특집

의료영상에서의 GPGPU활용

강동수·신병석 (인하대학교)

I. 서 론

1980년대부터 CT(Computed Tomography)와 MRI(Magnetic Resonance Imaging)를 시작으로 PET(Positron Emission Tomography), 초음파 영상장치(Ultrasound)등 다양한 단층촬영 장치들이 개발되면서 우리의 몸속을 입체적으로 관찰할 수 있는 계기가 마련되었다. 무엇보다도 영상처리 기술과 컴퓨터 그래픽스 기술이 급속도로 발전하면서 단순히 2차원에 머물던 의사들의 시선이 3차원 영상 진단 쪽으로 옮겨가게 되었다. 3차원 영상 진단이란 인체를 연속 단층촬영한 후 단면 영상들을 종축으로 쌓아서 이산 샘플(discrete samples)들의 3차원 집합인 볼륨데이터를 만들고 여기에 특수한 렌더링 기술을 적용하여 실물을 보는 것과 같은 3차원 영상을 만들어 내는 것이다.

초기의 3차원 의학 영상들은 기본이 되는 단면영상의 조악한 품질과 IT기술의 미비로 말미암아 의사들의 주목을 끌지 못했다. 무엇보다 '정확성'을 생명으로 하는 의료진단 분야에서 어설픈 3차원 영상은 실제 임상에 적용하기에는 부족할 수밖에 없었다. 그러나 하드웨어 기술, 특히 그래픽처리장치(GPU : Graphics Processing Unit)와 컴퓨터 그래픽 기술이 급속도로 발전하면서 현재에는 1㎜이내의 해상도를 가지는 multi-detector CT로부터 얻어진 데이터를 고해상도의 영상으로 만들어 낼 수 있게 되었으며, 초당 10~30프레임의 속도를 보장하는 영상기법들이 만들어지고 있다. 또한, 프로브(prove)로부터 초당 30장 이상의 영상을 획득할 수 있는 기계적 특성을 지닌 초음파 장치에도 GPU기술을 활용하는 실시간 필터링 기술과 3차원 가시화 연구가 진행되고 있다.

3차원 의학 영상은 영상유도 수술(image-guided surgery), 가상 내시경(virtual endoscopy), 가상 수술 시뮬레이션(virtual surgical simulation)과 같이 이전에는 없던 새로운 기술 분야를 창출하며 21세기의 새로운 영역으로 발전해 나가고 있다.

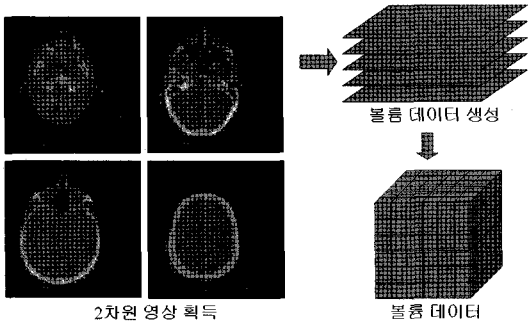
여기서는 의료영상 분야 중에서 가상내시경과 4차원 초음파영상에 대해 소개하고, 여기에 GPGPU(General-Purpose computation on Graphics Processing Units)의 활용사례를 소개하고자 한다.

II. CASE 1 : 가상 내시경

1. 가상 내시경 기술

가상 내시경은 CT, MRI와 같은 전산화 단층 촬영 장치로 인체 내부에 대한 연속적인 단면 영상을 얻어낸 후 인체의 3차원적 구조를 재구성함으로써 내시경 카메라로 보는 것과 같은 가상의 3차원 영상을 만들어내는 기법을 말한다. 가상 내시경을 이용하면 환자에게 고통을 주지 않으면서 외상이나 감염의 우려 없이 대장, 기관지, 혈관과 같은 관상(pipe-like) 인체기관의 내부를 3차원 영상으로 재구성 할 수 있다.

가상 내시경을 위해서는 2차원의 영상들로부터 3차원 볼륨모델을 생성해야 한다. 볼륨 데이터는 복셀(voxel)이라 불리는 규칙적인 값들의 3차원 배열로 구성되며, 이것은 <그림 1>에서 보는 바와 같이 영상(image) 혹은 슬라이스(slice)라 불리는 2차원 배열들의 스택 형태를 갖는다. 이 배열의 값들은 3차원 구조에서 특정 위치 (x,y,z)의 물리적 특성을 표현하는 색상, 밀도, 압력 등의 스칼라값 혹은 속도와 같은 벡터값이 될 수 있는데, 의료영상에서는 주로 장기의 밀도값을 표현한다. 최근의 영상장치는 일반적으로 512×512 픽셀 해상도의 영상을 0.5~0.1㎜ 사이 간격으로 획득한다.



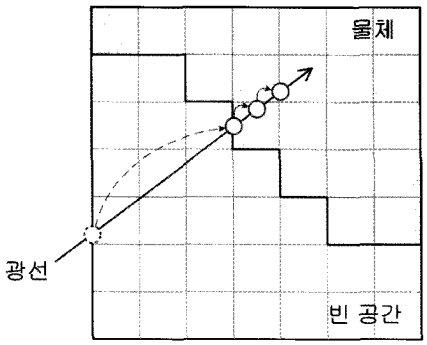
〈그림 1〉 CT, MRI에서 촬영한 영상으로부터 볼륨데이터를 재구성하는 과정

볼륨 데이터를 이용한 가상 내시경의 가시화 기술은 크게 표면 렌더링 기법[1]과 직접 볼륨 렌더링 기법[2]으로 나눌 수 있다. 표면 렌더링은 장기를 삼각형 메쉬(mesh)로 만든 후, 이것을 범용 그래픽가속기를 이용하여 렌더링하는 기법으로서 대표적인 방법으로는 Marching Cubes 알고리즘이 있다[1]. 표면 렌더링 방법은 범용의 그래픽가속기를 이용하기 때문에 빠른 속도로 렌더링을 할 수 있지만 메쉬를 생성하는 시간이 오래 걸리고 인체와 같이 반투명한 조직들로 구성된 물체를 표현하는 데는 적합하지 않다. 직접 볼륨 렌더링 기법은 볼륨 데이터 자체를 중간 처리과정 없이 직접 렌더링하는 기법으로서 크게 객체 기반 기법(object-based method)과 영상 기반 기법(image-based method)으로 나눌 수 있다. 객체 기반 기법의 대표적인 방법은 스플래팅(splatting) 이다[3]. 스플래팅 기법은 복셀이 임의의 크기를 가지고 있는 물체로 가정하여, 화면에 투영하는 기법이다. 최근에는 비기하 데이터(non-geometry data)인 점 데이터(point data)를 이용하여 렌더링하는 기법(point rendering) 기법도 많이 사용하고 있다.

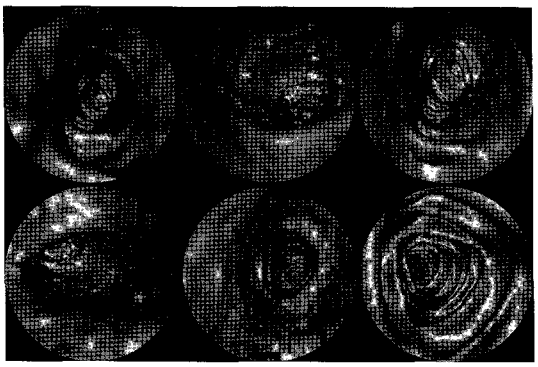
대표적인 영상기반 기법으로는 볼륨 광선 투사법(volume ray-casting)이 있다[2]. 볼륨 광선 투사법은 크게 두 가지 단계로 이루어진다. 첫 번째로, 영상의 픽셀에서 광선을 발사하여 빈 공간으로 간주되는 볼륨 영역을 도약한다. 두 번째로, 광선이 관심영역에 도달하면 직접 값을 샘플링하여 누적된 색상 값과 불투명도를 얻어내게 된다. 불투명도 값이 임계값에 도달하면 한 광선에 대한 색상계산이 완료된다. 이와 같은 과정을

영상의 모든 픽셀에 대해서 수행하게 된다. <그림 2>는 볼륨광선 투사법의 과정에 보여주고, <그림 3>은 본 연구팀에서 개발한 가상 내시경 시스템으로 대장을 렌더링한 결과이다.

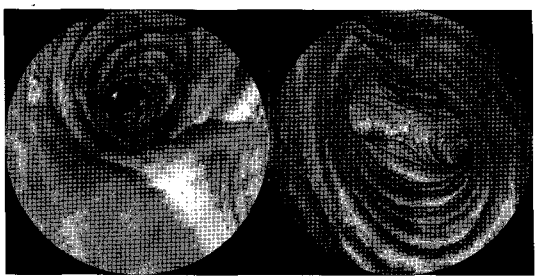
<그림 4>는 인체의 고해상도 컬러 영상데이터인 VKH(Visible Korean Human) 데이터를 이용하여 기관지를 렌더링한 영상이다.



〈그림 2〉 볼륨 광선 투사법



〈그림 3〉 가상 대장 내시경 결과 영상



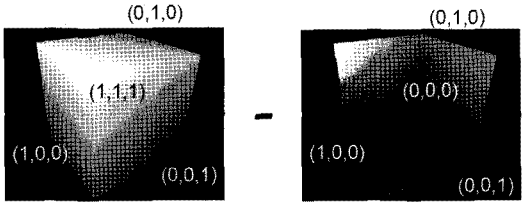
〈그림 4〉 VKH데이터를 이용한 가상 기관지 내시경

2. 가상 내시경 분야에서의 GPGPU 활용

볼륨 렌더링은 3차원 데이터를 분석하기 위한 강력한 도구이다. 하지만, 일반적으로 볼륨 데이터의 크기가 크기 때문에 렌더링 시간이 오래 걸린다는 단점을 가지고 있다. 따라서 대화식 수준의 볼륨 렌더링 속도를 내기위해서는 많은 하드웨어의 지원이 요구되며, 이런 이유로 초기의 볼륨 렌더링은 대용량 워크스테이션, 병렬 컴퓨터 또는 특수 목적 하드웨어에서만 수행되었다. 하지만 최근에는 GPU 성능의 향상으로 인하여 일반 그래픽 하드웨어에서 대용량 볼륨 렌더링의 가속화 연구가 본격화 되고 있다.

최근 연산 능력이 매우 커진 GPU는 3차원 의료영상 데이터를 매우 빠르게 가시화할 수 있도록 한다. 의료영상의 빠른 가시화는 질병에 대한 진단을 빠르고 효과적으로 할 수 있게 한다. GPU를 활용한 볼륨 렌더링의 대표적인 사례로는 2003년에 Krüger가 제안한 GPU기반의 볼륨 광선 투사법이 있다[4]. 해당 기법은 다단계 렌더링(multi-pass) 방법으로 GPU 내부의 그래픽스에 특화된 병렬 처리 장치인 정점 쉐이더(vertex shader)와 픽셀 쉐이더(pixel shader)를 사용하였다. 이것은 기존의 CPU에서 직렬로(serial) 처리되었던 모든 수행과정을 GPU를 이용하여 병렬(parallel)로 가속화하는 방법이다. 병렬 처리를 통한 렌더링의 가속화 부분은 크게 광선의 방향 계산과 각 광선별 색상 누적이 있다. 우선, 기존의 픽셀별 광선 방향의 계산은 GPU의 정점 쉐이더를 이용하여 <그림 5>와 같이 2차원 텍스처로부터 병렬로 계산된다.

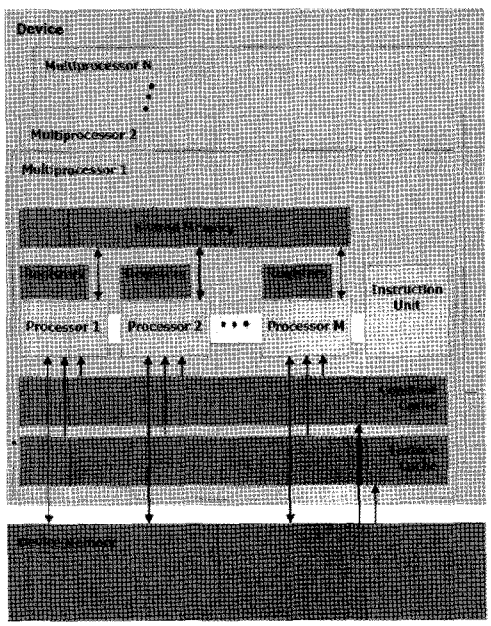
다음으로 각 픽셀별 광선 투사 및 색상과 불투명도 값의 누적 역시 하나의 픽셀별로 하나의 쓰레드(thread)가 할당되어 빠른 부동 소수점 연산



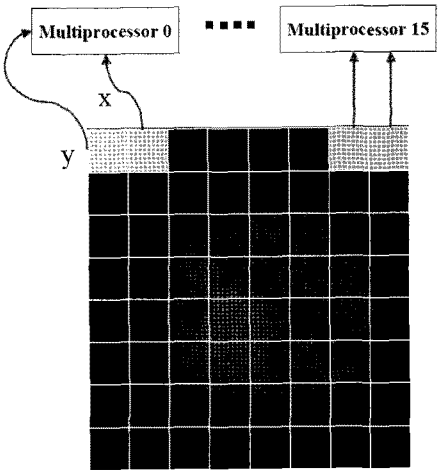
〈그림 5〉 GPU를 이용한 광선 방향 계산

이 가능한 GPU내에서 고속으로 처리된다. 현재 최신형 그래픽 카드에서 256^3크기의 볼륨데이터를 100*fps* 정도의 속도로 렌더링 할 수 있다.

최근에는 NVIDIA에서 개발한 새로운 아키텍처인 CUDA(Computer Unified Device Architecture)를 이용한 볼륨 렌더링 기법이 많이 연구되고 있다. CUDA는 기존의 그래픽연산처리에 특화되어 있는 GPU를 비그래픽스(non-graphics)분야에도 적용할 수 있도록 하는 범용성을 가진 병렬처리 아키텍처이다. CUDA는 <그림 6>과 같이 SIMD 멀티프로세서를 갖고, 각 프로세서들



〈그림 6〉 CUDA 하드웨어 아키텍처

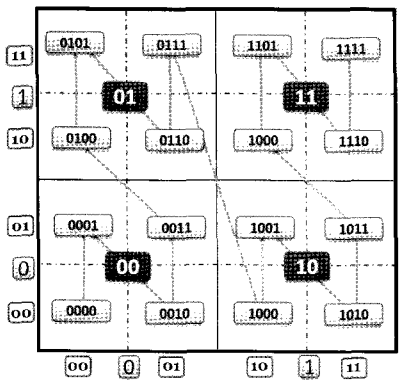


〈그림 7〉 CUDA에서의 광선 투사법

은 작은 지역 공유 메모리(shared memory)와 상수 캐시(constant cache), 그리고 프로세서들로 구성되어 있다[5].

CUDA기반의 광선 투사법은 기존의 CPU기반의 광선투사법을 병렬처리로 바꾼 형태이다. 즉, <그림 7>과 같이 최종 프레임버퍼(frame buffer)에 출력될 이미지의 각 픽셀을 하나의 쓰레드로 할당하여 광선을 진행하면서 불투명도 전이함수(OTF : Opacity Transfer Function)에 따라 색상을 누적한다. 그리고 동일 블록내의 쓰레드와 빠른 데이터 전송이 가능한 공유 메모리에는 OTF정보를 저장한 배열과 각 광선들의 방향이 저장되어 광선 투사를 더욱 가속화 한다[5].

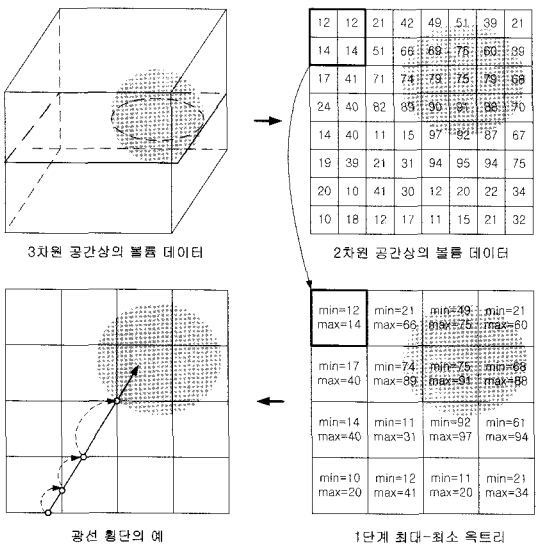
CUDA를 이용한 볼륨 렌더링의 가속화 방법에는 옥트리 기반의 빈 공간 도약과 광선 조기종료 기법이 있다[6]. 옥트리 기반의 빈 공간 도약은 기존 쉐이더 기반의 광선 투사법에서 그래픽적인 부분에 특화된 구조 때문에 적용이 불가능하였던 것에 비해 CUDA에서는 해당 기법을 트리를 이용하여 연속적인 배열 형태로 변형하여 적용할 수 있다[7]. 이 기법은 SFC(Space Filling



〈그림 8〉 CUDA에서의 space filling curve

Curve) 알고리즘을 이용하여 옥트리 기반의 데이터를 <그림 8>과 같이 연속적인 배열로 변형한 뒤, 각 노드를 이진화하여 GPU에서 처리 가능하도록 연속적인 순서로 재배치하는 방법이다. 이것은 병렬처리 시스템 상에서 옥트리를 상향식으로 구성하기 위한 하나의 자료구조이다.

CUDA내에서 SFC를 통해 변형된 데이터는 한 레벨 위의 노드 개수만큼 쓰레드를 생성하고, 각 쓰레드는 자신이 포함하는 8개의 하위 노드들을 비교하여 그 중 최대값과 최소값을 저장하는 방식으로 상위 노드를 생성한다. 해당 과정을 사용자가 원하는 크기의 옥트리 레벨까지 반복적으로 수행한다. 이렇게 만들어진 옥트리 형태의 자료구조는 <그림 9>와 같이 렌더링 단계에서 각 픽셀별로 발사된 광선이 점진적으로 진행하면서 위치한 공간의 최대/최소값을 바탕으로 공유메모리에 저장된 OTF값과 비교하여 해당 공간을 도약할지, 색상을 샘플링하여 누적할 지를 결정한다.



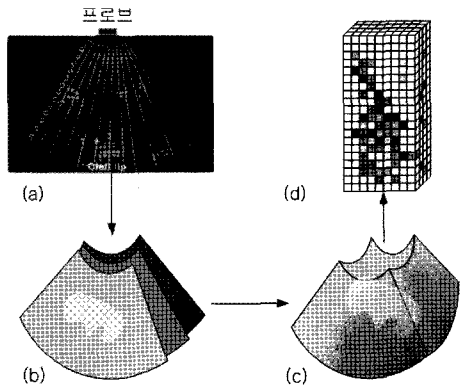
〈그림 9〉 최대-최소 옥트리

이런 가속화 방법은 혈관 렌더링과 같이 빈 공간이 많은 볼륨데이터의 가시화에 효과적이다. 마지막으로 광선 조기종료는 CUDA의 각 쓰레드별로 불투명도가 임계값에 도달할 때까지 색상 누적을 수행한 후 광선진행을 멈추도록 구현함으로써 적용할 수 있다. 추가적인 가속화 기법으로는 최근 블록 단위로 빠른 데이터 전송 속도를 제공하는 공유메모리의 사용률을 높이고, 뱅크충돌(bank conflict)을 줄이는 병렬프로그램의 최적화 방법이 많이 연구되고 있다[8].

III. CASE 2 : 4차원 초음파 가시화

1. 초음파 의료진단 기술

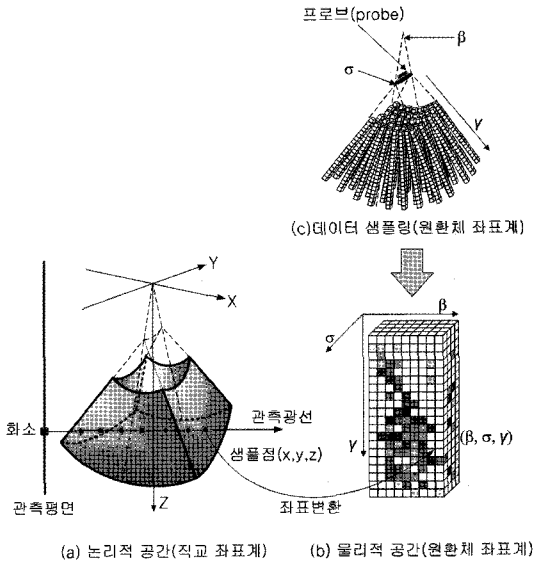
초음파 영상은 진단과 치료 목적에서 X선 영상과 더불어 가장 널리 사용되고 있다. 초음파 영상은 인체의 내부에 초음파를 발사한 후 반향된 신호를 시각화하여 인체 내부의 형상을 볼 수 있도록 하는 투시방법이다. 초음파 영상은 컴퓨터 단층 촬영기 또는 자기 공명 장치를 이용한 영상에 비해 화질은 떨어지지만 비용이 저렴하여 많



〈그림 10〉 초음파 볼륨 데이터 획득과정

이 사용되고 있다. 최근에는 초음파 데이터 획득 장치에 모터를 장착하여 특정 구간을 연속촬영 하도록 함으로써 3차원 볼륨데이터를 획득 할 수 있게 되었고, 이에 따라 초음파 볼륨가시화 방법에 대한 연구가 필요하게 되었다. 이러한 초음파 영상 데이터는 일반 CT나 MRI와는 상이한 방식으로 스캐닝을 하기 때문에 직교 좌표계가 아닌 동심구(concentric sphere) 좌표계를 사용하여 데이터를 표시한다[9]. 따라서 적절한 좌표변환 과정이 필요하다. 초음파 볼륨 데이터는 <그림 10>에서 보는 것과 같이 프로브(probe)에서 원환체 형태로 초음파 빔을 순차적으로 방사한 후(a), 프로브에서 얻은 연속된 2차원 영상들을 모아서(b) 초음파 볼륨 데이터를 만들고(c) 이것을 메모리에 순차적으로 저장한다(d).

초음파 데이터를 렌더링하기 위해서는 <그림 11>에서 보는 것과 같이 관측광선이 진행하는 논리적 좌표 공간(a)과 물리적 메모리에 데이터가 저장된 원환체 좌표 공간(b) 사이의 좌표변환 과정이 필요하다. 초음파 데이터의 가시화는 일반 볼륨모델의 렌더링 기법인 광선 투사법을 사용한다. 하지만 초음파 볼륨 모델의 경우 직교 좌표계를 사용하지 않기 때문에 옥트리와 같은 자



〈그림 11〉 초음파 볼륨 렌더링의 좌표변환

료구조를 생성할 수 없는 점과 획득된 영상의 화질이 좋지 않은 문제점을 가진다. 이것을 해결하고자 초음파 가시화 분야에서는 GPU를 이용한 고속의 스캔변환(scan conversion) 방법과 실시간 3차원 필터링 방법이 많이 연구되고 있다. 또한, 최근의 초음파 볼륨 데이터 획득 장치는 심장의 움직임, 태아의 모습 등을 실시간으로 획득할 수 있을 정도로 빠르기 때문에 렌더링도 실시간에 되어야 하며, 이를 위해 GPU를 이용한 4차원 고속 초음파 렌더링 기법이 필요하다[10].

2. 초음파 데이터 렌더링의 GPGPU 활용

초음파 렌더링분야에서 GPU의 활용에는 계산량이 많은 스캔변환의 고속화와 초음파 데이터의 실시간 필터링이 있다. 일반 볼륨데이터와는 다르게 초음파 렌더링에서는 좌표변화를 위한 스캔변환 과정이 필요하다. 하지만, 스캔변환은 앞 절에서 언급한 것처럼 광선이 한 단계씩 진행

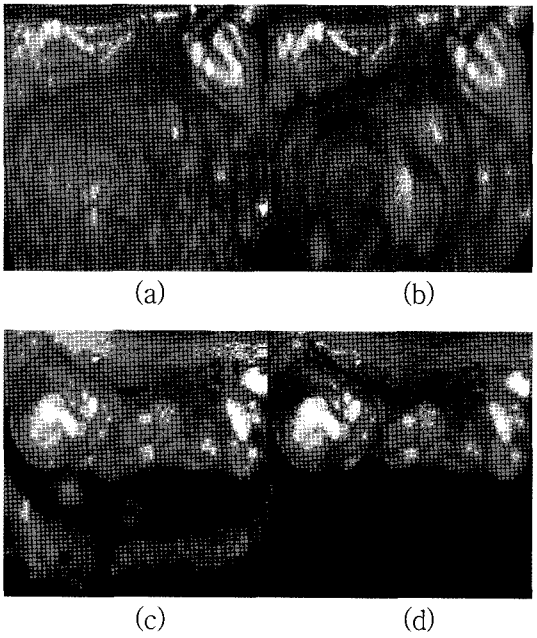
할 때 마다 픽셀별로 반복 수행해야 하기 때문에 부하가 클 뿐만 아니라, 동심구 좌표계 계산에 따른 부동 소수점 연산으로 CPU에서는 실시간 처리가 어렵다. 그러므로 현재 초음파 렌더링 프로그램에서는 스캔변환을 픽셀 쉐이더나 프레그먼트 쉐이더(fragment shader)에서 병렬로 처리하고 있다. 이것은 60*fps* 이상으로 렌더링할 수 있도록 한다.

또 다른 응용으로는 3차원 초음파 데이터의 실시간 필터링이 있다. 기존의 그래픽 특화된 GPU내에서는 간단한 3차원 텍스처의 선형 보간(interpolation)만 가능할 뿐, 잡음이 많은 초음파 데이터에 대한 실시간 필터링이 불가능하다. 이런 이유로 대부분의 시스템에서는 픽셀 쉐이더를 이용하는 후처리 필터링이나 OpenMP 또는 MMX(MultiMedia eXtension)와 같은 CPU 병렬처리 시스템을 이용한 볼륨 필터링 기법이 사용된다. 하지만 후처리 필터링의 경우, 최종 출력된 2차원 이미지만을 가공하는 것이기 때문에 임상적 요구를 만족시키기에는 한계가 있다. 또한, CPU 기반의 필터링의 경우에는 필터링 과정을 각 복셀별로 병렬처리 하여도 대용량의 3차원 초음파 데이터를 실시간으로 처리하기에는 역부족이다.

일반적으로 프로브로부터 데이터를 획득하는 초음파 획득 장치는 대략 초당 30장의 초음파 데이터를 렌더링 엔진으로 전송한다. 이렇게 전송된 데이터를 진단을 위한 의료영상으로 사용하기 위해서는 대화식 속도를 보장하면서 적절히 필터링 되어야 한다. 이런 이유로 현재 초음파 영상 가시화 분야에서는 실시간으로 렌더링하기 위한 연구가 GPGPU 기술을 이용하여 활발히 진행되고 있다. 현재 CUDA 기반의 3차원 초음파 데이터의 필터링 기법은 기본적으로 20*fps* 이상의 속도를 보장하며, 초음파 데이터의 특성에 맞도록 좌표계 기반의 필터링과 OTF기반의 필터링이 많이 사용된다. 좌표계 기반의 필터링은 관찰하고자 하는 장기의 위치가 프로브가 지나가는 동심구 좌표계 내에 어느 부분에 위치하는지를 미리 계산하여 해당 영역을 미리 추려내는 필터링 기법이다. 그리고 OTF기반의 필터링은 모든 복셀을 각 쓰레드로 계산하여 측정하고자 하는 물체에 대한 OTF를 기준으로 각 복셀별로 필터링의 적용유무를 선별하는 기법이다. <그림 12>는 본 연구진에서 수행하였던 실시간 필터링기법을 적용하여 렌더링한 영상이다. 해당 결과물은 초당 30장의 영상을 획득하는 초음파 스캐너가 장착된 시스템을 기반으로 CUDA를 이용하여 데이터를 실시간 필터링을 한 후, 초당 20 ~25*fps* 정도로 초음파영상을 3차원으로 가시화하는 시스템의 결과 영상이다. 사용된 시스템은 DirectX 기반의 렌더링 코어와 CUDA기반의 선 필터링 기법을 통합한 프레임워크를 가진다. 실험은 NVIDIA GeForce GTX 260 그래픽 장치와 Intel Core2Duo 6400을 갖춘 시스템에서 수행되었고, 256^3의 팬텀 데이터에서 특정 필터링을 위해 실시간으로 5×5×5 크기의 좌표계 기반 필터 커널을 6회, 2×2×2 크기의 OTF기반 필터 커널을 2회 적용시킨 결과영상이다.

Ⅳ. 향후 전망

그래픽 연산에 특화되어 있던 GPU가 범용 목적의 하드웨어로 진화되면서 의료영상 분야에서도 많은 변화가 일어나고 있다. 특히, 진단목적의 CT, MRI 그리고 초음파를 이용한 3차원 의료영상 가시화 분야에서 많이 사용되는 볼륨 광선 투



(a) (b)

(c) (d)

〈그림 12〉 초음파 영상의 실시간 3차원 필터링 및 렌더링 ((b)–25*fps*, (d)–28*fps*); (a)와 (c)는 원본 영상, (b)와 (d)는 필터링 된 영상이다

사법이 CUDA와 같은 GPGPU 응용으로 구현되면서, 기존에 사용하지 못했던 여러 가지 복잡한 기법들을 적용할 수 있게 되었다. 특히, 3차원 데이터의 실시간 필터링이나 부동소수점 연산이 많이 필요한 스캔변환 등이 대표적인 예라고 할 수 있다. 뿐만 아니라 실제 의학 영상 응용에서 그래픽 처리에 대해 CPU 사용률을 최소화 하고, 기존에 CPU에서 처리했던 여러 복잡한 부동소수점 계산을 GPU에서 빠르게 처리함으로서 CPU를 다른 여러 곳으로 분산하여 사용할 수 있는 여건을 제공하게 되었다. GPGPU는 향후 대용량 데이터를 기반으로 하는 고해상도의 CT나 MRI의 데이터 처리 등 여러 의학 영상 분에에서 큰 기여를 하게 될 것이다.

참고문헌

[1] Lorensen, W. and Cline, H., "Marching Cubes: A High Resolution 3D Surface Construction Algorithm," Proc. ACM SIGGRAPH, pp.163-169, 1987.

[2] Levoy, M., "Display of Surfaces from Volume Data," IEEE Computer Graphics and Applications, Vol.8, No.3, pp.29-37, 1988.

[3] Rusinkiweicz, S., Levoy, M., " QSplat: A Multiresolution Point Rendering System for Large Meshes," Proc. ACM SIGGRAPH, pp.343-352, 2000.

[4] Krüger, J., Westermann, R., "Acceleration Techniques for GPU-based Volume Rendering," Proceedings of IEEE Visualization, pp.38-46, 2003.

[5] NVIDIA CUDA Compute Unified Device Architecture Programming Guide, v. 2.0, 2008.

[6] Lim, S., Shin, B., "RPO: A Reverse-Phased hierarchical min-max Octree for efficient space leaping," Proceeding of Pacific Graphics, pp.145-147, 2005.

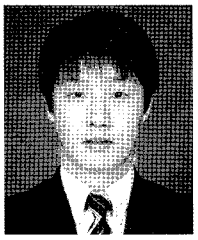
[7] Ajmera, P., Goradia, R., Chandran,S., Aluru, S., "Fast, Parallel, GPU-based Construction of Space Filling Curves and Octrees," Proc. Interactive 3D graphics and games, 2008.

[8] Marsakec, L., "High-Speed Volume Ray Casting with CUDA," Interactive Ray Tracing, 2008.

[9] 권의철, 이재근, "실시간 3차원 초음파 진단 시스템용 렌더링 장치 및 방법," 한국 특허-KR20040078352A.

[10] Kuo, J.,Bredthauer, G.R., Castellucci, J.B., Ramm, O.T., "Interactive Volume Rendering of Real-Time Three-Dimensional Ultrasound Images," IEEE Transaction on Ultrasonics, Vol.54, pp.313-318, 2007.

저자소개



강 동 수

2005년 2월 인하대학교 컴퓨터 · 정보공학과(학사)
2007년 2월 인하대학교 컴퓨터 · 정보공학과(석사)
2008년 3월 ~ 현재 인하대학교 컴퓨터·정보공학과 박사과정

주관심 분야 : 의료영상, 볼륨 그래픽스, 병렬처리

신 병 석

1990년 2월 서울대학교 컴퓨터공학과(학사)
1992년 2월 서울대학교 컴퓨터공학과(석사)
1997년 2월 서울대학교 컴퓨터공학과(박사)
2000년 3월 ~ 현재 인하대학교 컴퓨터공학부 부교수
2003년 1월 ~ 현재 한국 컴퓨터 그래픽스 학회 학술이사
2005년 1월 ~ 현재 한국 차세대 PC학회 부회장

주관심 분야 : 의료영상, 실시간 렌더링, 볼륨 그래픽스