

인체영상 어노테이션 DB 설계에 관한 연구

A Study on Design of Annotation Database for Visible Human

안부영, 이승복, 한 건, 이상호
한국과학기술정보연구원(KISTI)

Ahn bu-young, Lee seung-bock, Han Geon, Lee sang-ho
Korea Institute of Science and Technology
Information

요약

최근 들어 정보기술과 컴퓨터 네트워크가 급속도로 발전하면서 디지털 콘텐츠의 보급이 증가하고 있다. 디지털 콘텐츠는 멀티미디어 형태의 2, 3차원 정보로 표현되며 이 가운데 인간의 몸을 촬영하여 데이터베이스를 구축한 인체영상 데이터베이스는 다양한 분야에 활용될 수 있는 주요한 디지털 콘텐츠이다. 한국과학기술정보연구원(KISTI)에서는 한국인의 인체영상, 골격형상, 인체물성, 인체모델 등 다양한 인체정보를 현재 구축 중에 있으며 이러한 정보들을 인터넷을 통해 공개하고 있다. 그러나 인체영상의 각 이미지에 관한 설명자료는 제공되고 있지 않아 의학영상 전문가가 아니면 각 이미지에 대한 세부 내용을 알 수가 없다. 이에 본 논문에서는 인체영상 데이터의 정보접근 효율성을 향상시키고자 한국인의 인체영상 데이터베이스를 대상으로 하여 이미지에 관한 설명과 특이사항을 정리하여 이미지와 텍스트가 연계 가능한 어노테이션 데이터베이스 및 검색 인터페이스를 설계하였다. 이를 통해 보다 접근성이 좋은 인체영상 데이터베이스의 개발과 함께 데이터 활용이 더욱 촉진되리라 기대된다.

Abstract

As the IT and computer network technology is developed very rapidly, the quantity of digital contents is increased and disseminated more widely. The digital contents is generally expressed in 2 or 3 dimensional multimedia format and the visible human image that is taken from human body is very important because of its variety of usefulness. The KISTI(Korea Institute of Science and Technology Information) is now constructing various Korean human informations such as visible Korean, digital Korean, human bone property and human models. These informations are accessible through the internet. However, these human images are not easily understandable for general users because they are specialized in medical image field and there is no detailed explanation data. In this study, we designed the annotation database and searching interface for KISTI's visible Korean database. This annotation database involved the detailed explanation and special note of visible Korean data and it can connect the image and text data of visible Korean with each other. By studying this database and interface design, the KISTI's visible Korean database is more easily accessible and understandable to general users and it can promote the usage of visible Korean data more widely.

I. 서론

최근 들어 정보기술이 급속도로 발전하면서 디지털 콘텐츠의 중요성이 더욱 부각되고 있으며, 디지털 콘텐츠의 하나인 의료영상 또한 급속도로 양이 증가하고 발전하였다. 의료영상 분야에서 사용하는 장치로는 CR(Computed Radiography), DR(Digital Radiography), 컴퓨터 단층촬영(CT), 자기공명영상(MRI) 등 다양하다[1]. 이런 의료영상 장치를 이용하여 인간의 몸을 촬영하여 데이터베이스로 구축한 인체영상 데이터베이스는 의학 및 의학산업의 발전에 없어서는 안되는 중요한 데이터베이스이다.

미국의 NLM에서는 1994년부터 Visible Human Project를 수행하고 있으며 우리나라는 한국과학기술정보연구원(KISTI)에서 2000년부터 2002년까지 한국인의 인체영상 데이터베이스를 구축하여 서비스하고 있다.

그러나 디지털 콘텐츠인 인체영상에 관한 고화질의 다양한 이미지는 제공되고 있으나 각 이미지에 관한 부가적인 설명자료는 제공되고 있지 않아 의학영상 전문가가 아니면 각 이미지에 대한 세부 내용을 알 수가 없게 되어 있다. 이미지에 관한 세부내용을 표현할 수 있는 도구로는 메타데이터와 어노테이션이 있다. 메타데이터는 데이터에 관한 데이터이고, 어노테

이션은 콘텐츠에 관한 콘텐츠라고 말할 수 있다[2]. 이에 본 논문에서는 인체영상의 많은 이용자들의 정보접근 효율성을 향상시키고자 한국인의 인체영상 데이터베이스를 대상으로 이미지와 텍스트의 연계 가능한 어노테이션 데이터베이스와 검색 인터페이스를 설계하고자 한다.

II. 인체영상 데이터베이스

1. Visible Human Project

미국의 국가기관인 NLM(National Library of Medicine's)은 1994년에 사람 몸의 연속절단면의 영상(Visible Human)을 만들었다. Visible Human은 사람 몸을 연속절단해서 나타나는 CT, MRI 및 해부 영상이며, 전세계에서 의학영상 분야의 기본이 되는 데이터베이스로 사용되고 있다.

그러나 미국의 Visible Human 데이터베이스는 미국 백인 남성을 대상으로 만들었기 때문에 한국인한테 적용하기 힘들다. 왜냐하면 얼굴의 생김새는 인종과 민족마다 다를 뿐만 아니라 몸 속에 있는 기관의 생김새도 인종과 민족마다 다르기 때문이다. 또한 해부 영상의 간격(남자: 1 mm, 여자: 0.33 mm)이 넓고 해상도가 낮은 단점이 있다.

2. Visible Korean Human

한국과학기술정보연구원(KISTI)은 2000년-2002년 한국인의 인체영상 데이터베이스(VKH, Visible Korean Human)를 구축하여 서비스 하고 있다.



▶▶ 그림 1. Visible Korean Human 웹사이트[5]

한국인의 인체영상 데이터베이스는 의학, 기계공학, 사진학, 컴퓨터학 등의 학·연 연구로 수행하여 연속절단면 영상 데이터베이스를 구축하고 3차원으로 가시화하여 인터넷을 통해 제

공함으로써 국내 의학, 의학 산업 및 과학기술 발전에 기여하고 있다.

인체영상 데이터베이스 구축을 위해 70대와 40대의 남자 시신을 기증받았으며, 0.2mm 간격으로 연속절단하여 해부영상과 구역화 영상을 제작하여 정확하고 높은 해상도의 이미지를 보여주고 있다. 또한 한국인의 인체영상 데이터베이스는 국내외 산·학·연 60여개 관련 기관에서 데이터를 제공받아 각종 연구 및 프로젝트에 활용하고 있다.

III. 어노테이션 데이터베이스

1. 어노테이션 개요

1.1 어노테이션의 정의

논문 또는 보고서를 작성할 때 대부분의 저자들은 중요하게 설명하고자 하는 부분에 주석(각주, 미주 등)을 표시하여 부연 설명을 하게 된다. 또한 그 논문 또는 보고서를 읽는 사람들은 중요 부분에 밑줄을 긋거나 부가적인 설명을 적는 경우도 있다. 이처럼 원본 문서의 특별한 부분에 추가되는 부가정보와 이러한 부가정보를 추가하는 행위[2]를 어노테이션이라고 부른다.

어노테이션은 문헌정보에서 메타데이터(데이터의 데이터)를 상호 보완하는 의미로 사용되고 있지만 인체영상과 같은 디지털 콘텐츠에 있어 각 영상에 대한 부가설명을 위해서도 사용 가능하다. 부가설명을 위해 하는 행위도 꼭 텍스트로 주석을 다는 것 뿐만 아니라 밑줄과 도형 등의 심볼 및 동영상 자료를 이용하기도 한다. 그래서 메타데이터를 데이터의 데이터라고 하는 것처럼 어노테이션을 콘텐츠의 콘텐츠라고 한다.

1.2 어노테이션의 필요성

메타데이터는 각 분야의 원본 데이터를 설명하는데 없어서는 안될 중요한 정보로 자리잡았다. 원본 데이터를 직접 접근하거나 수정할 필요없이 메타데이터만을 가지고 원본 데이터에 관한 정보에 접근하고 원본 데이터에 대한 정보를 업데이트 할 수 있다. 요즘은 메타데이터만을 가지고 새로운 응용 콘텐츠를 개발하고 있는 실정이다. 예를 들면 세계생물다양성정보기구(GBIF)에서는 1억건이 넘는 세계 각지에 흩어져 있는 생물다양성 데이터의 메타데이터만을 서비스하고 있다. 이 메타데이터를 가지고 google-earth를 이용하여 생물다양성 분포정보를 제공하는 응용 소프트웨어를 개발하여 이용자들에게 큰 호응을 얻고 있다.

이런 메타데이터처럼 어노테이션 또한 콘텐츠를 부연 설명하는데 있어 필수적인 정보이기에 향후 어노테이션 정보만을 가지고 다양한 응용 소프트웨어 개발이 이루어질 것이라고 생

각된다. 지금도 어노테이션은 문헌정보분야 뿐만 아니라 생물정보학, 멀티미디어 분야 등에서도 많이 활용되고 있다. 미국 국립보건원(NLM)의 NCBI(National Center for Biotechnology Information)에서는 2005년도에 이미 유전자 주석(annotation) 데이터베이스를 완성하여 제공하고 있으며, 영화, 비디오, 음악 등의 멀티미디어 분야에서도 어노테이션을 이용하여 각 정보의 부가가치를 높이고 있다.

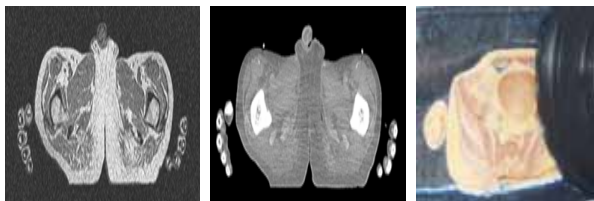
이에 KISTI에서 제공하고 있는 인체영상 데이터베이스의 부가가치를 높이고 이용자들에게 편리성을 제공하기 위해서는 어노테이션 데이터베이스를 설계하여 연계하여 제공하는 것이 바람직하다고 사료된다.

2. 어노테이션 데이터베이스 설계

2.1 인체영상 데이터베이스 구성

인체영상 데이터베이스는 절단면 영상, CT 영상, MRI 영상 등으로 구성되어 있다. 모든 영상 데이터는 편집 및 정렬과 피부, 뼈, 근육, 창자, 심장, 뇌, 척수 등의 기관을 구역화하여 데이터베이스로 구축하였다.

이렇게 구축된 인체영상 데이터베이스를 바탕으로 2차원 가시화 영상 소프트웨어를 개발하여 해부학적 이미지(수평, 수직, 비스듬)와 가상해부 이미지(피부, 뼈, 장기) 등을 인터넷(<http://vkh3.kisti.re.kr>)을 통하여 이용자들에게 제공하고 있다. <그림 2>는 시신을 기증받아 절단된 면을 MRI, CT, 카메라 등으로 촬영한 영상들이다.



▶▶ 그림 2. MRI/CT/절단면 촬영 영상

2.2 인체영상 어노테이션 데이터베이스 스키마

본 논문에서는 어노테이션 데이터베이스 스키마를 설계하기 위하여 다양한 정보와 웹사이트에 수록된 공개된 자료와 전자매체까지도 표현이 가능한 MODS(Metadata Object Description Schema)와 각 분야에서 메타데이터의 표준으로 자리잡고 있는 DC(Dublin Core)를 기본으로 어노테이션 데이터베이스 스키마를 구성하였다. 설계된 스키마를 적용한 어노테이션 데이터베이스를 구축하여 운영하면서 문제점이 발생하면 지속적으로 개선해 나갈 예정이다.

[표 2] MODS와 DC 요소 비교

MODS	한글명	DC	한글명
titleInfo	서명사항	Title	표제
name	저자사항	Creator Contributor	제작자 및 기타기여자
typeOfResource	자료유형	Type	유형
genre	장르		
originInfo	발행사항	Publisher Date	발행처 날짜
language	언어	Language	언어
physicalDescription	형태사항	Format	표현형식
abstract	초록	Description	내용기술
tableOfContents	목차	Description	내용기술
targetAudience	이용대상	Audience	이용대상자
note	주기	Description	내용기술
subject	주제	Subject	주제
classification	분류	Subject	주제
relatedItem	관련정보	Relation	관계
identifier	식별기호	Identifier	식별기호
location	소장위치		
accessCondition	이용제한	Rights	권리관리
extension	로컬정보		
recordInfo	레코드정보		

본 논문에서 설계된 어노테이션 데이터베이스는 인체영상 데이터 획득을 용이하게 이용자들을 지원하는데 목적이 있다. 그래서 이용자 누구라도 용이한 접근이 가능하도록 <표 3>과 같이 스키마를 단순화하여 설계하였다.

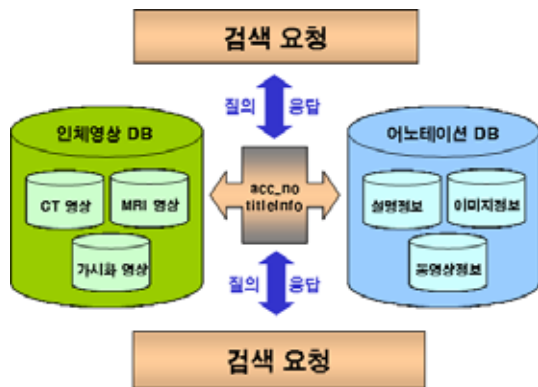
[표 3] 인체영상 어노테이션 DB 스키마

엘리먼트 명칭	설명
acc_no	영상 고유번호
titleInfo	영상 제목 정보
name_Creator_Contributor	저자(제작자 및 기여자) 정보
typeOfResource	자료 유형 정보
originInfo_Publisher_Date	발행 사항(제작일) 정보
language	언어 정보
physicalDescription_format	형태 사항 및 표현 형식 정보
abstract_description	초록, 내용 등의 정보
targetAudience	이용 대상 정보
note	주기, 특이사항 정보
subject	주제 정보
classification	분류 정보
relatedItem	관련(관계) 정보
identifier	식별자 정보
location	소장 위치 정보
accessCondition_right	이용제한 및 권리 정보
extension	확장 가능한 필드 정보
recordInfo	일반적인 레코드 정보

3. 인체영상 검색 인터페이스 설계

3.1 인체영상 데이터베이스 검색 구조

인체영상 데이터베이스와 어노테이션 데이터베이스는 다음 그림과 같이 영상 고유번호(acc_no, primary key)와 영상 제목(titleInfo, secondary key)을 키로 연결되며, 이 키를 통하여 이용자가 질의한 영상에 관한 설명(부가정보)을 불러와서 응답하게 된다.

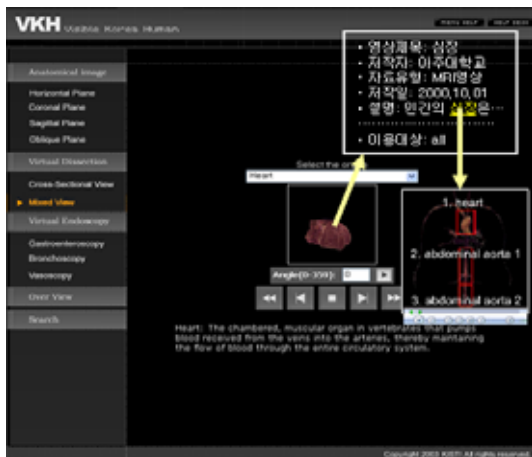


▶▶ 그림 3. 데이터베이스 검색 구조

3.2 인체영상 데이터베이스 검색 화면

기존의 인체영상 웹사이트를 수정하지 않고 어노테이션만 추가하여 <그림 4>와 같은 인터페이스를 설계하였다. 그림에서와 같이 사용자가 인체의 장기중 심장을 검색하면 심장에 관한 설명자료를 불러오고, 불러 온 설명자료에서 심장이라는 단어를 클릭하면 기존에 구축되어 서비스중인 심장에 관한 동영상 파일을 윈도우미디어로 재생시켜 볼 수 있다.

더불어 의학서적에 게재된 인체와 관련된 이미지(심장, 간, 뇌 등)를 데이터베이스로 구축한 뒤에 사용자가 검색한 부분의 이미지를 보여주는 인터페이스 또한 추가하여 좀 더 많은 정보를 제공하도록 설계하였다.



▶▶ 그림 4. 인체영상과 어노테이션 검색화면

입을 구현하여 이용자의 피드백을 거쳐야 하기에 아직은 완벽한 설계라고 단정지를 수는 없다. 그러나 인체영상 어노테이션 데이터베이스와 검색 인터페이스는 전문적인 인체영상정보를 일반 이용자로 쉽게 검색하고 내용을 획득할 수 있기에 전문가 뿐만 아니라 학생, 교사를 포함한 일반이용자들의 학습 콘텐츠로도 이용이 가능하리라 사료된다.

향후 검색을 돕기 위한 인체영상 관련 용어사전 및 어노테이션 데이터베이스의 메타데이터를 작성하여 본 시스템에 추가하는 연구를 수행하여야 한다. 본 논문의 설계를 바탕으로 시스템이 구현되어 이용자들에게 제공된다면 인체영상 이미지를 검색하고 설명을 찾기위해 여러 사이트를 방문해야 하는 이용자들의 번거로움을 해결하고 정확한 인체영상 관련 정보를 획득하는데 기여할 수 있으리라 기대된다.

참 고 문 헌

- [1] 박병래, 신용원, “방사선 의료영상 검색시스템에 관한연구”, 방사선기술과학 28(1): 19-24, 2005년
- [2] 박승진, “디지털 콘텐츠의 어노테이션에 관한 연구”, 한국문헌정보학회지 40(4): 267-286, 2006년
- [3] 안형근, 고재진, “MPEG-7기반 의미적 메타데이터 모델을 이용한 멀티미디어 주석 및 검색시스템의 개발”, 정보처리학회논문지 14(D): 573-584, 2007년
- [4] Visible Human Project 홈페이지, <http://www.nlm.nih.gov/research/visible/>
- [5] 한국인의 인체영상 홈페이지, <http://vkh3.kisti.re.kr>

IV. 결론 및 향후 연구방향

본 논문에서는 인체영상에 관한 많은 이용자들의 정보접근 효율성을 향상시킬 수 있는 한국인의 인체영상 데이터베이스의 이미지와 텍스트의 연계 가능한 어노테이션 데이터베이스와 검색 인터페이스 설계에 관하여 기술하였다.

본 논문에서 설계한 데이터베이스와 인터페이스는 프로토타