의 관계

디오, 비디오 등 모든 종류의 멀티미디어 정보와 함께 내용과 저작, 생산, 이용, 관리 등에 필요한 정보가 기술된다. 이들 정보는 색상이나 질감, 움직임 등 대체로 콘텐츠로부터 자동 추출이 가능한 저차원적인 시청각적 특성과, 의미 있는 객체나 사건, 추상적인 개념에 대한 정보, 그리고 콘텐츠 관리 과정이나 저장 미디어에 대한 정보와 같이 인간의 노력이 개입되어야만 추출될 수 있는 고차원적인 특성으로 분류된다. 시청각적 특성을 기술하는 저차원적인 디스크립터의 일부는 원정보로부터 자동 추출이 가능하지만 자동화가 불가능한 대부분의 디스크립터는 인간에 의해 수동으로 추출되어야 한다.

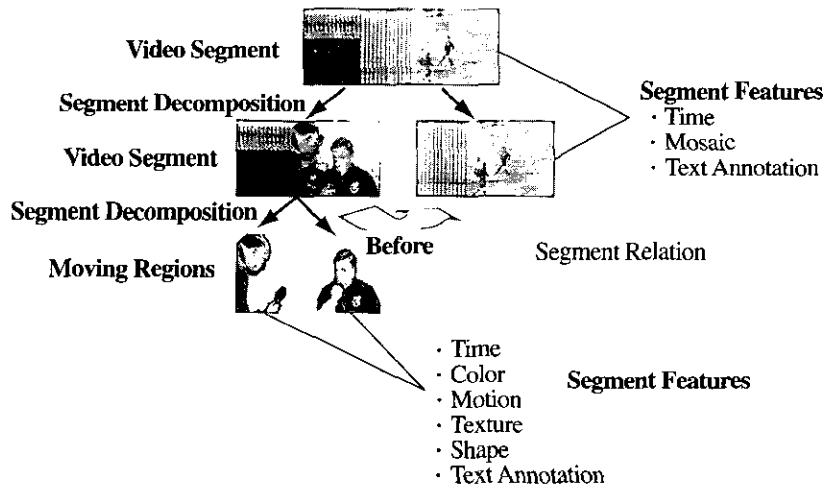
MPEG-7 멀티미디어 기술체계는 〈그림 2〉와 같이 기본요소, 스키마 도구, 콘텐츠 기술, 콘텐츠 관리, 콘텐츠 조직, 탐색과 접근, 이용자 상호작용이라는 7가지 부분으로 구성된다. 이 가운데 정보검색과 직접 관련된 부분은 콘텐츠 관리와 콘텐츠 기술부분으로서, 이를 통해 멀티미디어 정보의 내용이 기술되고 정보

의 제작과 생산, 이용에 대한 사항이 표현된다.

콘텐츠 관리에는 저작 정보와 이용 정보, 미디어 정보가 기술되며, 콘텐츠 기술에는 콘텐츠의 구조적인 관점과 개념적인 관점이 기술된다. 콘텐츠를 구조적으로 기술한다는 것은 콘텐츠를 물리적이거나 논리적 단위인 세그먼트(segment)로 분리한 다음, 각 세그먼트에 대한 시청각적 특성을 기술하고, 세그먼트간의 관계를 조직하고 규정하는 것으로, 콘텐츠의 시청각적 내용을 구조 측면에서 기술하는 것을 말한다. 예를 들어 비디오를 구성하는 각 프레임이나 이미지의 임의의 영역에 포함된 시청각 내용이 기술된다. 〈그림 3〉은 비디오 시퀀스를 구조적으로 기술한 예이다. 이 그림에서 하나의 비디오 세그먼트가 두 개의 비디오 세그먼트로 분해된 후, 다시 복수의 동영상 세그먼트로 분해되었다. 1차 분해에 의해 생성된 두 개의 비디오 세그먼트는 전후관계(before)로 연결되며, 세그먼트에 따라 상이한 특성에 의한 시청각 내용이 기술되고 있다.



〈그림 2〉 MPEG-7 기술체계의 구성



〈그림 3〉 비디오 시퀀스의 구조적 기술의 예

한편, 콘텐츠를 개념적인 관점에서 기술한다는 것은 콘텐츠로 표현된 실세계의 의미, 혹은 개념측면에서 본 콘텐츠의 시청각 내용을 의미한다. 즉 콘텐츠로부터 이용자의 관심 대상이 되는 의미 있는 개체를 식별하고, 이

개체의 속성을 기술한 다음, 각 개체간의 관계를 정의하는 것을 의미한다. 콘텐츠의 내용에 관심을 가진 이용자를 대상으로 하는 경우에는 개념적 관점에서 기술하는 것이 유용할 수 있다.

의미 있는 개체들은 의미 기술체계 (semantic DS)에 따라 기술된다. 의미 기술 체계는 추상적인 개념이므로 실제로 개체는 개념 기술체계, 객체 기술체계, 인간 기술체계, 사건 기술체계, 의미상태 기술체계와 같은 하위 항목으로 표현된다. <표 2>는 의미 기술체계에 대한 하위 기술체계와 이를 통해 기술되는 개체의 예를 제시한 것이다.

3 동영상 정보의 메타데이터

3.1 동영상 정보 모형

메타데이터를 기술하기 위해서는 기술 대상인 동영상 정보의 구조를 파악하여 메타데이터 체계의 토대가 되는 동영상 정보 모형을 구축해야 한다. 이 모형에서는 동영상이 어떤 정보로 구성되어 있는가를 파악하고, 이러한 물리적 구성단위로 메타데이터의 데이터 요소를 식별하고 이 식별단위요소별로 정보를 기술하게 된다. 본 연구에서는 동영상을

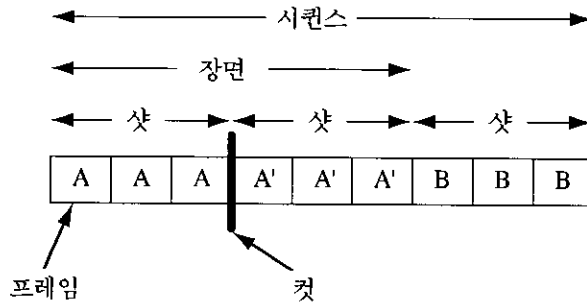
물리적으로 분할하여 기술단위를 식별한 뒤 더블린 코어와 MPEG-7으로부터 적절한 데이터 요소를 추출하여 기술하였다.

일반적으로 동영상은 물리적으로 시퀀스(sequence)와 장면(scene), 샷(shot), 프레임(frame)에 의해 계층적으로 조직된다(김우생 외 1998: Hunter and Armstrong 1999). 프레임은 동영상을 구성하는 최소 단위로 필름 한 장에 해당되며 하나의 화상을 표현한다. 동영상에서 장면의 전환이 이루어지는 부분을 컷(cut)이라고 하며, 컷으로 구분되는 동영상의 단위가 샷이다. 샷은 하나의 카메라 동작에 의해서 촬영된 프레임의 집합이다. 장면은 연속된 샷으로 이루어진 동영상 단위인 데 하나의 사건 및 행동을 여러 각도에서 본 것을 의미한다. 시퀀스는 논리적으로 관련된 장면의 집합을 말한다. 가령 뉴스 동영상의 경우 기자가 보도하는 각 사건이나 일기예보 등이 독립된 시퀀스를 구성할 수 있다. <그림 4>는 시퀀스와 장면, 샷, 프레임으로 구성되는 동영상의 물리적 구조를 제시한 것이다.

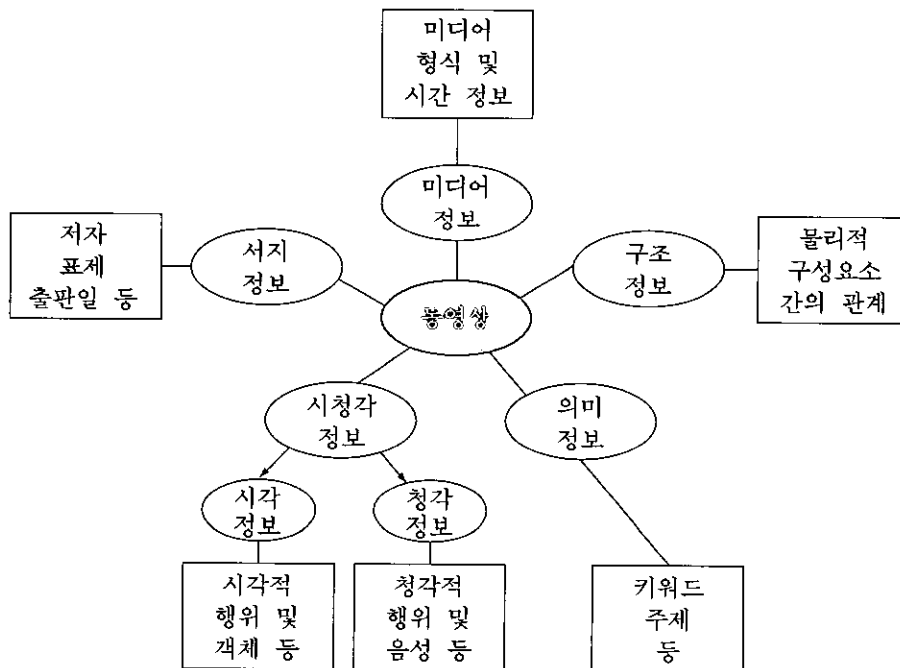
이와 같은 물리적 구조 이외에 메타데이터

<표 2> 의미 기술체계의 하위 기술체계

Semantic DS의 하위 기술체계	기술되는 개체
개념 기술체계 (Conceptual DS)	개발 중
객체 기술체계 (Object DS)	물리적이거나 추상적인 객체
인간 기술체계 (PersonObject DS)	사람
사건 기술체계 (Event DS)	의미 있는 상태의 변화
의미상태 기술체계 (SemanticState DS)	특정한 시공간에서 포착된 개체의 속성



〈그림 4〉 동영상의 물리적 구조



〈그림 5〉 동영상정보의 구성

에 사용될 동영상 정보의 적절한 데이터 요소를 선정하기 위해서는 동영상 정보의 구성을 파악할 필요가 있다. 즉 정보가 구체적으로 어떤 요소로 구성되어 있으며, 어떠한 요소들이 동영상 정보 기술에 필요한 것인지를 파악하여 이를 표현할 수 있는 메타데이터 요소를 선정하기 위한 것이다. 본 연구에서는

기존의 연구 결과에 기초하여 동영상 정보의 구성요소를 서지정보와 시각정보, 의미정보, 구조정보, 미디어정보로 구성하였다 (Hampapur and Jain 1998; Grosky 1994; Bohm and Rakow 1994).

‘서지정보란 동영상 정보의 제목이나 저작자, 저작일, 발행인, 정보원, 저작권 등의 정보

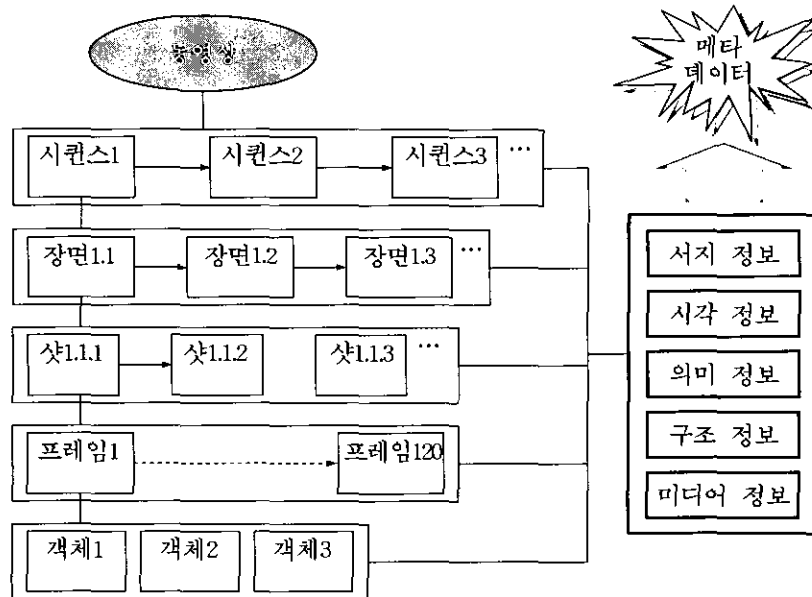
를 말하며 보통 동영상 전체에 대해 한번 기술된다. 의미정보는 동영상으로 전달되는 정보나 메시지를 의미하며, 시각정보는 이러한 의미정보를 전달하는 매개물로서 동영상에 나타난 시각적인 특성을 말한다. 가령 범죄 사건을 보도하는 뉴스 동영상의 경우, 범죄의 내용이나 발생 장소, 피해자 등의 정보는 의미정보가 되며, 범죄 사실을 보도하는 과정에서 볼 수 있는 범죄 발생 현장 사진은 시각정보가 되는 것이다. 구조정보는 전체 동영상의 물리적 구성요소들의 계층 관계에 대한 정보를 제공한다. 예를 들어 시퀀스1이 장면1.2.3으로 구성되어 있다는 등의 정보가 해당된다. 또 미디어 정보는 전체 동영상 및 그 구성요소들의 미디어 특성에 관련된 정보를 의미하는데, 파일 형식이나 미디어의 재생시간 등의 정보가 이에 해당된다.

이상에서 살펴본 동영상의 물리적 구조와

동영상 정보의 구성에 따라 동영상 정보의 모형을 제시하면 <그림 6>과 같다. 이 그림에서 보는 바와 같이 전체 동영상은 시퀀스와 장면, 샷과 같은 물리적 구성요소별로 분해되며, 서지정보와 의미정보, 구조정보 등이 전체 동영상 및 물리적 구성요소들의 메타데이터로 기술된다.

3.2 더블린 코어와 MPEG-7을 이용한 통합기술체계

정보자원의 적용분야나 형식의 다양성을 고려할 때 특정 형식의 단일 메타데이터로 모든 자원을 효과적으로 기술하는 것은 사실상 불가능하다. 따라서 본 연구에서는 동영상 정보자원의 기술을 위해 더블린 코어와 MPEG-7의 두 가지 메타데이터를 통합하여 사용하였다. 이들 메타데이터의 적용 범위가



<그림 6> 동영상 정보 모형

〈표 3〉 더블링크어와 MPEG-7으로부터 추출된 요소

요소명	D C	DC한정어	MPEG-7 기술체계 및 디스크립터		
서지정보	Title		Financial	Income	
	Creator				
	Publisher				
	Contributor				
	Source				
	Right				
	Date	Created			Cost
		Vaild			
		Available			
		Issued			
Modified					
의미정보	Subject				
	Description				
	Type				
	Coverage				Temporal
					Spatial
구조정보	Relation	HasPart			
미디어정보	Format	Medium	MediaTime	MediaTimePoint	
		Extent		MediaDuration	
	Identifer				
시각정보			Color	DominantColor	
			Texture	Texture-Browsing	Regularity
					Scale
					Direction
Motion	CameraMotion				

국제적이며, 사실상 국제 표준으로 사용되거나 공식적인 표준화 작업이 진행되고 있기 때문이다. 이와 더불어 더블링크어와 MPEG-7이 동영상 정보를 각각 독자적으로 기술하는 데에는 몇 가지 한계를 지니고 있다. 더블

링크어의 경우, 기본적으로 텍스트 정보 위주로 개발된 메타데이터이기 때문에 동영상의 시청각 정보의 기술에는 부적절하다. 한편 MPEG-7은 기술의 복잡성 때문에 웹 정보자원과 같이 신속한 접근과 검색이 요구되는 정

보를 기술하기에 부적합한 측면이 있다.

그런데 더블린 코어와 MPEG-7을 통합한 기술체계를 사용하게 되면 상호 보완적인 성격으로 인해 동영상 정보의 메타데이터에서 기술되는 서지정보와 시각정보, 의미정보, 구조정보, 미디어정보를 적절히 기술할 수 있다.

동영상 정보의 기술에 필요한 더블린 코어와 MPEG-7의 데이터 요소는 <표 2>와 같다. 이 표에서 서지정보와 의미정보, 구조정보는 더블린 코어와 그 한정어로부터, 그리고 시각정보와 미디어정보의 일부는 MPEG-7으로부터 추출한 것이다. 미디어 정보 중 MediaTimePoint는 전체 동영상 중 동영상 정보의 물리적 구성요소가 실행되는 시작 시간을 의미하며, MediaDuration은 해당 구성요소의 총 재생 시간을 의미한다. DominantColor는 기술의 대상인 동영상 구성요소의 대표 색상을 의미하며, Texture의 Regularity, Scale, Direction은 각각 질감의 규칙성, 촘촘함의 정도, 그리고 방향성을 기술하는데, 질감의 경우 MPEG-7에서는 이들 속성에 대한 이진 값의 범위를 제공하고 있다. CameraMotion은 boom, dolly, pan, tilt와 같이 동영상 촬영에 사용된 기본적인 카메라 동작을 기술하며, 이 속성 역시 MPEG-7에서 정의한 이진 값들 중에서 선택하여 사용된다.

4 RDF를 이용한 동영상 정보의 메타데이터 기술

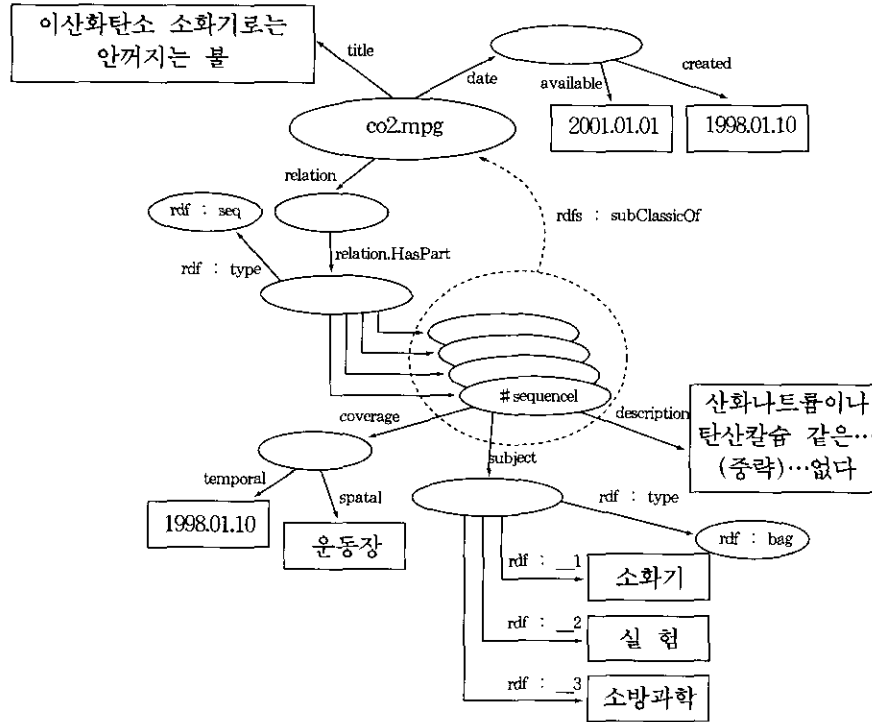
RDF는 메타데이터 기술을 위한 의미와 구분, 구조에 대한 공통 규약을 제공하며, 전송

구문으로는 XML을 사용한다. 이 RDF는 특정 분야의 자원 기술에 필요한 어휘를 제공하는 것이 아니라, 각 분야별로 필요한 메타데이터 요소를 정의하고 메타데이터 기술 및 전송의 상호 운용성을 구현하기 위한 것이다 (Miller 1998; Lassila and Swick 1999).

메타데이터 기술을 위한 기반구조로 RDF를 선택한 주된 이유는 RDF가 분산 환경에서 여러 메타데이터 형식이 가지는 다양한 의미와 구조를 수용할 수 있는 통합구조이기 때문이다. 따라서 동영상 정보 기술에 적합한 메타데이터 요소들을 더블린 코어와 MPEG-7으로부터 각각 추출하여 최적의 요소들만으로 동영상 정보를 기술하기 위한 것이다.

4.1 RDF 데이터 모형

RDF 데이터 모형은 속성과 속성 값으로 표현되는 방향성 그래프로 그릴 수 있는데, 앞에서 살펴본 동영상 정보 모형 및 통합 기술체계를 이용하여 동영상 정보의 RDF 데이터 모형을 구축하면 <그림 7>과 같다. 모든 메타데이터 요소들은 RDF 속성으로 표현되는데, 예를 들어 Date.Created 속성과 같이 더블린 코어 한정어로부터 추출된 속성이나 Media→MediaDuration과 같이 MPEG-7의 기술체계와 함께 사용되는 디스크립터의 경우 Seq이나 Beg 컨테이너를 사용하여 표현된다. 전체 동영상을 비롯하여 시퀀스 장면, 샷과 같은 메타데이터 기술의 대상은 RDF 기술문에서는 relation.HasPart 속성으로, 그리고 RDF 스키마상에서 rdfs:subClassOf를 통해 계층관계가 표현된다.



〈그림 7〉 동영상 정보의 RDF 데이터 모형

4.2 RDF 스키마 설계

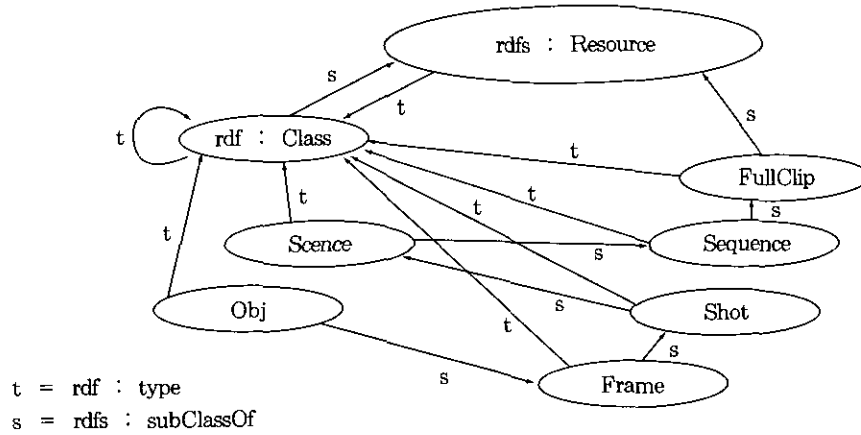
자원 기술에 사용될 속성을 파악하고 나면 새로운 스키마를 설계하거나 혹은 필요한 속성이 정의된 기존의 스키마를 RDF/XML 구문의 이름공간(namespace)을 통하여 참조하여 다시 사용할 수 있다. 그러나 현재 더블린 코어와 MPEG-7 모두 동영상 정보의 기술에 필요한 RDF 스키마를 고려하지 않고 있는 상태이다. 따라서 본 연구에서는 동영상 정보의 메타데이터 기술에 적합한 요소들만으로 구성된 새로운 스키마로 MovingImageCore (MIC)를 개발하였다. 새로운 스키마에 대한 이름공간은 가상의 공간에 정의하였으며 그

URI는 다음과 같이 선언하였다.

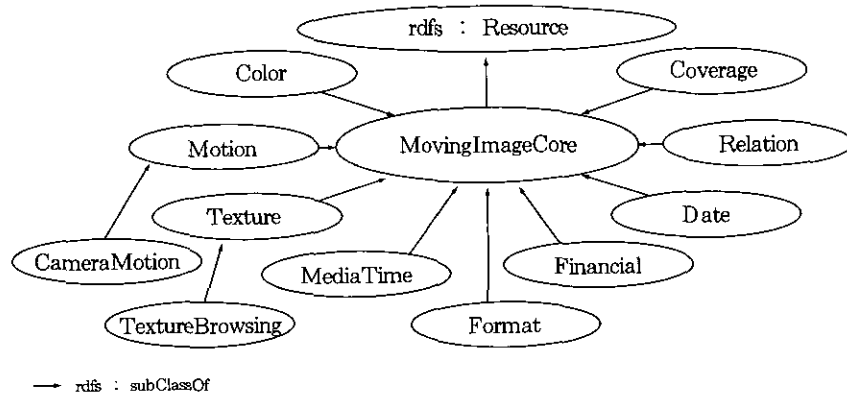
```
xmlns : mic = "http://lis.yonsei.ac.kr /enigma93/MovingImageCore#"
```

4.2.1 클래스의 지정

스키마를 작성하기 위해서는 클래스를 정의할 필요가 있다. 클래스는 물리적 측면과 논리적 측면에서 정의될 수 있다. 물리적 클래스는 메타데이터의 기술대상인 개체로 볼 수 있으므로 동영상 정보의 구성요소인 시퀀스 장면, 샷, 프레임, 객체, 그리고 전체 동영상을 물리적 개체로 식별하여 클래스명을 부여하였다. 물리적 클래스와 그들간의 계층관계는 〈그림 8〉과 같이 표현할 수 있다.



〈그림 8〉 물리적 클래스간의 계층관계



〈그림 9〉 논리적 클래스간의 계층관계

동영상 정보를 기술하기 위한 속성은 비교적 단순하기 때문에 논리적 클래스는 〈그림 9〉와 같이 정의하였다. 논리적 클래스는 Date, Created와 Coverage.Spatial처럼 속성이 더블린 코어 한정어인 경우와, MPEG-7에서 디스크 립터와 기술체계가 함께 식별되어야 하는 경우로 제한하여 클래스를 지정해주었다.

4.2.2 RDF 스키마(MIC) 기술의 예

위에서 검토한 물리적 클래스와 논리적 클래스를 기반으로 MovingImageCore 스키마를 설계하고 이를 XML로 기술하면 다음과 같다.

```

<?xml version = "1.0">
<rdf:RDF xml:lang = "en"
  xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs = "http://www.w3.org/2000/01/rdf-schema#">

<!-- ##### MovingImageCore CLASS ##### -->
<rdf:Class rdf:ID = "FullClip">
  <rdfs:subClassOf
    rdf:resource = "http://www.w3.org/2000/01/rdf-schema#Resource"/>
</rdf:Class>
<rdf:Class rdf:ID = "Sequence">
  <rdfs:subClassOf rdf:resource = "#FullClip"/>
</rdf:Class>
<rdf:Class rdf:ID = "Scene">
  <rdfs:subClassOf rdf:resource = "#Sequence"/>
</rdf:Class>

.....중 략.....

<!-- ##### MovingImageCore LOGICAL CLASS ##### -->
<rdf:Class rdf:ID = "Date">
  <rdfs:comment = "A date associated with an event in the life cycle of the resource">
  <rdfs:subClassOf rdf:resource = "#MovingImageCore"/>
</rdf:Class>

.....중 략.....

<!-- ##### Date ##### -->
<rdf:Property ID = "date_created">
  rdfs:label = "date_created"
  rdfs:comment = "Date of creation of the resource.">
  <rdfs:domain rdf:resource = "#FullClip" />
  <rdfs:domain rdf:resource = "#Date" />
  <rdfs:range rdf:resource = "http://www.w3.org/TR/xmlschema-2/#dateTime" />
</rdf:Property>
<rdf:Property ID = "date_valid">
  rdfs:label = "date_valid"
  rdfs:comment = "Date (often a range) of validity of a resource">
  <rdfs:domain rdf:resource = "#FullClip" />
  <rdfs:domain rdf:resource = "#Date" />
  <rdfs:range rdf:resource = "http://www.w3.org/TR/xmlschema-2/#dateTime" />
</rdf:Property>

.....중 략.....

<rdf:Property ID = "identifier">
  rdfs:label = "identifier"
  rdfs:comment = "An unambiguous reference to the resource within a given context">
  <rdfs:domain rdf:resource = "FullClip#" />
  <rdfs:range rdf:resource = "http://www.w3c.org/2000/01/rdf-schema#Literals" />
</rdf:Property>
</rdf:RDF>

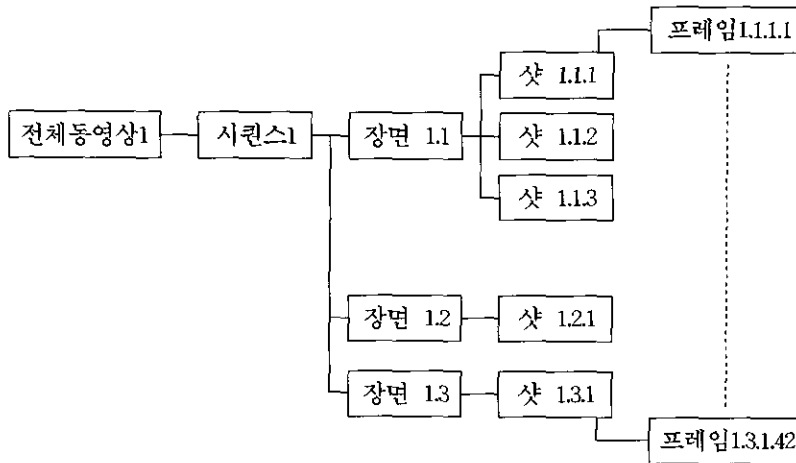
```

4.2.3 RDF 스키마를 이용한 메타데이터의 기술

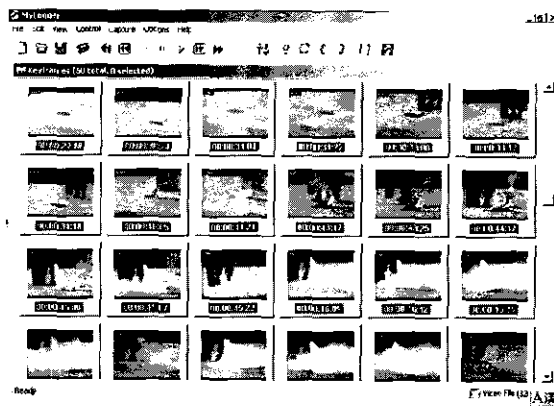
동영상 메타데이터는 전체 동영상은 물론, 시퀀스, 장면, 샷, 프레임, 객체 등 동영상의 물리적 구성요소별로 기술되어야 한다. 실제로 '좋은 나라 소방본부'에서 제작한 58초 분량의 소방 관련 동영상을 대상으로 기술하였다. 이 동영상은 <그림 10>과 같이 한 개의 시퀀스와, 3개의 장면, 5개의 샷, 그리고 42개의

프레임으로 구성되어 있다. 프레임은 동영상을 구성하는 최소 단위로 필름 한 장에 해당하는 하나의 화상을 의미하지만, 모든 프레임에 대해 메타데이터를 기술하는 것은 불가능할 뿐만 아니라 불필요하기 때문에 사실상 키 프레임(key frame)이 메타데이터 기술의 대상이 된다.

예에서 키 프레임 추출을 위해서는 Virage사에서 개발한 비디오 주석 시스템인



<그림 10> 예에서 사용된 동영상의 전체 구조



<그림 11> MyLogger를 이용한 키 프레임 추출

MyLogger를 사용하였으며 결과적으로 <그림 11>과 같은 42개의 키 프레임이 추출되었다. 키 프레임이라 할지라도 42개의 키 프레임이 모두 메타데이터 기술의 대상이 되어야 하는 것은 아니며 검색의 필요성이 있는 경우에만 메타데이터를 기술하도록 한다.

다음 예는 시퀀스1과 장면1.3에 대한 메타데이터를 시범적으로 기술한 것이다. 기술의 대상이 되는 자원에는 고유한 식별기호가 부여되어야 하므로 위의 두 자원에 각각 SQ1, SC1.3이라는 ID를 임의로 부여하였다.

```

<?xml version = "1.0"?>
<rdf:RDF xmlns:rdf = "http://www.w3c.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs = "http://www.w3c.org/2000/01/rdf-schema#"
  xmlns:mic = "http://lis.yonsei.ac.kr/enigma93/MovingImageCore#">

  <rdf:Description ID = "SQ1">
    <rdf:type resource = "#Sequence" />
    <mic:title>이산화탄소 소화기로 안 꺼지는 불</mic:title>
    <mic:creator>좋은 나라 소방본부</mic:creator>
    <mic:publisher>좋은 나라 소방본부</mic:publisher>
    <mic:source>http://myhome.naver.com/feublot/netian/asf.htm
    </mic:source>
    <mic:subject>
      <rdf:Bag>
        <rdf:li>화재</rdf:li>
        <rdf:li>실험</rdf:li>
        <rdf:li>이산화탄소</rdf:li>
        <rdf:li>소방과학</rdf:li>
      </rdf:Bag>
    </mic:subject>
    <mic:description> 산화나트륨이나 탄산칼슘 같은 것들은 불에 타면서 산소를 내보내기 때문에 이산화탄소로 불을 덮어봤자 밖에서 들어오는 산소만 없을 뿐, 안에서 생기는 산소로 불이 붙을 수 있기 때문에 이산화탄소 소화기로 불을 끌 수 없다.
    </mic:description>
    <mic:format_medium>
      <rdf:Bag>
        <rdf:li>Asf</rdf:li>
        <rdf:li>window media player</rdf:li>
      </rdf:Bag>
    </mic:format_medium>
    <mic:identifier>http://www.feublot.net/asf/500/co2.asf
    </mic:identifier>
    <mic:media_duration>00:00:58</mic:media_duration>
    <mic:relation_haspart>
      <rdf:Seq>
        <rdf:li>SC1.1</rdf:li>
        <rdf:li>SC1.2</rdf:li>
        <rdf:li>SC1.3</rdf:li>
      </rdf:Seq>
    </mic:relation_haspart>
  </rdf:Description>

```

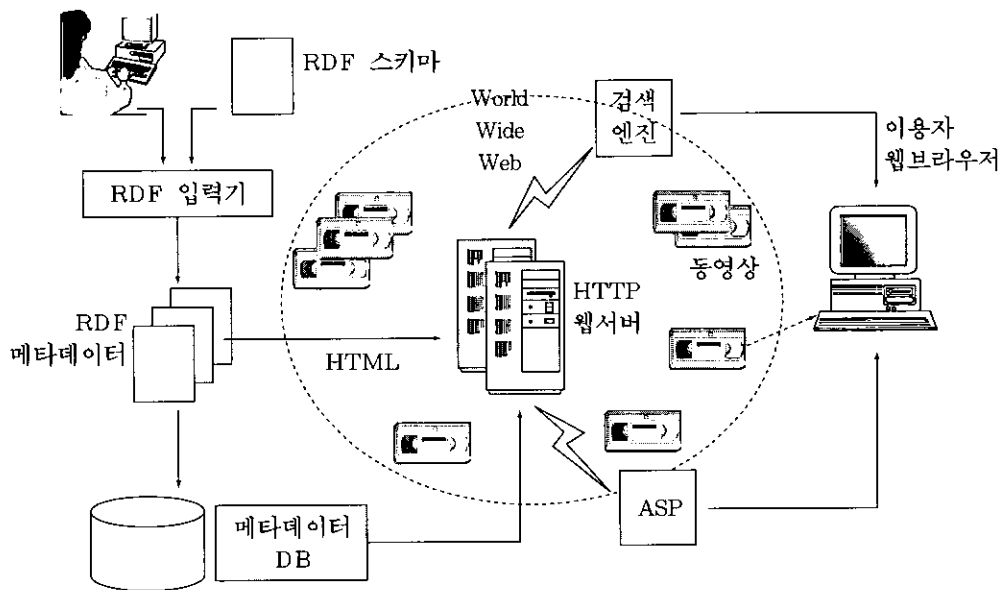
```

</rdf:Description>
<rdf:Description ID = "SC1.3">
  <rdf:type resource = "#Scene">
    <mic:subject>
      <rdf:Bag>
        <rdf:li>소방관</rdf:li>
        <rdf:li>이산화탄소 소화기</rdf:li>
        <rdf:li>진화</rdf:li>
        <rdf:li>진화</rdf:li>
        <rdf:li>화재</rdf:li>
      </rdf:Bag>
    </mic:subject>
    <mic:description>소방관이 이산화탄소 소화기를 이용하여 실험용 화재를 진
      압하고 있다.
    </mic:description>
    <mic:media_time_point>00:00:45:01</mic:media_time_point>
    <mic:media_duration>00:00:13</mic:media_duration>
    <mic:dominant_color>255:255:255</mic:dominant_color>
    <mic:camera_motion>1101</mic:camera_motion>
    <mic:relation>SH1.3.1</mic:relation>
  </rdf:Description>
  
```

5 메타데이터 검색 시스템

메타데이터를 이용하여 웹에 분포되어 있

는 동영상 정보를 검색하기 위한 전체적인 시스템 구성은 <그림 12>와 같다. 메타데이터 작성자가 RDF 입력기를 통해 메타데이터를



<그림 12> 메타데이터 검색 시스템 구성

입력하면, 입력기는 관련된 RDF 스키마를 참조하여 RDF 형식으로 기술된 메타데이터를 생성한다. 생성된 메타데이터는 메타데이터 데이터베이스에 저장되거나 HTML페이지로 생성되어 웹서버에 저장된다. 웹서버에 저장된 메타데이터의 HTML페이지는 자원에 대해 독립적으로 존재하는 메타데이터로서, 이용자는 기존의 검색엔진을 통해 동영상 정보에 접근할 수 있다. 한편 데이터베이스에 저장된 메타데이터는 특정 검색 서비스를 통해 이용자에게 제공되는데, 이때 이용자는 웹브라우저와 메타데이터 데이터베이스와의 연동 기술에 의해 웹에서 데이터베이스 검색이 가능해진다.

본 연구에서는 동영상 정보 검색을 위하여 웹과 연동되는 메타데이터 데이터베이스를 구축하였다. 검색 양식은 이용자의 요구와 능력에 따라 단순검색과 상세 검색의 두 가지 방식을 선택할 수 있다. 필드별 검색을 하지 않고, 단순히 키워드를 입력하는 경우에는 단순검색 양식이 적합하며, 제목이나 저자명, 주제, 그리고 날짜에 따라 검색을 제한하고자 하는 경우에는 상세 검색을 선택할 수 있다. 시각정보는 상세 검색 양식에서 검색할 수 있으며 색상, 질감, 카메라 이동의 측면에서 검색 조건을 지정할 수 있다.

데이터베이스 검색 결과는 <그림 13>과 같다. 검색 결과를 제시하는 화면에서 이용자는



<그림 13> 검색결과 화면 보기

제목이나 정보원을 클릭함으로써 해당 동영상 상에 접근할 수 있다. 제목은 동영상 파일과 직접 연결되며 정보원은 검색된 동영상을 제공하는 홈페이지와 연결된다. 이용자는 제목을 클릭하여 바로 미디어 플레이어와 같은 동영상 재생기를 실행시킴으로써 동영상에 접근하거나, 정보원을 클릭하여 일단 동영상을 제공하는 홈페이지로 이동한 후 이동한 홈페이지 내에서 검색된 동영상에 접근할 수 있다.

6 결 론

네트워크 기술과 동영상 압축 기술의 발달로 인터넷에서 동영상 정보의 이용이 증가되고 있다. 그런데 동영상 정보는 텍스트 정보 위주로 개발된 일반적인 검색엔진을 통해 효과적으로 검색되기 어렵기 때문에 이에 대한 적절한 검색 도구를 필요로 하고 있다. 이 연구는 동영상 정보를 물리적인 구성요소로 분석한 다음 더블런 코어와 MPEG-7의 데이터 요소로 이를 통합하여 RDF 구문으로 기술

하였으며, 메타데이터 데이터베이스를 이용한 동영상 정보의 웹 검색 시스템을 제시하였다. 이를 통해 동영상 정보의 메타데이터는 자원 기술의 단순성과 함께 시청각 정보를 효과적으로 표현할 수 있으며, 웹에 분포된 동영상 정보에 대한 검색 도구이면서 교환 형식으로 더욱 유용하게 사용될 것이다.

향후 연구에서는 청각 정보의 추출 및 검색기술의 개발이 요청되고 있으며, RDF 스키마와 관련하여 다양한 데이터 유형이 구현될 수 있는 방안이 제시되어야 할 것이다. 아울러 메타데이터 기술의 효율성을 향상시키기 위해 자동 추출이 가능한 데이터 요소에 대한 자동화 기술이 개발되어야 할 것이다.

동영상 정보가 지닌 우수한 정보 전달 능력을 고려할 때, 앞으로 각종 영역에서 멀티미디어 정보가 차지하는 비중 및 중요성은 계속해서 높아질 것이고, 이에 따라 동영상 정보에 효과적으로 접근할 수 있는 메타데이터 개발이 요청된다.

참 고 문 헌

- 김우생, 김진웅, 임분철. 1998. "MPEG-7의 표준화 및 내용기반 정보검색" 전자공학회지. 25(8): 772-785.
- Boll, S., W. Klas, and A. Sheth. 1998. "Overview on Using Metadata to Manage Multimedia." In *Multimedia Data Management*, Edited by A. Sheth and W. Klas. New York: McGraw-Hill.
- Böhm, Klemens and Thomas C. Rakow. 1994. "Metadata for Multimedia Documents." *SIGMOD Record*, 23(4): 21-26.
- Grosky, William I. 1994. "Multimedia Information Systems", *IEEE Multimedia*. 1(1): 12-24.
- Hampapur, Arun and Ramesh Jain. 1998. "Video Data Management Systems Metadata and Architecture." In

- Multimedia Data Management*,
 Edited by A. Sheth and W. Klas.
 New York: McGraw-Hill.
- Hunter, Jane and Renato Iannella. 1998. "The Application of Metadata Standards to Video Indexing." [cited 2001-01-09].
 <<http://archive.dstc.edu.au/RDU/staff/jane-hunter/ECDL2/final.html>>
- Hunter, Jane and L. Armstrong. 1999. "A Comparison of Schemas for Video Metadata Representation." [cited 2001-01-09].
 <<http://archive.dstc.edu.au/RDU/staff/jane-hunter/www8/paper.html>>
- Lagoze, Carl. 2000. "Accommodating Simplicity and Complexity in Metadata: Lessons from the Dublin Core Experience." Presented at Seminar on Metadata. [cited 2001-02-12]. <<http://cs-tr.cs.cornell.edu/Dienst/Repository/2.0/Body/ncstrl.cornell/TR2000-1801/pdf>>
- . 2001. "Keeping Dublin Core Simple: Cross-Domain Discovery or Resource Description?" *D-Lib Magazine*, 7(1). [cited 2001-02-12] <<http://www.dlib.org/dlib/january01/lagoze/01lagoze.html>>
- Lassila, Ora and Ralph Swick. 1999. "Resource Description Framework(RDF) Model and Syntax Specification" W3C Recommendation 22 February 1999. <<http://www.w3.org/TR/REC-rdf-syntax-19990222>> 한국어판: 김태수 역. 2001. 자원기술구조 모델과 구분 명세서. *지식처리연구* 2(1): 60-126. <http://dewey.yonsei.ac.kr/courses/metadata/rdf/rdf_model_syntax_spec_kor.html>
- Miller, Eric. 1998. "An Introduction to the Resource Description Framework", *D-Lib Magazine*, May. [cited 2000-09-19] <<http://www.dlib.org/dlib/may98/miller/05miller.html>>