

다해상도용 영상 데이터 관리 시스템 설계

김 성재, 조 승호
강남대학교 지식정보공학부

Design of Management System for Multiresolution Image Data

Seong Jae Kim, Seung Ho Cho
School of Information and Knowledge Engineering, Kangnam University
E-mail : {kimsj, shcho}@kangnam.ac.kr

요 약

본 논문은 광학 현미경으로 관찰된 데이터들을 분산 시스템이나 병렬 시스템에 구현한 소프트웨어 시스템의 설계에 대한 것으로, 이 시스템이 처리하는 데이터들이 대용량이라는 특성과 함께 다중 해상도의 특성을 갖는다. 본 시스템은 고객/서버 모델을 기반으로 하였으며, 대용량 데이터 처리시 성능에 중요한 디스크 입출력의 대역폭을 높이기 위해 힐버트 곡선 기반의 분산 알고리즘을 적용하였다. 서버부는 조정자 노드와 서비스 노드로 구성되며, 시스템의 제 구성 요소들간에는 정해진 통신 규약에 따라 메시지를 주고 받고, 상호 독립적이다. 이 시스템은 의학 교육, 원격 병리, 가상 학술 회의 등에 응용될 때 활용 가치가 높을 것으로 기대된다.

1. 서론

컴퓨터의 성능이 향상되고 센서가 고기능화 됨에 따라 대규모 시뮬레이션이나 실험 측정치들이 비교적 장기간에 걸쳐 수집 가능해지고 있다. 데이터 측정이나 컴퓨터 시뮬레이션으로부터 데이터를 수집하는 과학자나 엔지니어의 주목적은 당면한 문제를 이해하고 확인 및 검증해 보고자 하는 것이다. 이들은 시뮬레이션이나 실험에 의해 수집된 데이터를 분석해 봄으로써 문제에 대한 이해를 하는데 도움을 받게 되는 것이다. 또한, 고성능 CPU와 대용량 디스크들을 탑재한 상용 컴퓨터들이 네트워크 또는 내부 스위치[1]로 연결되면서 대용량의 과학 데이터들을 분산하여 저장 및 처리하는 것이 가능해졌다.

일반적으로 고객(client)은 데이터 집합(dataset)으로부터 관심 영역을 정의하고 서버에게 정보를 보내 서비스를 요청한다. 서버는 데이터 집합으로부터 해당 데이터를 추출하고 계산 또는 변환을 통해 새로운 데이터 출력을 생성한다. 이렇게 서버로부터 생성된 데이터는 다른 프로그램이나 고객에게 전송된다. 많은 과학용 데이터들은 다차원적(multidimensional)으로 구성되는데, 이들은 공간 좌표, 시간, 또는 온도, 속도,

자기장과 같은 가변적인 실험 조건 등으로 표현된다.

이러한 관심 영역에 대한 참조는 다차원적인 영역 박스(bounding box)을 표현하는 영역 질의(range query)를 통해 이루어진다. 따라서 이 경우 다차원적 영역 박스내 데이터들만이 선택되어 검색되고 처리된다[2,5,9,14].

본 시스템은 환자의 조직 샘플에 대한 고해상도 디지털 이미지(high resolution digital image)를 인터넷상에서 제공할 목적으로 궁극적으로는 의과 대학, 의료 기관 및 의학 연구 기관의 조직 샘플들로 저장된 캐비넷들을 디지털 저장 시스템(digital storage system)으로 대체하는 것이 목적이다. 이들 이미지 데이터 집합에 대해 본 시스템의 고객은 마치 현미경 사용자가 관심있는 샘플을 선택한 후 이를 통해 실제 샘플을 관찰하듯이 컴퓨터 상에서 조직 샘플들을 관찰할 수 있게 제공한다. 따라서 컴퓨터가 인터넷에 접속되어 있기만 하면 언제 어디서든 의사나 연구원들은 원하는 조직 샘플들을 검색할 수 있다.

이외에 실제 현미경에서는 제공할 수 없고 본 시스템만이 얻을 수 있는 이점으로는 첫 번째, 의학 교육 분야이다. 예를 들면 의과 대학 강의실에서 수업중의 모든 학생들이

이 시스템을 통해 동시에 똑같은 샘플을 보면서 교수의 설명을 들을 수 있으므로 강의의 보다 효과적으로 진행할 수 있다. 둘째로는 원격 병리(telepathology) 분야[11]에서 활용될 수 있는데, 이 분야에서는 세계 어디서든지 간에 인터넷을 통해 의사가 환자와 연관된 조직 샘플을 보면서 상담할 수 있게 해준다. 셋째로는 조직 샘플이 분실될 염려가 없고 조직 샘플이 필요한 경우 이를 찾기 위해 시간과 노력을 절약할 수 있다. 넷째로는 영상 회의와 결부되어 사용되는 경우로, 가령 세계의 의학 전문가들이 시스템을 이용하여 특정 조직 샘플을 인터넷을 통해 같이 보면서 영상 회의를 이용하여 학술 토론을 할 수 있다. 이러한 응용은 연구자들이 물리적으로 모여 토론하는 대신에 인터넷 상에서 토론하는 것을 가능하게 하므로 이 시스템을 이용한 가상 학술 회의를 개최할 수도 있다.

일반적으로 원격 병리에서는 지리적으로 멀리 떨어진 곳에서 조직 샘플들을 검사한다. 여기에는 두 가지 방식이 사용되는데, 하나는 정적(static)인 방식이고, 다른 하나는 동적(dynamic)인 방식이다. 동적인 방식은 다른 곳에 위치한 병리 상담자가 서버로부터 원하는 이미지를 선택하고 배울 등을 임의로 조절하여 검색할 수 있는데 비해, 정적인 방식에서는 병리 상담자가 미리 정해진 이미지 집합만을 검색할 수 있기 때문에 다양한 환경의 변화에 대응하기가 어렵다. 이러한 관점에서 이 시스템은 디지털 동적 소프트웨어(dynamic software)라 볼 수 있다[2].

이 시스템은 원격 병리 분야에 한정되지 않고 대용량 다차원 데이터 집합을 저장, 검색 기능들을 제공하는 보다 일반적인 소프트웨어로 발전할 수 있다. 예를 들면, 과학 및 공학 시뮬레이션의 결과를 보여준다든지, 인공 위성 등에 의해 포착된 이미지들을 분석한다든지 할 경우에도 역시 활용 가능하다. 그런데, 이러한 응용 분야들은 관심 영역의 데이터에 대한 매우 간단한 질의에 의해서도 수 MB의 데이터를 질의 결과로 생성하는 특징이 있다.

2장에서는 시스템 구성에 대해 기술하고, 3장에서는 디스크 입출력의 대역폭을 높이기 위해 데이터 저장 및 질의 처리가 어떻게 이루어지는지에 대해 기술하고 4장에서는 고객에게 데이터 시각화가 어떻게 이루어지는지 설명하고, 5장에서는 관련 연구들과 비교를 하며, 6장에서 결론을 맺는다.

2. 시스템 구성

본 시스템은 크게 클라이언트(client)와 서버(server)부로 구성된다. 기본적으로 분산 시스템[8]으

로 구성되거나 병렬 시스템[1]으로 구성되므로 서버 부분은 다시 조정자 노드(coordinator node)와 서비스 노드(service node)로 구성된다. 조정자 노드는 고객과 상호 작용하면서 서비스 노드들에게 질의를 배분하고 서비스 노드들은 저장된 다차원 데이터들을 처리 생성하여 출력을 고객에게 전송한다.

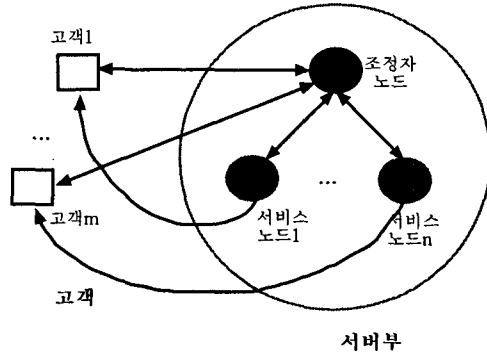


그림 1. 시스템 구성

1. 조정자 노드

이 노드는 정해진 통신 규약(communication protocol)에 따라 고객의 요청을 서비스 노드에게 배분한다. 고객은 고객 질의 CQ(Client Query) 메시지를 조정자 노드에게 보내면, 조정자 노드는 고객 질의에 대해 응답할 서비스 노드 수를 포함하는 고객 질의 응답 CQA(Client Query Acknowledgement) 메시지를 고객에게 돌려보낸다. 그리고, 서비스 노드에게는 전처리 질의 FQ(Frontend Query) 메시지를 보낸다. 서비스 노드들은 영역 질의로 표현된 FQ 메시지들을 수신한 후, 지역 디스크상에 분산 저장된 다차원 데이터를 검색하여 데이터를 처리한다. 처리후 생성된 출력 데이터로 구성되는 응답 R(Reply) 메시지를 직접 요청한 고객에게 전송한다. 고객은 서비스 노드들로부터 모든 이미지 조각들을 받은 후 이들을 조합하여 완전한 이미지를 보여준다.

2. 서비스 노드

일반적으로 분산 시스템이나 병렬 시스템으로 구성되는 서비스 노드들은 다해상도 이미지 데이터들을 서비스하기 위하여 현미경의 최고 배율로 캡처된 데이터들을 보관한다. 각 서비스 노드들은 고객에게 최종 이미지를 제공하기 전에 디스크로부터 읽혀진 데이터들을 고

객이 요청한 배율에 따라 데이터 변환을 수행한다. 이러한 서비스 노드의 작업들은 주로 디스크 입출력 연산이므로 병렬 디스크 접근(parallel disk access)을 수행하는 것이 성능에 중요하다.

3. 데이터 저장

서비스 노드에서 실행되는 서버 프로그램은 네트워크나 병렬 컴퓨터의 각 노드들에서 수행되는 분산 프로그램이며, 서비스 노드에는 입출력의 기본 단위인 블록(block)들로 데이터들이 분할되어 저장된다. 블록내 모든 데이터들을 포함하는 MBR(Minimum Bounding Box)에 의해 블록은 다차원 좌표계와 연관된다.

각 블록들은 디스크 대역폭을 높이기 위해 시스템내 디스크들에게 분산 배치되어야 한다. 영역 질의에 의한 데이터 접근 패턴을 이용하기 위해 서로 인접한 블록들은 서로 다른 디스크들에게 할당시키는데, 이 부분은 힐버트 곡선 기반의 분산 알고리즘(Hilbert curve-based declustering algorithm)[3]을 사용한다. 이 알고리즘은 빠르면서 균집/분산 특성을 잘 나타낸다. 그러나 그래프 분할(graph partitioning)[7]과 같은 다른 알고리즘도 채택하여 사용할 수 있다.

각 블록 단위로 개별 디스크에 할당되고, 지역 프로세서에 의해서만 읽기/쓰기가 행해진다. 이들 블록에 대한 다차원 인덱스(multidimensional index)는 각 블록의 MBR를 이용하여 구성되며, 이들 인덱스에 의해 요청된 영역 질의와 교차하는 MBR를 가진 블록들을 찾게 된다. 본 시스템의 기본 인덱싱 방법으로는 R-트리[4]를 사용한다.

4. 데이터 시각화

고객 프로그램(client program)은 사용자들이 선택한 슬라이드 중에서 관심있는 영역을 찾아 마우스로 이동하거나 버튼을 눌러 관찰할 수 있도록 그래픽 사용자 인터페이스(graphical user interface)를 제공한다[14]. 본 시스템 고객의 사용자 인터페이스가 제공하는 주요 기능들을 보면,

- ①관심 영역을 찾기 위한 슬라이드의 빠른 브라우징(fast browsing)
- ②현 화면 주변의 영역을 관찰하기 위한 지역 브라우징(local browsing)
- ③배율 변화(magnification change)

④초점 평면 변화(focal plane change)

고객 프로그램은 다중스레딩(multithreading) 기술을 적용하여 사용자 인터페이스 스레드(user interface thread)와 망 접속 스레드(network interface thread), 두 개의 스레드로 구성되며, 사용자 인터페이스 스레드가 망 접속 스레드보다 높은 우선 순위를 갖는다. 사용자 인터페이스는 그림 2와 같이 두 개의 창으로 이루어지는데, 그림에서 우측에 위치한 것이 제어 창(control window)이고, 좌측에 위치한 것이 관찰창(observation window)이다. 제어창은 다시 다음의 4개 요소들로 구성된다.

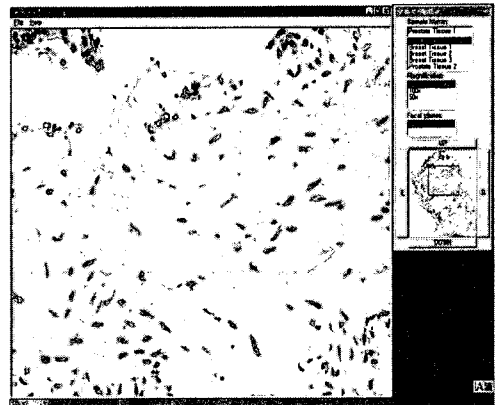


그림 2. 사용자 인터페이스

①샘플명(sample name)

여기에는 서비스 노드가 저장하고 있는 조직 샘플 슬라이드들의 목록을 보여준다. 이 샘플 목록은 사용자가 조정자 노드 접속시 노드로부터 전달받는다. 그림 2에서 현재 선택된 슬라이드가 전립선 조직 1(prostate tissue 1)임을 보여주고 있으며, 그 외에도 다른 4종의 조직 샘플들이 검색 가능함을 알려 준다.

②배율(magnification)

이것은 사용자가 관찰하고자 하는 현미경의 배율을 나타내며, 사용자가 배율을 선택하면 해당 배율에 맞춰 샘플의 이미지를 관찰할 수 있다. 위 그림은 전립선 조직1에 대한 50배,100배,200배 확대하여 관찰할 수 있음을 보여준다.

③초점 평면(focal planes)

슬라이드를 구성하는 여러 초점 평면들의 목록을

보여 준다. 위 그림에서 전립선 조직¹은 한 개의 초점 평면만을 포함하고 있음을 알려준다.

④전체 축소 이미지(thumbnail image)와 버튼들

이 이미지는 전체 슬라이드를 축소된 것으로 슬라이드에 대한 사용자의 브라우징을 도와 준다. 브라우징은 이 이미지 안에 위치한 뷰 박스(view box)를 마우스로 이동시켜 가면서 관심 영역의 이미지를 찾아 갈 수도 있고, 다른 방법으로는 이미지 옆에 붙어있는 4 방향의 버튼(Up,Down,L,R)들을 누름으로써 현재 보여지는 영역과 매우 인접한 부분으로 세밀한 이동을 할 수 있다.

다른 하나의 창은 서비스 노드로부터 도착한 이미지 조각(image chunk)들을 조합해서 사용자가 요구한 슬라이드상의 관심 영역에 해당하는 이미지를 보여 주는 관찰 창이다. 물론 이 창의 크기는 사용자가 원하는 대로 크기의 조절이 가능하다. 또한 축소 이미지 상에서 다른 부분으로 빠르게 뷰 박스를 이동시킬 때에는 이 전체 축소 이미지를 사용하며 원하는 이미지를 개략적으로 보여주기 위해 축소 이미지를 확대하여 보여 준다. 사용자가 마우스 이동을 마치고 마우스 버튼을 놓게 되면 고객 프로그램은 해당 부분의 검색을 위한 질의를 생성하여 조정자 노드에게 전송한다.

5. 관련 연구

현재 영상 분석, 저장, 검색 등의 기능을 제공하는 많은 종류의 소프트웨어 제품들이 유통되고 있다. 그 중에서 의학 분야에서 활용되고 있는 PACS(Picture Archiving and Communication System)들과 영상 분석을 위주로 하는 소프트웨어들이 본 시스템과 많은 기능적 유사점들을 가지고 있다[13].

본 시스템이 이들간의 유사점은 어떠한 영상 입력 장치로부터든지 영상들을 획득한 후, 이들을 저장하여 검색 기능을 제공한다는 점이고, 차이점으로는 본 시스템은 조직 샘플들로부터 획득된 고해상도 이미지들에 대해 현미경적 관찰을 제공하는 소프트웨어인 반면에, PACS용 소프트웨어들은 의학 영상들을 획득하여 이를 전송 및 저장하여 의사나 연구원들이 진료시 또는 학술 연구용으로 이들을 필름대신에 사용할 수 있도록 제공한다. 이러한 차이점에도 불구하고, 외형상 본 시스템과 비교될 수 있는 제품들로는 우리나라의 메디페이스에서

개발한 PACS를 구성하는 PiView, Interview, Spectra[12]와 미국 Media Cybernetics사의 영상 분석 소프트웨어인 Optimas[6] 등을 들 수 있다.

영상을 표현해 주는 메디페이스사의 PiView는 윈도우즈 환경에서 실행되며 초음파(Ultrasound), 컴퓨터 단층 촬영, 자기공명 촬영 등 여러 영상 획득 장치로부터 입력된 의학 영상들뿐만 아니라 16비트 회색 범위(gray scale) 또는 24비트 컬러 영상 등도 모니터에 표현할 수 있고, 화면은 1024x768 화소부터 지원한다. PiView는 DICOM(Digital Imaging and Communications in Medicine) 3.0과 호환되는 영상은 물론 호환되지 않는 영상들에 대해서도 처리가 가능하며 DICOM 3.0 질의/검색과 저장 기능을 제공한다. PiView는 DB 관리자, NetGate, 영상 표시기(image viewer), TeleGate, VisualGate로 구성되고, 단순히 영상 표시 기능만이 아닌 영상 분석 및 DB 관리 기능까지 통합 제공하고 있다. 그 중에서도 PiView의 영상 표시기는 영상의 축소/확대, 영역 표현, 칼라 변환, 반전 등의 영상 처리 기능들을 제공한다. 또한 인터넷에서 메디페이스의 PACS에 접속하기 위한 수단을 제공하는 InterView는 비교적 단순한 기능을 제공하는 웹 기반의 소프트웨어이고, 이미지 저장을 위한 서버인 Spectra는 단기 저장 장치와 장기 저장 장치로 구성되고 이들은 각각 50MB의 RAID(Redundant Arrays of Inexpensive Disks)[10]와 512GB 저장 용량을 지원한다. 단기 저장 장치의 영상은 일정 기간동안 원래대로 저장하다가 이 기간이 경과하면 약 2:1의 평균 압축 비율로 압축한 후 장기 저장 장치인 테이프 장치에 보관한다. Spectra는 Oracle 데이터베이스로 구축되었다.

Media Cybernetics사의 Optimas는 기본적으로 영상 분석을 위한 목적으로 만들어진 소프트웨어로서 임의의 영상에 대해 다양한 분석 기능들을 제공한다. 분석 대상 영상에 대해 영역 분할, 특징 인식, 입자 인식, 점 분석, 거리 측정 등의 영상 분석 기능들을 제공하고 자동적으로 이들을 처리해 주며, 이들에 대한 분석 결과를 통계 처리한 후 그래프로 보여 주는 기능 등을 또한 제공한다. 일반적으로 영상 관련 소프트웨어들이 제공하는 영상 색깔 처리, 음영 조절, 노이즈 제거, 영상 축소/확대, 회전, 스크롤링, 주석 달기 및 출력 등의 기능들이 함께 제공된다. 이 소프트웨어는 카메라, 스캐너 등의 영상 입력 장치와 접속하여 영상을 입력받아 이들을 파일로 저장한 후 처리하고 있다.

6. 결론

본 논문은 광학 현미경을 컴퓨터 상에서 구현한 소프트웨어 시스템의 설계에 관한 것으로서, 실제 현미경의 기능을 그대로 재현하는 것은 물론 더 나아가 실제 현미경을 사용하는 것보다도 더 유용한 기능들을 제공한다. 궁극적으로 이 시스템은 의과 대학, 의료 기관 및 의학 연구 기관의 슬라이드로 채워진 캐비닛을 디지털 저장 시스템으로 대체하기 위한 고객 서버 소프트웨어이다. 샘플 목록으로부터 사용자가 관심있는 슬라이드를 선택한 후 현미경을 통해 관찰하듯이 컴퓨터 상에서 조직 샘플들을 관찰할 수 있게 하는 이 시스템은 의학 교육, 임상/원격 병리, 원격 교육, 가상 학술 회의 등에 응용될 수 있다.

고객 서버 모델에 따라 설계된 이 시스템은 고객, 조정자 노드, 서비스 노드 3 부분으로 구성되었고 각 부분이 상호 독립적인 이들은 정해진 통신 규약에 따라 메시지를 주고 받는다. 이러한 모듈화에 의해 각 부분은 별도로 개발될 수 있었으며 변경이 필요한 경우에도 다른 모듈에 영향을 끼치지 않고 성능 향상과 기능 추가가 가능하다.

향후에는 인터넷이나 근거리 통신망 또는 내부 스위치에서 소요되는 전송 시간으로 인한 이미지 검색 시간을 줄일 수 있는 사전 인출(prefetching)과 캐싱(cacheing) 기술에 대한 연구가 요구되며, 대용량의 데이터들을 대상으로 저장 및 관리하게 되므로 전송시어의 크기를 줄이기 위한 압축 기술의 적용에 대한 연구가 수행될 필요가 있다. 또한 사용자 관점에서는 이미지 분석 기능, 통계 처리 기능 등이 요구되며, 성능에 대한 연구가 향후에 진행되어야 할 것이다.

[참고문헌]

- [1]T. Agerwala, J. Martin, J. Mirza, D. Sadler, D. Dias, and M. Snir, "SP2 System Architecture," IBM System Journal, Vol.34, No.2, 1995.
- [2]S. Cho et al., "Digital Dynamic Telepathology - the Virtual Microscope," In *Proceedings of the 1998 AMIA Annual Symposium*, pp. 449-453, Oct. 1998.
- [3]C. Faloutsos and P. Bhagwat, Declustering using fractals, In *Proceedings of the 2nd International Conference on Parallel and Distributed Information System*, pp.18-25, Jan. 1993.
- [4]A. Guttman, R-Trees: A dynamic index structure for spacial searching, In *Proceedings of the 1984 ACM SIGMOD International Conference on Management of Data(SIGMOD84)*, pp.47-57, Boston, MA, Jun. 1984.
- [5]T. Kurc, U. Catalyurek, C. Chang, A. Sussman and J. Saltz, "Exploration and Visualization of Very Large Datasets with the Active Data Repository," *IEEE Computer Graphics and Applications*, 21(4):24-33, Jul/Aug. 2001.
- [6]Media Cybernetics, <http://www.optimas.com/>, 1998.
- [7]B. Moon, and J. Saltz, "Scalability analysis of declustering methods for multidimensional range queries," *IEEE Trans. on Knowledge and Data Engineering*, 10(2):310-327, Mar/Apr. 1998.
- [8]R. Orfali, D. Harkey and J. Edwards, *Client/Server Survival Guide*, 3rd edition, John Wiley and Sons, 1999.
- [9]H. Samet, *The Design and Analysis of Spatial Data Structure*, Addison-Wesley Publishing Co. 1994.
- [10]A. Silberschatz, H. Korth, and S. Sudarshan, *Database System Concepts*, 3rd ed. McGraw-Hill, 1999.
- [11]R. Weinstein, A. Bhattacharyya, A. Graham, and J. Davis, "Telepathology: A ten-year progress report," *Human Pathology*, 28(1):1-7. Jan. 1997.
- [12]김 영훈, "Mediface PACS," 대한 PACS 학회지, 제4권 제1호, pp.29~32, 1998.6.
- [13] 김 영훈, 최 승욱, 최 형식, "Archiving PACS for CT, MR, DSA," 대한 PACS 학회지, 제4권 제1호, pp.45~50, 1998.6.
- [14]조 승호, "현미경적 분석을 지원하는 이미지 저장/검색 시스템," 한국정보처리학회, 제6권 제2호, pp.DB-38~45, 1998.10.