

멀티미디어 내용표현을 위한 메타데이터 비교·분석

The Analysis of Metadata for Representation of Multimedia Content

김 성 희(Kim Seonghee)*

임 은 주(Im Eunju)**

〈 목 차 〉

I. 서론	5. MPEG7
II. 멀티미디어 메타데이터 유형	IV. 멀티미디어 메타데이터 비교·분석
III. 멀티미디어 메타데이터 initiatives	1. 멀티미디어 유형 및 개발 시기
1. Dublin Core	2. 기술 영역 및 수준
2. CDWA	3. 멀티미디어 내용표현
3. VRA	V. 결론
4. DIG35	

초 록

본 연구에서는 멀티미디어 표현 및 기술을 위한 주요 메타데이터 initiatives를 분석하고 문제점을 제시하였다. 그 결과 대부분의 메타데이터 개발은 1999년부터 2004도에 개발되었으며 주로 멀티미디어 내용자체를 표현하기보다는 예술작품 또는 동영상의 서지사항에 기반 한 요소들로 구성되어 있는 것으로 나타났다. 또한 생성방법은 인간이 개입하는 방식을 취하고 있으며 MPEG-7은 멀티미디어 전체를 포괄하면서 표현의 일관성을 유지할 수 있는 포괄적인 표준으로 나타났다.

주제어: 멀티미디어, 메타데이터, Dublin Core, CDWA, VRA, DIG35, MPEG-7

ABSTRACT

This paper provides a state-of-the-art on multimedia metadata. After a brief introduction to multimedia metadata, the paper described and analyzed the main metadata initiatives. The results show that these initiatives described the bibliographic elements rather than metadata for contents representation. On the other hand, MPEG-7 covers a rich and detail aspects for representing the complete and consistent multimedia content.

Key Words: Multimedia, Metadata, Dublin Core, CDWA, VRA, DIG35, MPEG-7

* 중앙대학교 문헌정보학과 부교수(seonghee@cau.ac.kr) (제1저자)

** 중앙대학교 문헌정보학과 대학원 (공동저자)

• 접수일: 2005년 11월 20일 • 최초심사일: 2005년 11월 25일 • 최종심사일: 2005년 12월 5일

I. 서론

최근 비약적인 멀티미디어 기술의 발전과 함께 수많은 정보들이 디지털화 하고 있으며, 다양한 디지털 입력 장치와 저장 장치의 발달로 인해 많은 양의 멀티미디어 콘텐츠가 만들어지고 있다. 또한 인터넷의 확산은 멀티미디어의 양적 팽창을 돕는 데 큰 역할을 하고 있다. 따라서 찾고자 하는 데이터를 인터넷상이나 데이터베이스에서 효율적으로 검색할 수 있는 시스템의 필요성이 증대되고 있다. 특히 멀티미디어 콘텐츠에서 동영상에 차지하는 부분이 많은데, 이는 동영상이 정보 전달력이 가장 강하기 때문이다.¹⁾ 동영상 데이터는 이미지 데이터뿐만 아니라 음성 데이터, 문자 데이터 등이 유기적으로 조직되어 있어 정보의 내용이 풍부하며, 가장 대중적인 매체이기도 하다.

효율적인 검색을 위해 반드시 필요한 요소 중 하나가 검색 대상의 데이터를 어떻게 기술(description)하는가의 문제이다. 특히 멀티미디어의 경우 이미지나 동영상 등의 내용을 바탕으로 특징을 추출하기 때문에 어떠한 특징을 추출한 것인지는 매우 중요하다. 이러한 특징을 기술하는 것이 내용 기반 메타데이터라고 할 수 있다.

멀티미디어의 내용에 기반 한 기술(description)과 검색은 아직 문헌정보학에서는 연구가 많이 이루어지지 않은 분야이다. 데이터의 기술과 검색은 문헌정보학의 핵심적인 부분이며, 멀티미디어 자료가 폭발적으로 증가하고 있는 최근의 상황에서 멀티미디어 자료의 기술과 검색의 중요성은 아무리 강조해도 지나치지 않을 것이다. 텍스트, 그래픽, 이미지 등 여러 미디어의 조합인 멀티미디어는 많은 정보를 전달할 수 있는 매우 효과적이고 대중적인 매체이지만 멀티미디어 데이터의 조직과 검색은 기술(technology)의 부족으로 이용의 효율성이 제한되어 있는 상태라 할 수 있다. 따라서 이러한 멀티미디어 내용을 적절하게 표현하는 것이 필요하다.

그 동안 멀티미디어의 표현 및 검색 방법에 대한 연구들은 많이 수행되어 왔다.

국내에서의 선행연구들을 살펴보면 이미지와 동영상을 기술하기 위한 메타데이터에 대한 연구는 있었지만 내용 기반 검색을 할 수 있는 메타데이터에 관한 연구는 부족한 실정이다. 대부분의 내용 기반 검색 관련 연구는 이미지와 동영상의 특징 추출이나 특징들의 결합 알고리즘에 대한 연구 등 공학적인 측면에서 수행된 연구들이다. 또한 현재 범용적으로 사용되고 있는 메타데이터의 이미지나 동영상에의 적용 가능성에 대해 연구한 논문들도 거의 없는 실정이며, 기술(description)을 위한 MPEG-7을 구체적으로 응용할 수 있는 연구가 거의 이루어지지 않았다. 또한 대부분의 내용기반 이미지 및 동영상에 관한 연구는 주로 내용 기반 검색 효율성과 관련한 다양한 색인 알고리즘 및 특징 추출 방법에 대한 연구이며 멀티미디어 콘텐츠 메타데이터를 체계적으로 연구한 논문은 거의 없는 실정이다.

따라서 본 연구에서는 내용 기반 검색을 위한 멀티미디어 콘텐츠 메타데이터에 대해 조사 분석

1) 강행봉, "내용 기반의 비디오 검색 및 브라우징," 제 2회 디지털도서관 컨퍼런스 논문집(1999 11), pp.117-126.

하고자 한다. 선정된 메타데이터는 Dublin Core, CDWA, DIG35, MPEG-7이다. 이러한 내용 기반 메타데이터를 비교 연구함으로써 멀티미디어의 효율적인 검색과 이용에 대한 효과를 기대할 수 있을 것이다.

II. 멀티미디어 메타데이터 유형

멀티미디어는 텍스트, 그래픽, 이미지, 애니메이션, 오디오, 동영상 등의 두 가지 이상의 미디어 어로 구성되며 멀티미디어 유형은 크게 이미지, 오디오, 비디오, 그래픽, 멀티미디어로 나눌 수 있으며 구체적인 종류는 <표 1>과 같다.

<표 1> 멀티미디어의 유형

유형	종류
이미지	photographs, prints, maps, manuscripts, documents, drawings, paintings, movie stills, posters
오디오	songs, music, plays, interviews, oral histories, radio programs, speeches, lectures, performances, language recordings
비디오/필름	full feature films, documentaries, news' clips, anthropological/expedition footage, home movies, animation
그래픽스	3D models, decorative designs and patterns, simulated walk-throughs of buildings, archeological sites, VRML
멀티미디어	presentations, slide shows, SMIL files, QuickTime VR

이처럼 멀티미디어는 여러 타입의 데이터 자체, 혹은 데이터들의 조합을 말한다. 이러한 멀티미디어는 컴퓨터 기술의 발전에 따라 더욱 그 영역이 넓어지고 다양해질 것으로 예상된다.

컴퓨터의 처리 대상으로 보면 이미지, 오디오, 동영상 등의 데이터는 크기도 크고 처리도 복잡하지만 인간의 입장에서 보면 다양한 미디어를 혼합해 여러 가지 시각적 효과와 청각적 효과를 느낄 수 있다. 따라서 이러한 멀티미디어를 효율적으로 검색하기 위해서는 멀티미디어 내용을 추출하여 표현하는 것이 중요하다.

멀티미디어 메타데이터에 대한 사용이 중요하게 대두되고 있는 데에는 크게 두 가지 요인이 있다. 하나는 복잡한 멀티미디어 정보를 효율적으로 관리하기 위한 방안으로서 메타데이터를 사용한다는 것이고, 다른 한 가지는 이종의(heterogeneous) 시스템들 사이에서 야기되는 데이터의 불일치(inconsistency)문제를 해결하기 위한 방법으로서 메타데이터를 사용한다는 것이다. 이러한 메타데이터는 기존의 전통적인 데이터가 아닌 이미지나 동영상 등과 같은 멀티미디어 데이터에 대해 내용 기반 검색을 가능하게 한다.

그동안 여러 학자들에 의해서 멀티미디어 메타데이터를 분류되어 왔다. 그 중 대표적인 것을 살펴보면 먼저 del Bimbo는 멀티미디어 메타데이터는 크게 내용독립메타데이터(content-independent), 내용에 대한 메타데이터(content-dependent), 기술적 메타데이터(content-descriptive) 로 구분하였다.²⁾ 구체적으로 살펴보면 다음과 같다.

- (1) 내용 독립 메타데이터 (content-independent): 멀티미디어 내용과 직접관련이 없는 부분으로 멀티미디어에 나타난 이미지 및 동영상에 대한 데이터이다. 예를 들면 이미지 이름, 타이틀, 날짜 등이 여기에 속한다. 이 부분은 수작업으로 수행되어야 하는 부분이며 이용자들이 멀티미디어 데이터를 검색할 때 사용하는 요소이므로 매우 중요하다.
- (2) 내용의존 메타데이터(content-dependent): 이는 이미지 및 동영상의 내용의 특징을 의미하는 것으로, 예컨대 칼라, 질감, 모양, 움직임 등이 여기에 해당된다. 컴퓨터에 의해 자동으로 이런 특징들을 추출할 수 있다.
- (3) 내용기술 메타데이터(content-descriptive): 멀티미디어의 내용의 주제를 설명하는 요소로써 분석자에 따라 상당히 주관적일 수 있다. 컴퓨터에 의해 자동으로 추출하기 어려운 분야라 할 수 있다.

Hunter³⁾는 멀티미디어 자원과 관련된 메타데이터를 서지 메타데이터, 포맷 메타데이터, 구조 메타데이터, 내용 메타데이터, 사건과 저작권 메타데이터의 5가지 카테고리로 분류하였다.

- (1) 서지 메타데이터(bibliographic metadata) - 멀티미디어 자원의 창작/생산(시간, 장소), 창작과 관련된 개인이나 기관(프로듀서, 디렉터, 캐스트), 자원의 주제 정보와 관련된 메타데이터이다. 제목, 초록, 주제, 장르 등이 이 메타데이터에 속한다.
- (2) 포맷 메타데이터(formatting metadata) - 멀티미디어 자원의 포맷, 인코딩, 저장매체와 시스템의 사양 등에 관한 정보를 포함한다. <표 2>는 각각의 매체 유형에 따른 포맷 메타데이터를 보여준다. 파일 사이즈, 버전, 차원, 해상도, 화면비율, 압축률, 시스템 요구사항 등이 포맷 메타데이터에 속한다.

구체적으로 살펴보면 다음과 같다.

2) Di Bono, M. G., Pieri G. and Salvetti O. *WP9: A Review of Data and Metadata Standards and Techniques for Representation of Multimedia Content. Multimedia Understanding through Semantics* (Computation and Learning, 2004).

3) Hunter J, *MetaNet - A Metadata Term Thesaurus to Enable Semantic Interoperability Between Metadata Domains*. 2002 <<http://jodi.ecs.soton.ac.uk/Articles/v01/i08/Hunter/>>. [cited 2005. 10.10]

<표 2> 매체 유형에 따른 포맷 메타데이터

Image	Video	Text	Multimedia
format (image/tiff)	format (Quicktime, MPEG1)	format(text/msword)	format (text/html)
filesize(bytes)	filesize(bytes)	filesize(bytes)	filesize(bytes)
version(v 4.0)	version(v 1.1)	version = 97	version(v 3.0)
resolution(600dpi)	dimensions(640x480)	compression(zip)	software (MS FrontPage 2.0)
dimensions(1024x768)	aspect ratio(4:3)	charset(Unicode)	storage type (HD serve)
aspect ratio(4:3)	duration (32min 12sec)	template(summary.dtd)	template(Program.xml)
colourdepth(8-bit grayscale, 24-bit colour)	compression(mp2)		bandwidth requirements
colourpalette(CMYK, RGB, GrayScale)	encoding(mp2)		system requirements (OS, software, hardware, peripherals)
framerate(25fps)	sound(Yes/No)		
colourLUT(base64)	storage type(DVD)		
orientation (Portrait, Landscape)	colour (Colour or B/W)		
compression(CCIT4)	special Effects (ChromaKey)		
storage type(CD-ROM, Jazz, hard drive)	delivery and presentation requirements(bandwidth, operating system, hardware)		
	software, peripherals		
scanner (make, model, serial #)	camera details and settings (make, model, serial #, aperture, focallength, filter)		

- (3) 구조 메타데이터(structural metadata) - 멀티미디어 자원의 시간적, 공간적, 시공간적 세그먼트(컨텐츠, 샷, 프레임, 이미지 영역)와 이러한 세그먼트들 간의 관계를 다루는, 구조적 분해에 대한 메타데이터이다.
- (4) 내용 메타데이터(content metadata) - 멀티미디어 자원에 기록되어 있는, 혹은 설명하고 있는 실제적인 내용에 대한 색인을 다룬다. 내용 메타데이터는 사람, 객체, 장소나 사건을 묘사하는 자연어에서부터 컬러 히스토그램이나 볼륨과 같은 저수준 오디오, 비디오 특성에 이르기까지 매우 다양하다.
- (5) 사건과 저작권 메타데이터(Events and rights metadata) - 이 메타데이터는 자원의 라이프 히스토리(life history)를 설명한다. 입수와 재배포에서부터 리포매팅(reformatting), 편집, 재포장, 배포, 메타데이터 속성, 사용, 저작권 동의에 이르기까지의 모든 것을 포함한다.

Di Bono는 멀티미디어 메타데이터는 데이터 기술의 수준(level), 자동생성(producibility), 영역

의존성(domain dependence)에 따라 나눌 수도 있다고 하였다.⁴⁾

- (1) 기술 수준 : 멀티미디어 콘텐츠가 기술되는 저수준 측면의 기술적 수준(technical level)과 멀티미디어 콘텐츠가 다루는 내용에 대한 고수준의 추상적인 의미적 수준(semantic level)로 구분할 수 있다.
- (2) 자동 생성 : 메타데이터의 생성은 저수준의 기술(technical) 메타데이터에서 사용되는 경제적 측면에서는 바람직한 특성이 자동적 생산이 가능한 메타데이터와 멀티미디어 콘텐츠에서 묘사하는 의미를 다루며 인간의 지식을 필요로 하는 의미적 메타데이터 두 가지로 나눌 수 있다. 후자의 경우에 메타데이터 생성은 수동으로 행해져야 한다.
- (3) 영역 의존성 : 메타데이터는 영역 의존적일 수 있다. 예를 들면, 대부분의 응용 프로그램에서는 이미지의 칼라 분포가 유용하게 쓰이지만 의학을 다루는 응용분야에서는 종양의 위치가 더 중요하다. 또 메타데이터는 미디어 타입에 의존적일 수도 있다. 예를 들면 창작관련 데이터는 모든 미디어에 적용 가능하지만 컬러분포는 시각적 미디어에만 적용할 수 있다.

Shatford⁵⁾는 멀티미디어 내용을 표현하기 위한 메타데이터를 3가지 수준으로 구분하고 있다. 이는 Erwin Panofsky⁶⁾ 이론에 기초를 둔 것으로서 크게 pre-iconographic level, iconographic level, iconologic level으로 구분하고 있다.

- (1) pre-iconographic level: 이 수준은 이미지의 주제가 무엇인지를 기술하는 것으로 사실적(factual) 수준과 표현적(expressional)수준으로 구분할 수 있다. 사실적 주제로는 오브젝트, 액션, 이벤트에 대해 일반적으로 기술하는 것으로서 여자, 어린이, 울고 있는 것(crying) 등이 여기에 해당될 수 있다. 반면에 표현적 주제(expressional)는 어떤 이미지에 나타나는 감정이나 분위기(mood)를 기술하는 것으로 “평화로움”, “우울함”, “즐거움” 등이 여기에 해당된다.
- (2) iconographic level: 이 수준은 어떤 이미지의 의미를 부여하는 좀더 구체적인 요소이다. 예를 들면 여자라고 일반적으로 표현하는 대장면에 성모마리아를 표현한다든지 모자를 벗어 위로 들어올리는 행위 보면서 서양의 인사스타일이라고 이해하고 의미를 부여하려면 먼저 서양의 문화를 이해해야 하는 것이다.
- (3) iconologic level: 앞에서 설명한 2개의 수준의 복합된 형태이다. 즉 어떤 이미지에 대한 구체적이고 깊이 있는 이해 및 해석부분이 여기에 해당된다. 이 요소를 기술하기 위해서는 이미

4) Di Bono, M. G., Pieri G. and Salvetti O., *Ibid.*

5) Shatford, Sara. "Analyzing the subject of a picture." *Cataloging & Classification Quarterly*, Vol. 6, no. 3 (1986), pp.39-62

6) Panofsky, Erwin. *Iconography and iconology: An introduction to the study of Renaissance art*. In *Studies in iconology: Humanistic themes in the art of the Renaissance*(New York: Oxford University Press. 1939)

지가 포함하고 있는 사회적 문화적 배경을 철저하게 이해하고 있어야 함을 의미한다. Shatford는 이상의 Panofsky의 이론을 정보학 분야에서 유사하게 적용하고 있다. 즉, Pre-iconographic의 주제의 사실적 묘사는 기술적(descriptive)인 것과 표현적인 의미로 구분하고 있다. 즉 기술적인 주제는 이미지는 구체적으로 무엇을 나타내고 있는지("of")에 대한 수준이며 따라서 지식은 상대적으로 객관적일 수 있다. 한편 표현적 주제는 이 그림은 일반적으로 무엇에 대한 것 즉, 이미지가 우리에게 이야기하고자 하는 것이 무엇인지("about")에 대한 것으로 정의하고 있다. 따라서 이 표현적 수준인 "aboutness"는 좀더 주관적일 수 있다. 따라서 Shatford가 제시하고 있는 "of"와 "about"은 Panofsky의 iconographic 수준을 기술적 일반적 수준(descriptive and generic layer)으로 나누고 있다. Shatford는 이러한 수준의 구분을 객관적 주관적으로 구분하기 보다는 오히려 객관적 또는 상징적 의미로 구분하려 하고 있다. 예컨대 "of"의 수준은 구체적 이라기보다는 일반적 으로 표현하는 것이다. 예를들면 Ms Jensen이라고 구체적으로 표현하기보다는 "여성"이라고 일반적으로 표현하고 있으며, "about"은 상징적인 것으로 예컨대 "Ms. Jensen"은 장면 화적 캐릭터, 또는 추상적 인격(abstract personification)으로 나타낼 수 있다. Shatford는 이상의 3개의 기술 수준 중에서 iconologic 수준은 너무 주관적이고 추상적이어서 일관성 있는 방법으로 기술하는 것이 불가능하다고 생각하고 있기 때문에 여기서는 구분대상에서 제외시켰다. 이상의 내용을 표로 나타내면 <표 3>과 같다.

<표 3> Panofsky와 Shatford의 이미지 분류

	Panofsky	Shatford
-	-	-
descriptive level:	pre-iconographic, factual	generic "of"
	iconographic	specific "of"
derived level:	pre-iconographic, expressional	generic "about"
	iconographic	specific "about"
interpretative level:	iconological	-

이러한 멀티미디어 메타데이터들은 기존의 서지 정보뿐만 아니라 미디어 자체에 대한 정보, 미디어의 구조에 대한 정보, 내용에 대한 정보 등 다양하고 확장된 정보들을 담고 있음을 알 수 있다.