특집 03

다중 영상으로부터의 3차원 형상 복원 기법

목 차

1. 서 론
2. 미들베리(middlebury) 스테레오 비전 홈페이지
3. 주요 알고리즘 소개 및 성능 비교
4. 결 론

박인규 · 장주용
(인하대학교 · 삼성전자)

1. 서 론

영상으로부터 3차원 형상을 복원하는 것은 컴퓨터 비전 및 컴퓨터 그래픽스의 고전적인 문제이다. 3차원 형상 복원 기법의 응용 분야는 로봇 분야, 3차원 애니메이션, 문화재 복원, 도시 모델링, 전자 상거래, 엔터테인먼트, 건축, 예술 분야에 걸쳐 광범위하므로 현재에도 매우 활발히 연구가 진행되고 있다. 현재까지 제안된 수많은 기법들은 조명이나 레이저, 프로젝터등과 같은 부가적인 발광 장비를 이용하는 능동적 기법과, 오직 영상들 또는 비디오 프레임들을 이용하는 수동적 기법으로 나뉜다. 이중 수동적 기법은 카메라만을 이용하기 때문에 취득할 수 있는 물체 또는 장면의 범위에 제한이 없고 상대적으로 저렴한 가격으로 시스템을 구축할 수 있는 장점이 있으나, 능동적 기법에 비하여 취득된 데이터의 정밀도가 떨어지는 단점이 있다.

수동적 기법 중에서 실제 산업 분야에 활발히 응용되는 기법은 (1) 영상 두 장을 사용하여 깊이 영상(depth image)을 생성하는 양안 스테레오 (binocular stereo) 기법, (2) 시점이 다른 여러 장의 영상을 사용하여 기준 영상에 대한 깊이 영상 또는 절대 좌표계에서의 전체적인 3차원 형상을 생성하는 다중뷰 스테레오(multi-view stereo) 기법, (3) 순차적으로 변하는 시점 또는 장면에 대한 비디오 프레임으로부터 추출된 특징점의 3차원 정보와 카메라의 위치 및 자세를 추정하는 structure from motion (SFM) 기법 등이 있다. 이러한 기법의 완전한 이해에 필요한 이론적인 배경은 대부분의 관련 대학원 교과목에서 채택되어 강의되는 Hartley와 Zisserman의 교과서 [1] 및 Moons의 최근 서베이 논문 [2]를 참고하기 바란다.

고전적인 양안 스테레오 기법은 로봇 비전에서 많이 이용되며 매우 정확한 3차원 거리 추정보다는 대략적이고 상대적인 거리 추정 및 비교에 적당하다. 따라서 주행 중인 로봇이나 자동차에 부착된 두 대의 카메라로부터 장애물 인식, 충돌 방지 등의 분야에 응용 가능하다.

최근 10년간 많은 연구가 진행된 다중뷰 스테레오 기법은 기존의 양안 스테레오(binocular stereo) 기법에 비해 매우 정확한 3차원 형상을 취득할 수 있다. 최근 기법은 640×480 정도의 저

해상도 입력 영상에 대해 물체 크기 대비 0.5% 이내의 적은 오차를 보인다. 본 기법은 사전에 카메라 보정(camera calibration)이 필요하고 전반사(specular reflection)등에 의한 잡음을 줄이기 위하여 스튜디오와 같이 카메라의 위치, 조명의 위치와 속성 등을 제어 가능한 환경에 주로 이용되지만, 최근에는 도시 모델링과 같이 대규모 장면에 대한 3차원 복원에도 응용되는 추세이다.

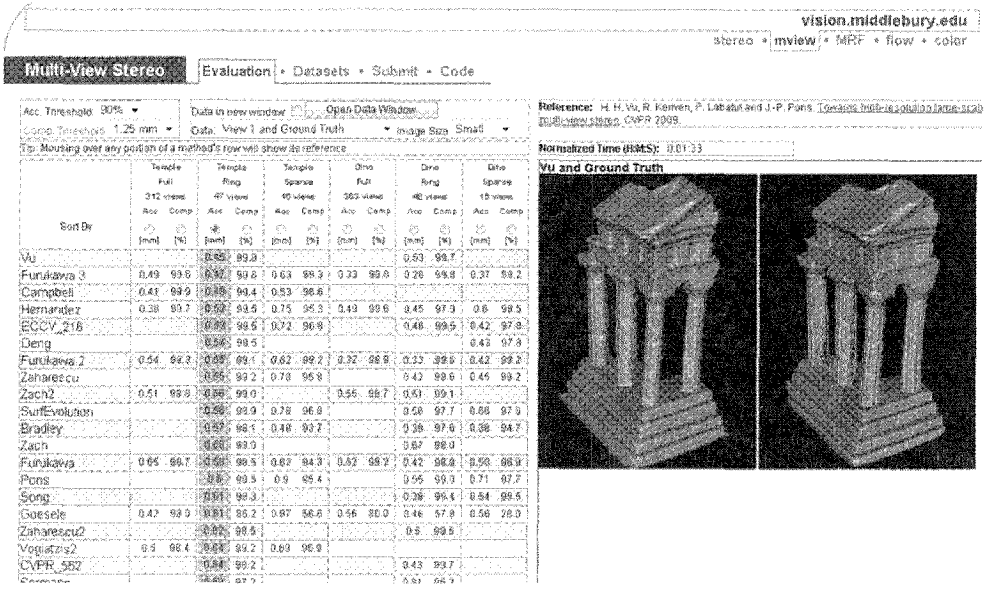
Structure from motion 기법에서의 핵심 기술은 각 프레임에서의 안정적인 특징 추출(feature extraction) 및 추적(tracking)을 통한 대응 문제 (correspondence problem) 해결이다. 일반적인 다중뷰 스테레오에 비해 본 기법은 비디오 프레임과 같이 시간 축으로 조밀하게 샘플링 되어 장면의 변화가 적은 영상을 대상으로 하기 때문에 광류(optical flow)와 같은 기법을 통한 추적 기법이 활발하게 이용된다. 다만, 프레임간의 카메라 간격(baseline)이 작기 때문에 추출되는 3차원 정보의 정확도가 낮고, 추적하는 특징점의 개수가 영상 해상도에 비해 매우 적기 때문에 조밀한(dense) 3차원 복원을 위해 안정적인 후처리 알고리즘이 필요한 단점이 있다.

본 기고에서는 이 세 가지 기법 중에서 최근 가장 많은 관심을 받고 있는 다중뷰 스테레오 기법에 대해 보다 구체적인 기술 분류와 관련 논문 및 최신 연구 동향을 소개한다.

2. 미들베리(middlebury) 스테레오 비전 홈페이지 [3]

다중뷰 스테레오 기법은 최근 10년간 가장 활발히 연구가 진행된 분야이다. 보다 객관적인 성능 비교를 위하여 Seitz등은 미들베리 다중뷰 스테레오 홈페이지[3]를 통하여 Temple, Dino등의 참값(groundtruth)이 알려진 테스트 데이터를 제공한다. 또한 Seitz등은 정확도 (accuracy) 및 완성도(completeness)등의 두 가지 성능 측정 기준을 제시하고 2006년까지의 주요 알고리즘들의 성능을 이에 의해 평가하고 분석하였다[4].

개발된 주요 알고리즘은 모두 미들베리 홈페이지를 통해 성능이 객관적으로 비교되고 그 순위는 실시간으로 (그림 1)과 같은 테이블로 제



(그림 1) Middlebury 사이트에 게시된 다중뷰 스테레오 알고리즘의 성능 비교표

공되고 있다. 알고리즘 개발자가 'submit' 탭에 설명되어 있는 방식으로 복원된 형상 데이터를 담당자(Daniel Scharstein)에게 이메일로 전송하면 담당자는 참값 데이터와의 비교를 수행하고 비교 결과가 수록된 홈페이지의 링크를 다시 보내주게 된다.

3. 주요 알고리즘 소개 및 성능 비교

본 기고에서는 기존의 다중뷰 스테레오 기법을 3차원 형상의 표현 방식에 따라 다음의 세 가지로 분류하고 각 기법의 특징과 장단점을 기술하고, 또한 2006년 이후 현재까지 미들베리 스테레오 사이트의 성능 평가에서 상위에 랭크된 주요 알고리즘을 선별하여 소개한다.

- 복셀(voxel)과 같은 분할된 공간 요소를 통한 형상 표현 및 전역적 최적화 기법 [5][6]
- 초기 3차원 형상 정보로부터 변형 가능한 (deformable) 곡면의 반복적 수렴 기법 [7][8]
- Patch 또는 surfel을 이용한 3차원 표면 요소 모델링 및 영역 확장 기법 [9][10]

소개하는 개별 알고리즘의 성능은 [3]에 제시된 데이터를 〈표 1〉에 정리하여 수록하였다. 본 기고에서 제시한 형상 복원 실험 결과는 각 논문에 제시된 결과를 인용하였다.

3.1 복셀(voxel)과 같은 분할된 공간 요소를 통한 형상 표현 및 전역적 최적화 기법

복셀과 같이 균일하게 분할된 공간 요소는 3차원 장면의 표현 방식이 매우 단순하고 구현이 쉬우며 임의의 형상을 표현할 수 있다는 장점을 가지므로 가장 일반적으로 사용된다. 또한 이러한 표현 방법은 특정 뷰에 독립적이고 부드러운 곡면을 만들기 위한 제한 조건을 쉽게 모델링할 수 있기 때문에 알고리즘의 개발이 용이하다.

그래프 컷 (graph cuts) 기법은 이산 그래프 (discrete graph)에서의 최대 흐름 (maximum flow) 문제를 리만 기하학에서의 최소 곡면 (minimal surface) 추출 문제에 적용한 조합적 최적화 (combinatorial optimization) 기법으로서 2차원 또는 3차원 영상의 분할 (segmentation) 문제에 응용되어 왔다. 복셀과 같이 분할된 공간 요소를 이용하는 경우, 다중뷰 스테레오 문제는 결국 3차원 공간 요소의 집합을 물체와 배경으로 이진 분할하는 문제로 귀결된다.

본 분류에 속하는 기법들은 이와 같이 균일하게 분할된 공간 요소 표현 및 그래프 컷을 통한 전역적 최적화의 구조를 가지며 기존의 최대-흐름 (max-flow) 알고리즘을 이용하면 다룰 수 있는 (tractable) 계산량으로 임의의 형상 복원이 가능한 장점을 가진다.

〈표 1〉 주요 알고리즘의 성능 비교 [3]

알고리즘	Temple 데이터						Dino 데이터					
	Full (312 views)		Ring (47 views)		Sparse (16 views)		Full (363 views)		Ring (48 views)		Sparse (16 views)	
	Acc	Comp	Acc	Comp	Acc	Comp	Acc	Comp	Acc	Comp	Acc	Comp
Vogiatzis [5]	0.5	98.4	0.64	99.2	0.69	96.9	–	–	–	–	–	–
Sinha [6]	–	–	0.79	94.9	–	–	–	–	0.69	97.2	–	–
Vu [7]	–	–	0.45	99.8	–	–	–	–	0.53	99.7	–	–
Pons [8]	–	–	0.6	99.5	0.9	95.4	–	–	0.55	99.0	0.71	97.7
Furukawa [9]	0.49	99.6	0.47	99.6	0.63	99.3	0.33	99.8	0.28	99.8	0.37	99.2
Habbecke [10]	0.66	98.0	–	–	–	–	0.43	99.7	–	–	–	–

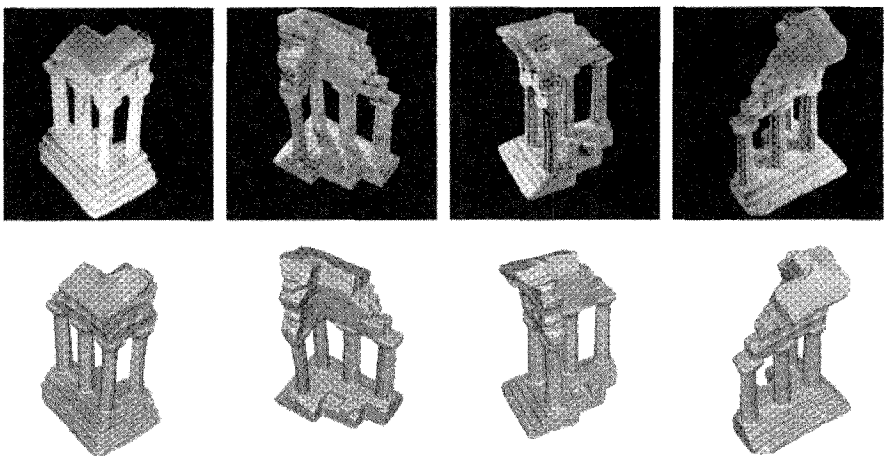
3.1.1 균일한 복셀 그리드(voxel grid)에서의 그래프 컷 기법 [5]

본 기법은 그래프 컷 기법을 다중 뷰 스테레오에 적용한 최초의 기법이며 이후 다른 연구자들의 알고리즘에 많은 영향을 주었다. 분할된 공간 요소에 대해 전경/배경 구분은 다중뷰로 투영했을 때 물체의 전경으로 투영되는지 여부에 의해 결정된다. 본 기법이 제안하는 영상-일관성(photo consistency)은 기존의 기법에 비해 강건한데, 카메라 중심과 기준 영상의 화소를 지나는 직선상에서 샘플링된 공간 요소를 각 다중뷰 영상에 투영하고 epipolar line을 따라 NCC (normalized cross correlation) 값을 계산하여 국부 최대값을 구한 후 Parzen 윈도우 기법을 통해 국부 최대치가 최대인 공간 요소를 찾아 곡면상의 곡면 요소로 간주한다.

그래프 컷 기법에서의 비용 함수를 구성하는 두 항인 불연속 비용은 곡면 표면에서의 영상-일관성 값의 적분으로 모델링하고, 전경/배경 비용은 각각 실루엣에 의해 결정되는 부피를 최대화하는 ballooning 항으로 모델링한다. 또한 그래프의 노드는 각 표면 요소를, 노드간의 에지는 두 노드가 나타내는 표면 요소 위치의 중점에서의 영상-일관성에 비례하는 값으로 결정된다. 이후 최대-흐름 알고리즘에 의해 각 표면 요소를 전경과 배경으로 분할하면 전경에 해당하는 표면 요소는 3차원 형상의 곡면을 구성하는 3차원 점을 구성하게 된다. (그림 2)에 본 기법으로 복원한 Temple 모델의 예제를 제시하였다.

3.1.2 적응적인 사면체 메쉬(tetrahedral mesh)에서의 그래프 컷 기법 [6]

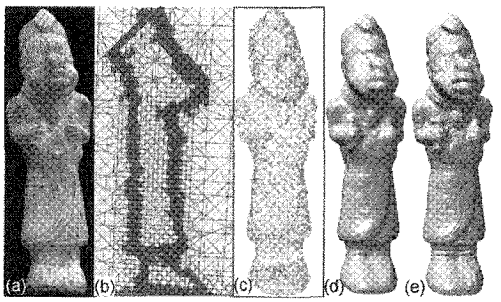
[5]와 같이 균일한 복셀 그리드에서의 그래프 컷은 실제 곡면 후보 여부에 관계없이 메모리를 소요하게 되므로 임의의 고해상도 그리드에서 그래프를 구성하기 어렵고 또한 그래프 구성이 그래프 컷 자체보다 더 많은 시간이 소요되기도 하는 단점을 가진다. 반면 [6]에서 제시하는 기법은 반 균일하면서 다중 해상도를 가질 수 있는 사면체 메쉬의 듀얼 (dual) 그래프에서의 그래프 컷으로 3차원 형상을 복원한다. 사면체 메쉬는 육면체 복셀에 비해 보다 더 다양한 법선 방향을 가질 수 있고 1:2, 1:4, 1:8등의 다양한 분할(subdivision)이 가능하다. 이에 의해 초기 사면체 메쉬의 각 면은 다중뷰에서의 영상-일관성 테스트에 의해 적응적으로 분할된다. 즉, 영상-일관성이 큰 점을 포함하는 사면체는 더 세부적



(그림 2) 균일한 복셀 그리드에서의 그래프 컷 기법[5]에 의한 형상 복원 결과

으로 분할되어 영상-일관성을 다시 테스트함을 반복한다. 이렇게 생성된 다중 해상도의 사면체 메쉬에 대해 듀얼 그래프를 생성하고 영상-일관성 항과 부피 확장 (ballooning) 항에 의해 비용 함수를 설정한 후 그래프 컷을 수행하게 된다. 이렇게 생성된 3차원 형상을 다중뷰 영상에 재투영하여 실루엣 제한 조건을 생성한 후 다시 한 번 그래프 컷을 수행하여 최종 3차원 형상을 복원하게 된다.

본 기법의 장점 중 하나는 실루엣 정보에 의한 비주얼 헐 (visual hull) 의 내부에 복원하는 곡면이 존재한다고 제한하지 않는 것이다. 또한 본 기법은 물체의 형상에 대한 가정을 하지 않고 충분한 영상에서 취득되었다면 심하게 움푹한 부분도 복원이 가능하다. 그래프 컷은 두 번 수행하는데 첫 번째 그래프 컷으로 얻어진 메쉬 모델은 실루엣 제한 조건에 의해 지역적 변형을 통하여 세부적인 기하정보까지 복원하는 구조를 가진다. (그림 3)에 본 기법에 의한 3차원 형상 복원의 사례를 제시하였다.



(그림 3) 적응적 사면체 메쉬 기반의 다중뷰 스테레오 기법[6]. (a) 입력영상 (36개중 하나), (b) 적응적 사면체 메쉬의 한 단면, (c) 메쉬 보정을 통해 얻어진 초기 패치 모델, (d) 그래프 컷을 통해 얻어진 3차원 형상, (e) 최종 3차원 형상.

3.2 변형 가능한 (deformable) 곡면의 수렴을 통한 반복적 수렴 기법 [7][8]

본 분류에 해당하는 기법에서는 대상 물체의 3차원 형상이 수학적인 곡면 모델로서 표현 되며, 이는 다각형 메쉬(polygonal mesh)와 같은 명시적(explicit)이거나 레벨 셋(level set)과 같은 내재적인(implicit) 표현 방법을 포함한다. 이러한 곡면 모델을 통한 대상 물체의 3차원 형상 복원은 3차원 모델과 주어진 입력 영상과의 정합 오차를 의미하는 에너지 함수(energy function)의 최적화를 통해 이루어지는데, 이러한 에너지 함수는 보통 곡면 상의 3차원 점에서 각각의 입력 영상으로 투영된 점들의 밝기나 컬러 정보의 일관성을 나타내는 영상-일관성 제약조건으로 표현된다. 3차원 곡면 모델을 변수로 하는 에너지 함수의 최적화에는 고전적인 방법인 변분 방법(variational method)이 주로 사용되며, 이를 통해 곡면 모델의 시간에 따른 전개 과정을 묘사하는 편미분 방정식(PDE)을 유도할 수 있다. 이러한 곡면 전개 방정식의 해를 구하기 위해서는 초기 곡면 정보가 필요한데, 이를 위해 일반적으로 대상 물체를 포함하는 경계 상자(bounding box) 혹은 대상 물체의 윤곽선(silhouette) 정보와 일치하는 최대 형상인 비주얼 헐(visual hull)이 사용된다. 초기화된 곡면 모델은 곡면 전개 편미분 방정식(surface evolution PDE)을 통해 시간적으로 전개되며, 최종적으로 수렴된 곡면의 형태로 대상 물체의 3차원 형상이 복원된다.

이러한 변형 가능한 곡면(deformable surface) 기반의 3차원 형상 복원 방법은 각 반복 단계(iteration step)마다 주어진 입력 영상들에 대한 가시성(visibility) 정보를 쉽게 추정할 수 있기 때문에 영상-일관성 제약조건을 통한 에너지의 계산에 있어서 가리어짐(occlusion)으로 인한 오차를 줄일 수 있다는 장점을 가진다. 반면에 초기 곡면이 전개 도중 에너지 함수의 국부적 최소점(local minima)에 빠질 수 있으며, 에너지 함수의 특성 상 더 적은 면적을 가지는 곡면이 선호되어 높은 곡률을 가지는 영역이 과도하게 평탄화되는 소위 최소 곡면 편향(minimal surface bias)과 같은 단점이 있다.

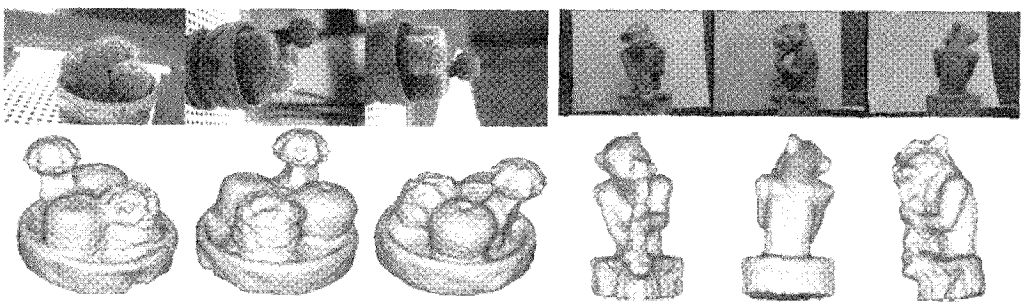
3.2.1 레벨 셋 (level set) 및 전역적 영상기반 정합 점수 (global image-based matching score) 기반의 다중뷰 스테레오 [7]

본 기법에서는 레벨 셋(level set) 기반의 다중뷰 스테레오를 예측 오차(prediction error)의 최적화에 기반한 영상 정합(image registration) 문제의 형태로 정식화하는 방법이 제안된다. 일반적으로 다중뷰 스테레오에서는 잡음(noise) 및 여러 현실적인 영상 취득 조건에 강인한 정합 함수를 사용하기 위해서 특정한 크기의 윈도우를 사용한다. 그런데 이러한 윈도우는 위치와 방향을 함께 고려해야 하므로, 많은 수의 기법들이 윈도우가 카메라의 초점 면(focal plane)에 수평을 이루고 있다는 fronto parallel 가정을 사용하여 정합 함수 계산을 간략화 한다. 하지만 이 가정은 대상 물체의 실제 기하학적 성질을 제대로 반영하고 있지 못하므로, 소위 투영 왜곡(projective distortion)으로 인한 3차원 복원 결과의 오차를 야기한다. 이를 해결하기 위해 윈도우를 사용한 정합 함수를 계산할 때에 현재 곡면의 접평면(tangent plane)을 고려하는 방법이 제안되었으나, 곡면의 1차 근사를 사용함에도 불구하고 결과적인 곡면 전개 방정식은 매우 복잡한 형태를 가지게 되어, 방정식의 풀이 과정이 다루기 힘들고, 많은 계산 시간을 요구하며, 불안정하다는 단점을 가지고 있다.

이러한 점을 해결하기 위해 본 기법에서는 영상 정합 문제에서 사용되었던 예측 오차에 기반한 에너지 함수가 제안된다. 이는 기준 영상으로부터 현재 곡면 모델에 기반하여 목표 영상을 예측하고, 이렇게 예측된 목표 영상과 실제 목표 영상을 비교하여 계산된 오차로 정의된다. 현재 곡면 모델을 참고하여, 기준 영상으로부터 목표 영상을 생성하는 것은 대부분의 그래픽 하드웨어가 지원하는 텍스처 매핑(texture mapping)을 통해 쉽고, 효율적으로 수행될 수 있다. 구체적으로는 OpenGL의 projective texture mapping 함수를 사용하여 기준 영상의 다른 시점에 대한 재투영 영상(reprojection image)을 계산하고, shadow mapping 함수를 통해 가리어짐 영역을 제거하는 방법이 언급되어 있다. 이렇게 예측된 목표 영상과 실제 목표 영상을 NCC 등의 여러 가지 정합 함수를 통해 비교하여 최종 에너지 함수를 생성한다. 생성된 에너지 함수는 현재 곡면의 위치(position)에만 의존하므로 그 최적화가 매우 간단하다. (그림 4)는 이러한 광역적 영상-기반 정합 함수에 기반한 다중뷰 스테레오 알고리즘의 복원 결과를 보여준다.

3.2.2 변형 가능한 메쉬 (deformable mesh) 기반의 다중뷰 스테레오 [8]

본 기법의 주된 목적은 높은 해상도를 가지는 많은 영상으로부터 대규모 장면(large-scale



(그림 4) 레벨 셋과 전역적 영상기반 정합 점수 기법의 실험 결과 [7]

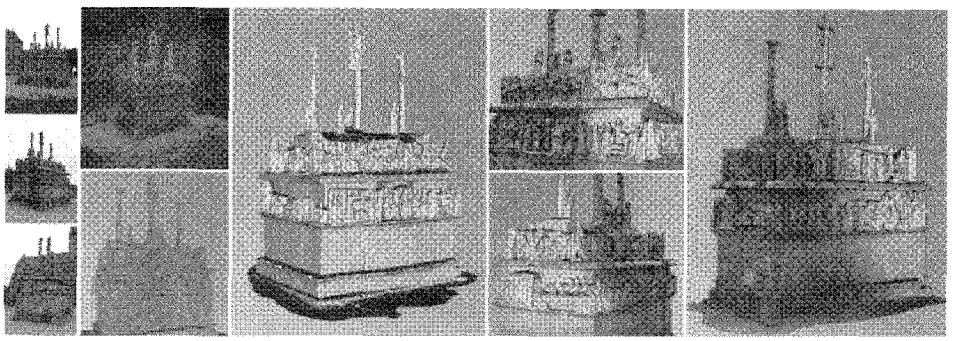
scene)에 대한 효율적이고, 정확한 다중뷰 스테레오 알고리즘을 제안하는 것이다. 복셀(voxel) 혹은 레벨 셋(level set) 표현을 포함하는 다양한 기존 방법들은 계산량 및 메모리의 한계로 인해 대규모 장면을 표현하는 데에 문제가 있으므로, 본 기법에서는 변형 가능한 메쉬를 이용하는 방법을 제안하였다. 이는 다음의 두 가지 주요한 단계로 구성된다: (1) 밀집된 3차원 점의 집합(dense 3D point cloud)으로부터 그래프 컷 최적화에 기반하여 가시성을 고려한 메쉬 모델 생성; (2) 생성된 메쉬에 대한 주어진 영상에 일관된 변분 최적화(variational optimization).

첫 번째 단계에서는, 주어진 영상으로부터 추출된 많은 수의 SIFT, difference of Gaussian (DOG), Harris 특징점들과 NCC의 합으로 표현되는 함수를 사용한 점 정합(point matching)을 통해 밀집된 3차원 점의 집합을 생성한다. 이때 정합의 강인함을 높이기 위해 NCC 함수를 계산하기 위한 윈도우의 크기 및 방향을 다양하게 고려하며, 또한 이러한 특징점 추출 및 NCC 계산의 모든 과정은 빠른 수행속도를 위해 그래픽 프로세서(GPU)를 통해 구현된다. 추출된 많은 수의 특징점들은 Delaunay 삼각화를 통해서 Delaunay tetrahedra로 변환되며, 각 카메라에 대한 가시성(visibility)을 고려한 에너지 함수의 광역적 최적화를 통해 초기 메쉬 모델이 생성된다. 이러한 에너지 함수의 최적화에는 그래프 컷 기법이 사용된다.

앞서 얻어진 초기 메쉬 모델은 세밀한 부분에서의 복원 정확도가 떨어지는데, 이를 개선하기 위해 두 번째 단계에서는 초기 메쉬 모델에 입력 영상으로부터 정의된 에너지 함수의 경사 하강 방정식(gradient descent equation)을 적용하는 변분 다중뷰 스테레오(variational multi-view stereo) 기법을 사용한다. 일반적으로 에너지 함수는 입력 영상에 대한 영상-일관성 및 곡면 평탄화 항(smoothness term)으로 구성된다. 이러한 국부적 최적화(local optimization)를 통해 초기 메쉬 모델이 국부적 최소점에 빠지지 않고, 올바른 곡면 모델로 수렴하기 위해서는 초기 메쉬 모델이 충분히 대상 3차원 형상에 근접해야 하는데, 본 기법에서는 앞선 단계의 그래프 컷을 통한 일관된 가시성 메쉬 모델 생성이 그러한 역할을 잘 수행하고 있다고 주장하고 있다.

(그림 5)는 전체적인 알고리즘의 수행 단계를 보여주고 있는데, 주어진 고해상도의 입력 영상으로부터 얻어진 많은 수의 3차원 점의 집합과, 이로부터 얻어진 초기 메쉬 모델, 그리고 메쉬의 변형을 통해 개선된 최종 메쉬 모델 및 다양한 시점에서의 렌더링 결과를 확인할 수 있다.

3.3 Patch 또는 surfel을 이용한 3차원 표면 요소 모델링 및 영역 확장 기법 [9][10]



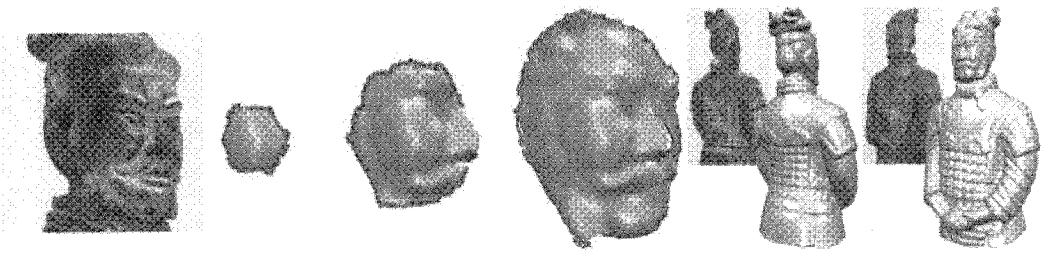
(그림 5) 변형 가능한 메쉬 기반의 다중 스테레오 기법의 실험 결과 [8]

본 분류의 기법에서는 대상 물체의 곡면을 작은 평면 패치(planar patch)들의 집합으로 표현한다. 이러한 독립적인 패치들의 집합에서는 패치들 사이의 연결 정보를 고려할 필요가 없으므로 매우 간단하고, 효율적으로 대상 곡면을 표현할 수 있으며, 또한 곡면의 시각화(visualization)에 있어서도 추가적인 다면체 메쉬 모델로의 변환 없이 점 기반 렌더링(point-based rendering)과 같은 기법을 직접적으로 적용할 수 있다는 장점을 가진다. 반면에 각 패치들을 독립적으로 고려하므로 잡음(noise)에 강인한 3차원 형상 복원을 위해 필요한 곡면의 평탄화 제약조건(smoothness constraint)을 적용하기 어려우며, 패치 마다 각각의 영상에 대한 가시성(visibility)을 판별하기가 쉽지 않고, 또한 텍스처(texture) 정보가 부족한 영역에서는 패치가 제대로 생성되지 않아서 구멍(hole) 영역에서의 추가적인 패치 생성이 필요하다는 단점이 있다.

3.3.1 Surfel 기반의 다중뷰 스테레오 기법 [9]

본 기법에서는 방향성을 가진 점(oriented particles)을 기반으로 한 곡면 표현 방법을 활용하는 다중뷰 스테레오 기법이 제안된다. 구체적으로 각각의 곡면 요소들은 중심 점, 법선 벡터, 그리고 반지름으로 정의되는 평면 디스크(planar disk), 즉 surfel의 형태를 가진다. 이러한 디스크 기반의 곡면 표현 방법은 다음과 같은 장점을 가진다. 먼저 3차원 공간에서의 샘플링 밀도와 디스크의 크기가 임의로 그리고 독립적으로 조절될 수 있어서 매우 융통성 있는(flexible) 다중해상도 3차원 형상 복원을 가능하게 한다. 또한 제안하는 방법은 경계 상자(bounding box)나 비주얼 헐(visual hull) 등의 초기화 정보를 필요로 하지 않는다.

제안된 방법은 다음과 같은 두 단계로 이루어진다. 첫 번째 단계에서는 자동으로 결정된 입력 영상 쌍으로부터 영상-기반 호모그래피(homography) 정합을 통해서 시드 디스크(seed disk)를 계산한다. 계산된 시드 디스크는 임의의 개수의 입력 영상을 고려하는 평면 근사(plane fitting) 알고리즘을 통해 올바르게 수정된다. 두 번째 단계에서는 앞서 복원된 시드 디스크 정보를 greedy growing 방법을 사용하여 주변 영역으로 확장한다. 새로운 디스크들은 기존의 디스크 영역의 경계에서 생성되는데, 이러한 확장 알고리즘은 대상 장면의 모든 가시적인 부분들이 디스크로 모두 채워질 때까지 동작하여, 최종적으로 정확한 법선 정보를 포함하는 3차원 디스크의 밀집된 집합을 생성하게 된다. 복원된 디스크 집합은 가상 시점 생성(virtual view synthesis)과 같은 목적으로 직접적으로 적용될 수 있으며, 복원된 모델의 다양한 응용을 위해 Poisson surface reconstruction[11] 기법을 통하여 삼각형 메쉬 모델로 변환될 수도 있다. (그림 6)은 제안된 알고리즘의 surface growing 과정과 이를 통해 생성된 3차원 메쉬 모델을 보여주고 있다.



(그림 6) Surfel 및 곡면 확장 기법의 사례 및 이를 통해 복원된 3차원 형상 [9]

(그림 7) PMVC 기법의 예제 [10]. 좌로부터 입력 영상, 검출된 특징점, 초기 정합을 통해 복원된 3차원 패치, 확장 및 필터링을 통해 생성된 최종 패치, 그리고 변환된 메쉬 모델

3.3.2 Patch 기반의 다중뷰 스테레오 기법 [10]

본 기법은 단일한 물체의 복원뿐만이 아니라 여러 물체들 및 다양한 배경 물체들로 이루어져 가리어짐 관계가 복잡하고, 사전에 경계 상자(bounding box)를 추정하기 어려운 복잡한 장면에 대한 3차원 형상 복원을 그 목표로 한다. 이를 달성하기 위해서는 기존의 복셀 기반이나 메쉬 기반의 곡면 모델보다 좀 더 융통성 있는 표현 방법이 필요하며, 따라서 본 기법에서는 작은 사각형 패치들의 집합으로 대상 물체를 복원하는 패치 기반의 다중뷰 스테레오 기법(patch-based multiview stereo)을 제안한다.

제안하는 방법은 정합, 확장, 필터링의 세 단계로 이루어진다. 먼저 정합 단계에서는 Harris 및 DOG 연산자로부터 얻어진 여러 입력 영상의 특징점들이 서로 정합되어 조밀하지 않은 3D points가 생성된다. 확장 단계에서는 초기 정합된 점들을 인접한 픽셀들로 확장해서 더욱 밀집한 3차원 패치 집합을 생성 하며, 필터링 단계에서는 가시성을 고려하여 올바르지 않은 패치들을 제거한다. 제안하는 방법은 Lhuillier와 Quan[12]에 의해서 제안되었던 방법의 greedy expansion 단계를 확장 및 필터링 단계로 개선하여 복잡한 곡면을 좀 더 잘 다루고, outlier들을 좀 더 효율적으로 제거한다. 제안하는 방법에서는 정합 단계로부터 얻어진 초기 3차원 점들로부터 확장 및 필터링 단계가 몇 번 반복이 되는데, 이를 통해 텍스처가 부족한 영역이나 가리어진 영역을 제외하고 대상 물체를 조밀하게 포함하는 방향성을 가진 사각형 패치들의 집합이 생성된다. 생성된 패치 집합은 추가적인 Poisson surface reconstruction 기법을 통하여 삼각형 메쉬 모델로 변환될 수 있다. 제안된 방법의 전체적인 절차는 그림 7에 예제를 통해 나타내었다.

5. 결 론

본 기고에서는 최근 각광받고 있는 다중뷰 스테레오 기법의 개요와 최신 알고리즘을 소개하고 비교 분석하였다. 지난 10년간 활발히 연구된 다중뷰 스테레오 기법의 완성도는 이제 성숙기에 접어들어 실제 산업 분야에서 응용할 수 있는 수준에 이르렀다고 볼 수 있다. 현재는 기존의 기법을 확장하여 인터넷상의 수많은 영상을 이용한 도시 모델링이나 건물, 얼굴, 역사적 유적 등의 특정 객체에 최적화된 기법들이 속속 발표되고 있다. 또한 많은 계산량과 긴 수행시간을 GPU를 이용한 대용량 병렬처리를 통해 극복하는 연구도 활발히 진행되고 있다. 독자들은 본 기고의 내용을 시작점으로 보다 쉽게 다중뷰 스테레오를 이용한 3차원 형상 복원 기법의 현황과 전망을 파악할 수 있을 것이라고 기대한다.

감사의 글

본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (NIPA-2010-(C1090-1011-0003)).

이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(2010-0015486).

참고문헌

[1] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Second Edition, Cambridge University Press, 2004.

[2] T. Moons, L. Van Gool, and M. Vergauwen, "3D reconstruction from multiple images, Part 1: Principles," Foundations and Trends® in Computer Graphics and Vision, vol. 4, no. 4 pp. 287 - 404, 2008.

[3] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, Multi-view stereo evaluation, http://vision.middlebury.edu/mview/.

[4] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," Proc. IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 519-526, June 2006.

[5] G. Vogiatzis, C. Hernández, P. Torr and R. Cipolla, "Multi-view stereo via volumetric graph-cuts and occlusion robust photo-consistency," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 12, pp. 2241-2246, December 2007.

[6] S. Sinha, P. Mordohai, and M. Pollefeys, "Multi-view stereo via graph cuts on the dual of an adaptive tetrahedral mesh," Proc. IEEE International Conference on Computer Vision, pp. 1-8, October 2007.

[7] J.-P. Pons, R. Keriven, and O.D. Faugeras, "Multi-View Stereo Reconstruction and Scene Flow Estimation with a Global Image-Based Matching Score," International Journal of Computer Vision, vol. 72, no. 2, pp. 179-193, 2007.

[8] H. H. Vu, R. Keriven, P. Labatut and J.-P. Pons, "Towards high-resolution large-scale multi-view stereo," Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 1430-1437, June 2009.

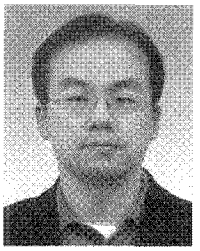
[9] M. Habbecke and L. Kobbelt, "A surface-growing approach to multi-view stereo reconstruction," Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2007.

[10] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32 no. 8, pp. 1362-1376, August 2010.

[11] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," Proc. Symp. Geometry Processing, 2006.

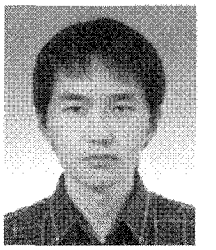
[12] M. Lhuillier and L. Quan, "A quasi-dense approach to surface reconstruction from uncalibrated images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27 no. 3, pp. 418-433, March 2005.

저자약력



박 인 규

1995년 2월 서울대학교 제어계측공학과(학사)
1997년 2월 서울대학교 제어계측공학과(석사)
2001년 8월 서울대학교 전기컴퓨터공학부(박사)
2001년 9월~2004년 3월 삼성종합기술원 멀티미디어랩 전문연구원
2007년 1월~2008년 2월 미국 Mitsubishi Electric Research Laboratories (MERL) 방문연구원
2004년 3월~현재 인하대학교 정보통신공학부 조교수
관심분야 : 컴퓨터 그래픽스 및 비전 (영상기반 3차원 형상 모델링 및 렌더링, computational photography), GPGPU
이 메 일 : pik@inha.ac.kr



장 주 용

2001년 2월 서울대학교 전기공학부(학사)
2008년 2월 서울대학교 전기컴퓨터공학부(박사)
2008년 2월~2009년 1월 미국 Mitsubishi Electric Research Laboratories (MERL) Postdoctoral Researcher
2009년 4월~현재 삼성전자 DMC연구소 멀티미디어팀 책임연구원
관심분야 : 컴퓨터 비전 (영상기반 3차원 형상 복원 및 3차원 자세 추정), 3차원 영상처리 및 디스플레이, GPGPU
이 메 일 : juyong.chang@gmail.com