

# 기하 인스턴싱 기법을 이용한 단백질 구조 가시화 및 속도 향상에 관한 연구

박 찬 용<sup>\*</sup> · 황 치 정<sup>\*\*</sup>

## 요 약

구조적 생물 정보학 분야는 단백질의 3차원 구조를 대상으로 단백질을 연구하는 분야이며, 구조적생물 정보학의 중요한 분야 중의 하나는 단백질 3차원 구조 가시화이다. 단백질의 3차원 구조를 규명하는 장비의 발달로, 규명되는 단백질의 크기와 개수가 증가함에 따라, 고성능의 단백질 가시화 시스템의 필요성도 크게 증가하였으나, 기존의 단백질 구조 가시화 시스템은 3차원 그래픽 하드웨어에 최적화 되지 못하여, 거대 단백질의 가시화에 충분한 성능을 가지지 못하였다. 본 논문에서 제안하는 단백질 3차원 구조 가시화 시스템은 거대 단백질의 가시화 하기 위하여, 3차원 그래픽 하드웨어의 최적화 기법중의 하나인 기하 인스턴싱 기법을 사용하여 빠르게 거대 단백질을 렌더링 한다. 성능 실험에서 7종의 다른 크기의 단백질을 대상으로, 4가지 가시화 방법에 대하여, 제안하는 시스템과 기존의 시스템과의 단백질 렌더링 성능 비교 실험을 하여, 대부분의 경우 우수한 성능을 보였다.

키워드 : 컴퓨터 그래픽, 바이오인포매틱스, 단백질 구조 가시화

## The Study of Protein Structure Visualization and Rendering Speed Using the Geometry Instancing

Chan-Yong Park<sup>\*</sup> · Chi-Jung Hwang<sup>\*\*</sup>

## ABSTRACT

Analysis of 3-dimensional (3D) protein structure plays an important role of structural bioinformatics. The protein structure visualization is the one of the structural bioinformatics and the most fundamental problem. As the number of known protein structure increases rapidly and the study of protein-protein interaction is prevalent, the fast visualization of large scale protein structure becomes essential. The fast protein structure visualization system we proposed is sophisticated and well designed visualization system using geometry instancing technique. Because this system is optimized for recent 3D graphics hardware using geometry instancing technique, its rendering speed is faster than other visualization tools.

Keywords : Computer Graphics, Bioinformatics, Protein Structure Visualization

### 1. 서 론

단백질은 생체 내에서 생명현상 유지를 위한 다양한 기능과 역할을 한다[1] 대표적인 기능으로 생체내의 화학반응을 촉매하는 여러 가지 효소, 정보전달에 관여하는 세포막 수용체, 생체를 방어하는 항체, 근육 수축, 이완 단백질, 혈액 응고인자, 산소, 탄소가스의 운반체, 영양소의 운반체, 뼈와 연골 등을 구성하는 콜라겐 등이다. 단백질들은 생체 내에

서 공유 결합이나, 수소 결합 등 자연계에 존재하는 힘에 의해 특정한 형태의 3차원 구조를 형성하고 있다. 단백질의 3차원 구조는 단백질의 기능과 밀접한 관련이 있기 때문에, 단백질의 기능을 규명하기 위하여 단백질의 3차원 구조 연구는 필수적이다.

단백질 3차원 구조 가시화는 NMR 이나 X-Ray crystallography 등을 이용하여 규명된 단백질을 3차원 컴퓨터 그래픽을 이용하여, 단백질의 구성요소를 사용자가 원하는 다양한 형태로 표현하는 것이다. 최근 규명된 단백질의 개수와 단백질 구조의 크기가 급속하게 증가하여 단백질 구조 가시화 시스템은 단백질을 구성하는 원자의 수가 매우 많은 (> 50K 원자), 거대한 크기의 단백질을 렌더링 할 필요성이 더욱 증대되었다. 그리고, 여러 종류의 단백질을 동시에 렌

\* 이 논문 또는 저서는 2007년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임(KRF-2007-611-D00029).

† 정 회 원 : 한국전자통신연구원 선임연구원

\*\* 종 신 회 원 : 충남대학교 전기정보통신공학부 교수

논문접수: 2009년 1월 22일

수정일: 1차 2009년 3월 13일

심사완료: 2009년 4월 3일

더링하여 단백질간의 상호작용을 가시화하기 위한 요구도 증대되었다.

단백질 구조 가시화를 위하여 많은 시스템이 제안되어 왔다. 대표적인 가시화 시스템은 Chimera[2], VMD[3], DS Visualizer[4], PyMol[5], Jmol[6]등이다. 단백질 3차원 구조 가시화 시스템들은 고품질, 고성능의 단백질 가시화를 위하여 3차원 그래픽 하드웨어에 크게 의존하고 있다. 최근 개인용 PC 시스템에서 3차원 그래픽 하드웨어의 성능이 매우 급속하게 발전하였지만, 기존의 가시화 시스템들은 최근에 개발된 고성능의 3차원 그래픽 하드웨어에 최적화 되어 있지 않아, 거대한 크기의 단백질을 렌더링 하거나, 여러 단백질을 동시에 화면에 렌더링 할 경우, 빠른 가시화가 이루어지고 있지 않다. 제안하는 단백질 3차원 구조 가시화 시스템은 거대 단백질을 가시화하고 여러 단백질을 동시에 가시화 하기 위하여, 3차원 그래픽 하드웨어에 최적화를 통하여, 3차원 구조를 빠르게 렌더링 한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련연구를 기술하고, 3장에서는 새로운 단백질 구조 가시화 시스템을 제안한다. 4장 실험 및 성능 평가에서는 단백질 3차원 구조 가시화 시스템의 성능 평가를 수행하고, 마지막으로 5장에서는 본 논문의 결론과 향후 연구를 기술한다.

## 2. 관련 연구

단백질 3차원 가시화 시스템들은 고성능의 3차원 그래픽을 위하여 3차원 그래픽 하드웨어에 크게 의존하고 있다. 최근 개인용 PC 시스템에서의 3차원 그래픽 하드웨어의 성능이 매우 급속하게 발전하여, 빠른 3차원 그래픽을 화면에 디스플레이 하기 위해서 3차원 그래픽 하드웨어를 직접적으로 이용하는 것은 필수적이다. 이들 하드웨어의 성능을 직접적으로 이용하기 위하여 그래픽 라이브러리는 다양한 형태의 3차원 렌더링의 최적화 방법을 사용자에게 제공하고 있다. 기존의 단백질 구조 가시화 시스템들은 최근의 고성능 그래픽 하드웨어에 최적화가 되어 있지 않아, 원자의 수가 큰 거대 단백질에서는 가시화의 속도가 급격히 떨어지고 있는 실정이다.

또한 기존의 시스템들은 거대 단백질이나, 여러 단백질을 동시에 가시화 하는 것을 고려하지 않고 개발되어, 화면에 가시화된 복잡한 단백질 구성 요소 중에서 사용자가 필요로 하는 원자나 잔기(Residue)를 정확하게 선택하는 유저 인터페이스에 관련된 기능이 많이 부족하다.

이러한 문제점을 해결하기 위하여, 제안하는 단백질 3차원 구조 가시화 시스템은 단백질 구조 가시화를 위하여 3차원 그래픽 하드웨어에 최적화된 렌더링을 수행한다. 그리고, 화면에 가시화된 복잡한 단백질 구성요소를 사용자가 정확하고 빠르게 선택하기 위하여 단백질의 계층구조를 트리 구조로 보여주며, 선택된 영역의 가시화를 하기 위하여 선택 영역 리스트를 사용하여 단백질의 특정 영역만을 가시화 한다.

## 3. 단백질 구조 가시화 시스템

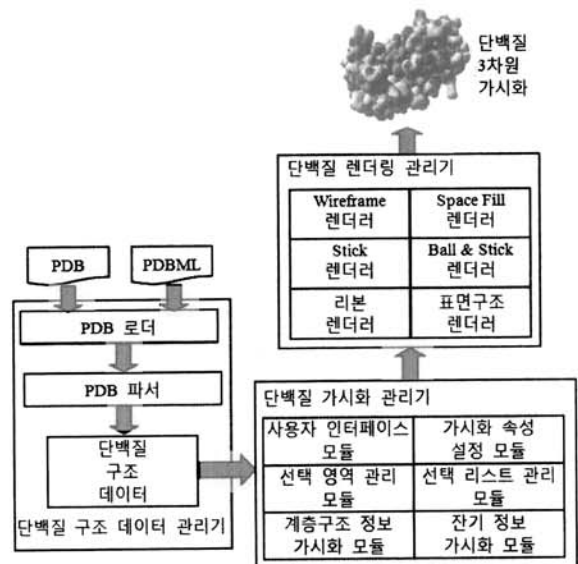
이 장에서는 제안하는 단백질 구조 가시화 시스템의 개요 및 구성을 기술하고, 단백질 구조의 3차원 가시화 방법 및 논리적 정보 가시화 방법을 설명한다.

### 3.1 시스템의 개요 및 구성

제안하고 있는 단백질 3차원 구조 가시화 시스템의 특징은 다음과 같다. (1) 단백질을 3차원으로 가시화 하기 위한 렌더링 모듈이 3차원 그래픽 하드웨어에 최적화되어, 거대 단백질 가시화나 여러 단백질을 동시에 가시화 할 경우, 가시화를 매우 빠르게 수행 할 수 있다. (2) 단백질의 특정한 선택 영역에 대해 사용자의 요구에 맞는 다양한 단백질 3차원 가시화가 가능하다. (3) 단백질 구조의 계층적 단위인 모델-체인-잔기-원자 단위로 단백질의 논리적 정보를 트리 형태로 가시화 하고, 단백질의 서열 정보를 가시화 한다.

시스템의 전체적인 구성은 3부분으로 구성되어 있다. 첫 번째로, PDB 단백질 구조 파일을 로드하고 파싱하는 단백질 구조 데이터 관리기, 두 번째로, 단백질 구조의 특정 부분을 선택하고 가시화하기 위한 단백질 가시화 관리기, 세 번째로, 단백질을 다양한 형태로 가시화 하기 위한 단백질 렌더링 관리기로 구성된다.

(그림 1)에 시스템의 주요한 구성 요소를 보여준다. 단백질 구조 데이터 관리기는 PDB 로더, PDB 파서로 구성되어 있다. PDB 로더는 PDB 파일을 로드하는 모듈로, 파일 시스템이나, PDB 데이터베이스에 저장된 PDB 파일을 로드 한다. PDB 파서는 로드된 PDB 파일을 파싱하는 모듈이다. PDB 파일 구조가 계층적인 구조로 되어 있으므로, PDB 파일에 정의된 모델, 체인, 잔기, 원자의 구조를 모두 파싱하여 파싱된 결과를 단백질 구조 데이터(Protein Structure Data)의 형태로 저장한다. 단백질 구조 데이터에는 PDB에 포함



(그림 1) 단백질 구조 가시화 시스템 구성도

된 모델, 체인, 잔기, 원자 데이터 뿐만 아니라, PDB의 헤더에 포함되는 단백질의 일반적인 정보, 단백질의 2차 구조 정보와, 원자간의 결합관계, 원자의 반데르 발스 반지름, 모델과 체인의 3차원 중심등의 정보가 분석되어 추가적으로 저장된다. 2차 구조 정보는 대부분 PDB 안에 포함되어 있지만, 만약 2차 구조 정보가 PDB 안에 존재하지 않을 경우에는 DSSP[7]를 이용하여 2차 구조 정보를 생성한다.

단백질 가시화 관리기는 단백질 구조 데이터를 다양한 형태로 화면에 가시화 하기 위해서 사용자가 단백질의 특정한 구조를 선택하고, 선택된구조에 대해서 다양한 가시화 속성을 설정하는 모듈로써, 사용자 인터페이스 모듈, 가시화 속성 설정 모듈, 선택영역 관리 모듈, 선택 리스트 관리 모듈, 계층구조 정보 가시화 모듈, 잔기 정보 가시화 모듈로 구성되어 있다.

단백질 계층 구조 정보 가시화 모듈은 단백질의 구조적 계층 구조를 트리 구조로 화면에 출력하게 하는 모듈로써, 사용자가 특정한 모델, 체인, 잔기, 원자를 트리 형태의 유저 인터페이스에서 선택할 수 있다. 잔기 정보 가시화 모듈은 잔기를 리스트 형태로 잔기의 구조를 화면에 출력하도록 한다. 선택 영역 관리 모듈은 단백질을 트리 구조, 잔기 리스트 구조, 3차원 구조상에서 단백질을 선택하여, 선택된 부분을 관리하는 모듈이다. 선택된 단백질의 일부분은 선택 리스트 관리 모듈의 선택 영역 리스트에 들어간다. 선택 영역 리스트의 각각의 항목인 선택 영역은 가시화 속성 설정 모듈에서 다양한 3차원 가시화 속성을 설정한다. 사용자 인터페이스 모듈은 사용자 인터페이스에 관련된 작업을 수행하는 모듈로 계층 구조 정보 가시화 모듈, 잔기 정보 가시화 모듈, 선택 영역 관리 모듈의 사용자 인터페이스를 관리한다.

단백질 렌더링 관리기는 wireframe 렌더러, Spacefill 렌

더러, Stick 렌더러, Ball&Stick 렌더러, 리본구조 렌더러, 표면 구조 렌더러로 구성되어, 단백질을 Wireframe, Spacefill, Stick, Ball&Stick, 리본구조, 표면 구조로 렌더링 할 수 있다. 6개의 단백질 렌더러는 기하 인스턴싱 방법을 이용한 최적화를 통하여 화면에 디스플레이를 한다.

### 3.2 단백질 구조의 3차원 가시화

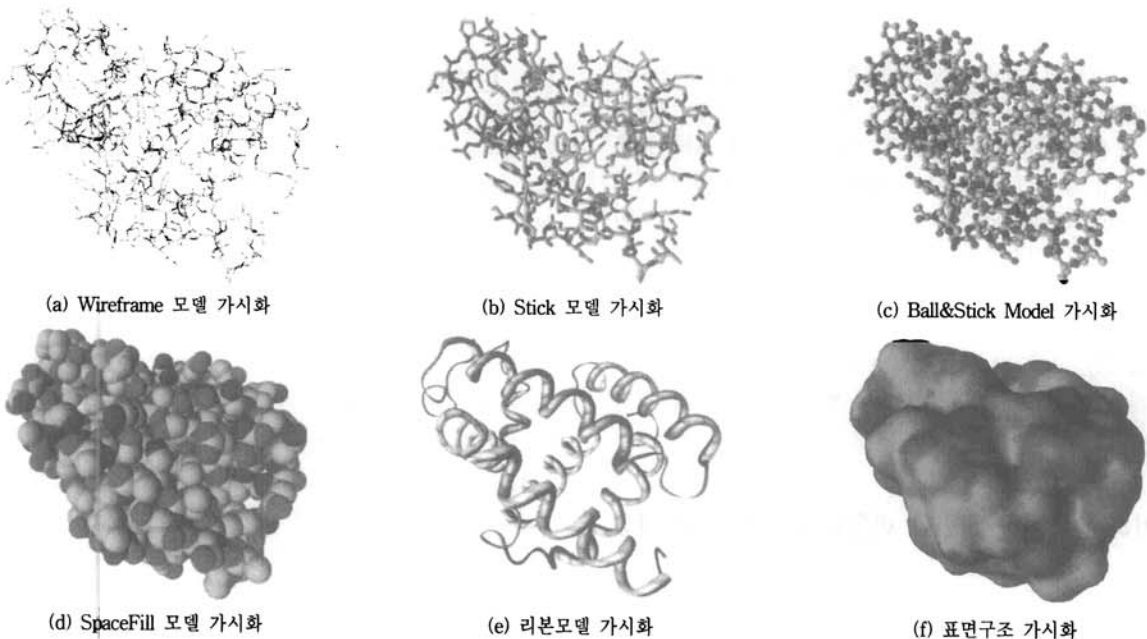
단백질 구조 가시화 시스템의 단백질의 3차원 가시화 방법은 6가지이다. 각각의 가시화 방법은 다른 가시화 방법이 보여주지 못하는 다양한 정보를 디스플레이 하기 때문에 제안하는 단백질 구조 가시화 시스템은 다양한 형태로 단백질 구조를 가시화 한다.

#### 3.2.1 Wireframe 모델 가시화

Wireframe 가시화 방법은 결합관계(bonds) 가시화 방법으로 불리는 것으로, 단백질을 구성하는 원자간의 결합관계를 선으로 표시한 방법이다(그림 2-(a)). 단백질 결합관계만 표시되어 있으므로, 단백질을 구성하는 원자의 이름이나, 크기가 가시화 되지 않는다.

#### 3.2.2 Stick 모델 가시화

Stick 모델 가시화 방법(그림 2-(b))은 Wireframe 모델 가시화 방법과 유사하지만, Wireframe 모델 가시화 방법과는 달리 원자간 결합관계가 일정한 두께가 있는 원통형으로 가시화를 한다. 원통형 구조의 끝에는 원자가 위치하고, 원자의 지름은 원통의 지름과 동일하다. 이 방법은 Wireframe 방법과 같이 원자간 결합(bond)를 중점적으로 가시화를 하지만, 원통의 색을 이용하여 원자의 종류도 가시화를 한다.



(그림 2) PDBID 1FLP의 다양한 가시화 방법

### 3.2.3 Ball&Stick 모델 가시화

Ball&Stick 모델 가시화 방법은 (그림 2-(c))와 같다. Ball & Stick 모델 가시화는 Stick 모델 가시화 방법을 기반으로 원자가 위치하는 곳에 일정크기의 구를 위치시키는 것이다. 이 방법은 원자간 결합뿐만 아니라, 원자의 위치도 가시화를 하지만, 원자의 크기는 나타내지는 못한다.

### 3.2.4 Space-Fill 모델 가시화

(그림 2-(d))은 Space-Fill 모델 가시화 방법으로 반데르 발스(van der Waals) 모델 가시화 방법이라고도 한다. 이 방법은 단백질을 구성하는 원자가 차지하는 체적(반데르 발스 크기)를 가시화 한다. 그러므로, 단백질 전체가 차지하는 부피나, 단백질의 전체구조를 가시화할 수 있다. 이 모델에서 각각의 원자는 반데르 발스 반지름의 크기를 가진 구 형태로 가시화 된다.

### 3.2.5 리본 모델 가시화

리본모델은 단백질의 2차 구조 정보를 가시화하는데 주로 사용된다(그림 2-(e)). 단백질의 2차 구조 정보는 단백질 원자의 3차원 좌표로 계산되어 PDB 파일에 정보가 포함되어 있다. 만약 2차 구조 정보가 PDB 안에 존재하지 않을 경우 DSSP[7]를 이용하여 2차 구조 정보를 계산한다.

2차 구조를 3차원으로 표시하기 위한 방법은 다양한 방법이 제안되었다[8, 9]. 본 논문에서는 기존의 [9]의 방법을 개선하여, 원자간 보간에 허미트 스플라인 커브(Hermite Spline curve)를 이용한다. 리본모델을 3차원으로 가시화 하기 위한 메쉬를 생성하기 위하여 Ca 와 Ca+1 간의 벡터를 식 1과 같이 계산한다.

$$\vec{A} = C\alpha_{i+1} - C\alpha_i \vec{B} = O_i - C\alpha_i \vec{C} = \vec{A} \times \vec{B} \quad (1)$$

### 3.2.6 표면 모델 가시화

단백질 표면의 개념은 단백질 폴딩과 단백질의 소수성(hydrophobicity)에 관련된 실험 중 단백질의 소수성 원자의 체적을 계산하기 위하여 처음 제안되었다. 여러 표면 모델 가시화 방법 중 본 논문에서는 solvent-excluded 단백질 표면 구조(SES, molecular surface)를 가시화 한다(그림 2-(f)) [10, 11]. SES의 기하학적 개념은 단백질의 반 데르 발스 표면에서 특정한 크기(probe radius)의 구를 움직였을 때, 구의 중심점을 모두 합하여 생성되는 또 다른 표면을 가리킨다. 대표적인 구의 반지름은 1.4Å 으로, 물 분자의 크기를 사용한다. 구현 알고리즘으로 [12]의 방법을 표면 모델 가시화로 사용한다.

## 4. 기하 인스턴싱 기법을 이용한 렌더링 성능 향상

단백질 3차원 구조 가시화 시스템은 고품질, 고성능의 단백질 가시화를 위하여 3차원 그래픽 하드웨어와 그래픽 라이브러리에 최적화 되어야 한다. 최근 개인용 PC 시스템에

서 3차원 그래픽 하드웨어의 성능이 매우 급속하게 발전하였지만, 많은 가시화 시스템들은 최근에 개발된 고성능의 3차원 그래픽 하드웨어에 최적화 되어 있지 않아, 거대한 크기의 단백질을 렌더링 하거나, 여러 단백질을 동시에 화면에 렌더링 할 경우, 빠른 가시화가 이루어지고 있지 않다.

3차원 데이터의 빠른 렌더링을 위하여, 그래픽 하드웨어에 최적화 되어 GPU(Graphic Processing Unit)로 3차원 데이터를 렌더링 하는 것은 GPU 기반 렌더링(GPU based Rendering)이라고 한다. 3차원 데이터를 GPU 기반 렌더링을 이용하여 렌더링 할 경우, CPU와 병렬처리가 가능하여 성능을 향상시킬 수 있다.

단백질 구조의 3차원 렌더링 데이터의 특성은 Ball 모델, Ball&Stick 모델, SpaceFill 모델, 리본 모델에서 구(Sphere)나 원통(Cylinder)의 기본 3차원 도형을 반복적으로 렌더링 하는 특성을 가지고 있다. 이러한 특성을 이용하여 GPU 기반 렌더링의 한가지 방법인 기하 인스턴싱(Geometry Instancing) 기법을 이용하여 단백질 구조를 빠르게 렌더링 한다.

실시간 3차원 그래픽 분야에서 지오메트리 인스턴싱 기법은 같은 오브젝트를 동시에 여러 번 렌더링 할 경우 빠른 렌더링 성능을 나타낸다. 이 기법의 특징은 공용 데이터(일반적으로 정점, 법선 데이터)와 인스턴스 데이터(일반적으로 변환행렬)가 분리되어, GPU가 공용 데이터를 인스턴스 데이터의 개수로 여러 번 렌더링을 수행한다. 또한 이 방법은 공용 데이터가 각 인스턴스에 중복되지 않으므로, 렌더링 성능 향상뿐만 아니라 그래픽 메모리의 최적화도 가능하다. 공용 데이터와 인스턴스 데이터의 분리를 위하여 2개의 정점 버퍼(Vertex Buffer)를 사용하여 렌더링 한다. 이와 같이 기하 인스턴싱 기법은 모든 3차원 데이터에 대하여 적용 가능한 것은 아니고, 반복적인 데이터 특성에 사용가능한 것이므로, 본 논문에서는 stick, ball & stick, spacefill, 리본 모델 가시화에 대해서만 기하 인스턴싱 기법을 사용하였다.

## 5. 단백질 구조 가시화 시스템의 실험 및 성능평가

제안된 가시화 시스템의 성능 평가를 위하여 단백질 원자의 개수에 따른 여러 시스템의 성능을 렌더링 속도로 비교한다. 렌더링 속도는FPS(Frame per second) 단위로 측정되며, 화면에 단백질 3차원 가시화를 얼마나 빨리 수행하는가를 측정하는 것이다. 렌더링 품질이 유사하다면 렌더링 속도가 시스템을 평가하는 중요한 요소이다.

비교를 위하여 제안하는 시스템은 두가지 시스템으로 첫 번째 시스템(Proposed)은 기하 인스턴싱 기법을 사용하여 렌더링을 수행한 것이고, 두 번째 시스템(Proposed-2)는 기하 인스턴싱 기법을 사용하지 않고 렌더링한 것이다. 비교 대상 가시화 툴은Chimera, VMD, DS Visualizer, PyMol, Chime, JMol 이다. 실험대상으로 사용한 PDB는 PDB의 원자의 개수가 1000(1K), 5000(5K), 10000(10K), 50000(50K), 100K, 500K and 900K 이다. 900K 모델은 PDB 데이터베이스에서 가장 원자수가 많은 단백질이다. 실험 시스템은 Pentium4

3.6GHz, 램4GByte, GeForce 8800 GTX, WindowsXP 이다. 실험결과는 (그림 3)에 보여진다. (a)는 wireframe 모델 가시화, (b)는 spacefill 모델 가시화, (c)는 리본모델 가시화, (d)는 표면 모델 가시화 방법의 성능이다. 세로축은 성능 측정치로 FPS(Frame Per Second) 이고, 높을수록 좋은 결과를 나타낸다. 일반적으로 가시화 시스템에서 30 fps 이상은 가시화의 성능에 영향을 미치지 않으므로, 30 fps 보다 큰 경우는 모두 30 fps로 조정하여 결과를 나타내었다.

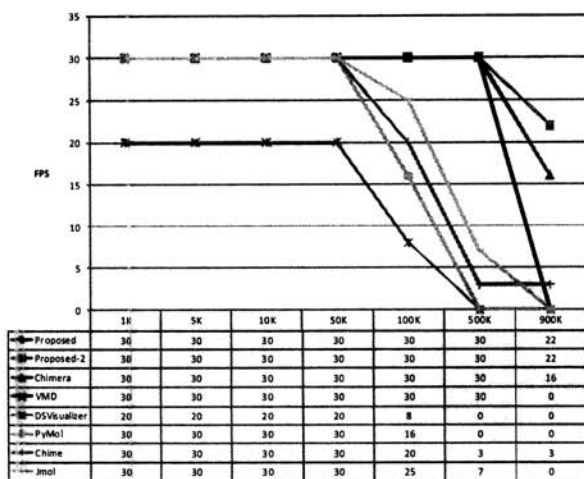
Wireframe 가시화의 경우는 대부분의 시스템에서 좋은 결과를 나타낸다. 제안하는 시스템은 900K개의 원자를 가진 단백질에 대해서도 약 20 FPS를 넘는 매우 우수한 성능을 보여준다. 또한 wireframe 모델에서는 기하 인스턴싱을 사용하지 않아, 실험시스템Proposed 와 Proposed-2는 동일한 성능을 나타내었다.

Spacefill 가시화의 경우는 렌더링을 해야 할 정점의 수가 매우 많아 모든 시스템에서 원자의 개수가 증가할수록 크게 성능이 떨어지는 결과를 보인다. 제안된 시스템은 900K 개의 원자에 대해 약 2 FPS 성능을 보인 반면, 기존의 시스템은 거의 0 FPS 에 가까운 성능을 나타낸다. 기하 인스턴싱을 사용하

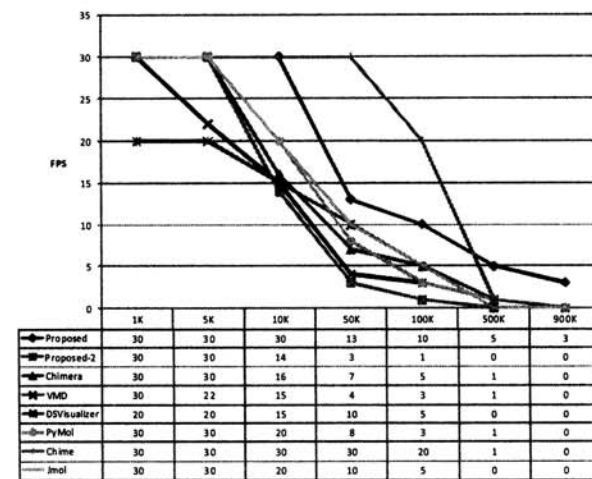
지 않는 Proposed-2 방법은 기존의 시스템과 거의 비슷한 결과를 나타내었다. Spacefill 모델은 기하 인스턴싱 기법이 많은 속도향상을 기대할 수 있는 렌더링 방법임을 알 수 있다. SpaceFill 모델 가시화에서 Chime 은 50K, 100K 원자에 대하여 매우 우수한 성능을 나타낸다. Chime은 OpenGL이나, Microsoft Direct3D 같은 기존의 3D 렌더러를 사용하지 않고, 자체적인 렌더러를 가지고 있어, 구형대만 가진 SpaceFill 렌더링시에 자체 렌더러에 최적화된 결과라고 생각된다.

리본 모델 가시화의 경우도 기존의 방법보다 우수한 결과를 보여주고 있고, Proposed-2 방법도 기존의 결과보다 약간 우수한 결과를 보여주고 있다. 리본모델의 경우에는 전체 모델에 대한 기하 인스턴싱 방법이 적용되지 않고, 2차구조의 특정 모델 반복에 대해서만 기하 인스턴싱이 적용되어 Proposed 방법과 Proposed-2 방법과의 성능향상이 크지 않다. Proposed-2 방법은 본 시스템의 정점 구조가 잘 정렬되어 다른 시스템과 비교하여 빠른 렌더링을 수행한다고 생각된다.

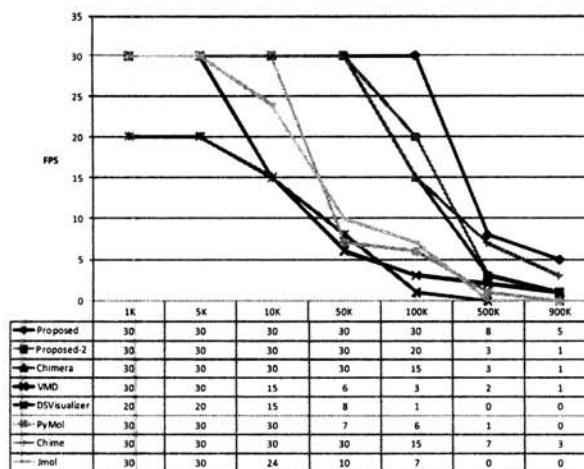
표면모델 가시화의 경우 기존의 시스템은 원자의 수가 많을 경우 단백질 표면 모델 생성에도 오류가 있어 표면 모델이 만들어지지 않는 경우도 발생하였다. 제안하는 시스템



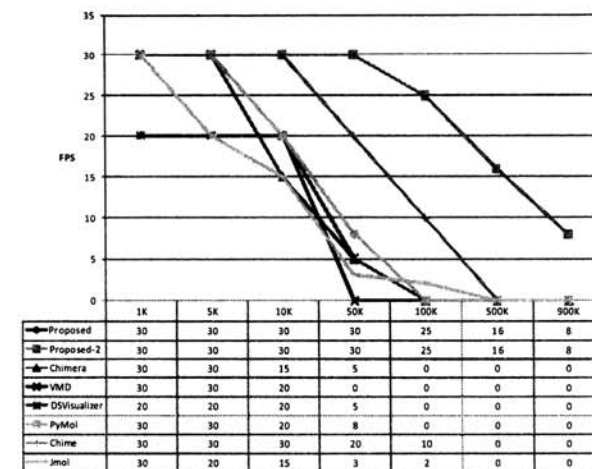
(a) Wireframe Display Model



(b) SpaceFill Display Model



(c) Ribbon Display Model



(d) Surface Display Model

(그림 3) 단백질 구조 가시화 시스템의 가시화 성능

은 단백질 표면 생성도 모두 정확하게 만들어 졌고, 정점 구조도 잘 정렬되어 기존의 방법보다 좋은 성능을 보이고 있다. 표면모델 가시화는 기사 인스턴싱이 적용되지 않아 Proposed 방법과 Proposed-2 방법의 성능이 같다.

### 6. 결론 및 향후 연구

본 연구에서 제안하는 단백질 3차원 구조 가시화 시스템은 거대 단백질을 빠른 속도로 가시화하고 여러 단백질을 동시에 가시화 하기 위하여, 3차원 그래픽 하드웨어에 최적화를 수행하였다. 최적화 기법으로 단백질의 3차원 렌더링 특성을 이용한 기하 인스턴싱 기법을 사용하여 단백질의 3차원 구조를 렌더링 하였다. 성능 실험에서 단백질 구조 가시화 시스템은 약 50만개의 원자를 가진 단백질에서 10 fps 이상의 우수한 성능을 보이고 있다.

향후 연구로 단백질 가시화의 질을 향상시키기 위하여 PRT 등과 같은 컴퓨터 그래픽 기법을 이용하여 단백질 렌더링의 질을 높일 예정이다.

### 참 고 문 헌

[1] R. Peter, J. George, M. Kenneth, L. Jonathan, S. Susan, 'Biology,' 8th ED. McGraw-Hill, 2007.

[2] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng and T.E. Ferrin, "UCSF Chimera - A Visualization System for Exploratory Research and Analysis," Journal of Computational Chemistry, Vol.25, No.13, pp.1605-1612, 2004.

[3] H. William, D. Andrew, and S. Klaus, "VMD - Visual Molecular Dynamics," Journal of Molecular Graphics, Vol.14, pp.33-38, 1996. (<http://www.ks.uiuc.edu/Research/vmd/>)

[4] <http://www.accelrys.com/>

[5] <http://pymol.org>

[6] <http://jmol.sourceforge.net/>

[7] W. Kabsch and C. Sander, "DSSP: definition of secondary structure of proteins given a set of 3D coordinates," Biopolymers, Vol.22, pp.2577-2637, 1983. (<http://swift.cmbi.ru.nl/gv/dssp/>)

[8] A.M. Lesk and K.D. Hardman, "Computer-generated schematic diagrams of protein structures," Science, Vol.216, pp.539-540, 1982.

[9] M. Carson and C.E. Bugg, "Algorithm for Ribbon Models of Proteins," Journal of Molecular Graphics, Vol.4, No.2, pp.121-122, June, 1986.

[10] M. L. Connolly, "Analytical molecular surface calculation," Journal of Applied Crystallography, vol.16, No.5, pp.548 - 558, Oct., 1983.

[11] L. Laurence, T. Annick, B. Robert, "Analysis of accessible surface of residues in proteins," Protein Science, Vol.12, No.7, pp.1406-1417, 2003.

[12] <http://local.wasp.uwa.edu.au/~pbourke/geometry/polygonise/>



### 박 찬 용

e-mail : cypark@etri.re.kr

1994년 광운대학교 컴퓨터공학과(학사)

1996년 광운대학교 컴퓨터공학과(석사)

2007년 충남대학교 컴퓨터과학과(박사)

2008년 미 UCSB 박사후 과정

1995년~현 재 한국전자통신연구원

선임연구원

관심분야: 컴퓨터 그래픽스, 바이오인포매틱스, 단백질 구조 등



### 황 치 정

e-mail : cjhwang@cnu.ac.kr

1975년 2월 서강대학교 수학과

1985년 2월 코네티컷주립대학교 전산학과

(석사)

1987년 2월 코네티컷주립대학교 전산학과

(박사)

1987년 2월 코네티컷주립대학교 객원교수

1999년 2월 충남대학교 전자계산소 소장

1988년 2월~현 재 충남대학교 전기정보통신공학부 교수

관심분야: 영상처리, 컴퓨터 비전, 컴퓨터그래픽스