

명시적인 3차원 물체 재구성 과정이 생략된 실시간 3차원 특징값 추출 방법*

홍광진[○] 이철한 정기철 오경수

송실대학교 대학원 미디어학과

{hongmsz, dashans, kcjung, oks}@ssu.ac.kr

Real-time 3D Feature Extraction without Explicit 3D Object Reconstruction

Kwangjin Hong[○] Chulhan Lee Keechul Jung Kyoungsu Oh

Department of Media, Graduate School of Soongsil University

1. 서론

인식 알고리즘은 상호작용이 가능한 컴퓨팅 환경에서 매우 중요한 요소이고, 특히 인식 정확도와 속도가 중요하게 다루어지는 문제이다. 따라서 대부분의 인식 알고리즘에 대한 연구는 이 문제를 중심으로 이루어지고 있다. 인식에 사용되는 특징값을 추출하는 방법은 2차원의 카메라 입력 영상을 이용하는 방법(2차원 특징값)과 재구성된 3차원 물체를 이용하는 방법(3차원 특징값)으로 나눌 수 있다. 2차원 특징값을 이용하는 인식 알고리즘은 인식 결과 도출 시간이 짧지만, 추출된 2차원 특징값으로 3차원 공간 안의 인식 대상을 표현하는데 한계가 있다. 반면에, 3차원 특징값을 이용하는 인식 알고리즘은 2차원 특징값에 비해 정확한 인식이 가능하지만, 특징값 추출을 위해 3차원 물체를 재구성하는 선행 과정이 필요하기 때문에 결과 도출 시간이 길다. 본 논문에서 우리는 명시적인 3차원 물체 재구성 과정을 생략하여 실시간 3차원 특징값 추출이 가능한 방법을 제안한다. 제안하는 방법은 기존의 특징값 추출 방법과 달리 3차원 물체 재구성 단계를 특징값 추출 단계에 통합시킴으로써 실시간으로 결과를 생성할 수 있다.

2. 본론

본 논문에서 우리는 명시적인 3차원 물체 재구성 과정을 생략하여 실시간 3차원 특징값 추출이 가능한 방법을 제안한다. 제안하는 방법은 기존의 특징값 추출 방법과 달리 3차원 물체 재구성 단계를 특징값 추출 단계에 통합시킴으로써 실시간으로 결과를 생성할 수 있다. 또한, 제안하는 방법은 각각 다른 정보를 가지는 3가지 종류의 특징값(프로젝션 맵)을 생성한다(그림 1).

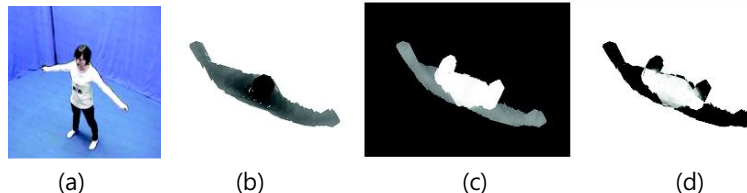


그림 1. 프로젝트 맵: (a) 입력 영상, (b)가장 가까운 경계면, (c)가장 먼 경계면, (d) 두께 맵.

프로젝션 맵은 인식 대상을 렌더링하여 생성할 수 있다. 이를 위해 비주얼 헐이 사용되는데, 비주얼 헐은 실루엣을 이용하여 3차원 물체의 형태를 재구성하는 방법을 통해 근사화된 기하학적 결과물이다. Li 등이 제안한 Hardware-accelerated Visual Hull(HAVH) 렌더링 기법[1]은 비주얼 헐을 재구성하고 렌더링하는 과정이 일반적인 그래픽 장치(GPU)를 사용하여 가속화될 수 있음을 보여준다. 또한, 우리는 비주얼 헐 렌더링 결과물로부터 특징값을 추출하기 때문에, 조명이나 텍스처 등과 같은 세부적인 정보를 필요로 하지 않는다. 따라서, 우리는 기존 알고리즘에서 불필요한 과정을 제거한 수정된 HAVH 알고리즘을 사용한다. 제안하는 방법은 입력 영상을 이용하여 실루엣 콘을 생성하는 단계와 프로젝트 맵이 생성되는 비주얼 헐 렌더링 단계의 두 부분으로 이루어진다. 실루엣 콘 생성 단계에서는 다시점 카메라 입력 영상에서 배경과 물체를 분리하고, 물체에 대한 실루엣 영역을 추출한 뒤, 카메라 각각에 대해 추출된 실루엣 영역과 카메라 위치를 연결하여 실루엣 콘을 생성한다. 또한, 비주얼 헐 렌더링 단계에서는 이전 단계에서 생성한 실루엣 콘에 실루엣 영상을 투영함으로써 교차 영역의 기하정보를 계산하지 않고 렌더링할 수 있다. 우리는 카메라 촬상면으로부터 3차원 물체의 가장 가까운 경계면까지의 거리와 가장 먼 경계면까지

* 이 논문은 한국과학재단(KOSEF)의 특정기초연구 No.(R01-2006-000-11214-0)의 지원에 의하여 연구되었음.

의 거리, 3차원 물체의 두께와 같은 3차원 정보를 포함하는 프로젝션 맵을 생성한다. 물체의 가장 가까운 경계면과 가장 먼 경계면 프로젝션 맵은 카메라 촬상면으로부터 물체 앞면을 구성하는 픽셀까지의 거리와 물체 뒷면을 구성하는 픽셀까지의 거리를 구함으로써 얻을 수 있다. 또한 물체의 두께 정보를 포함하는 프로젝션 맵(두께 맵)의 경우, 앞면을 이루는 픽셀까지의 거리와 뒷면을 이루는 픽셀까지의 거리의 차를 계산함으로써 구할 수 있다. 이 과정을 식으로 정리하면 다음과 같다.

$$P_f[i,j] = \begin{cases} \min_{z \leq 300} (z_p) & \text{if } \vec{v}[i,j,z] \cdot \vec{p}[i,j,z] > 0 \\ 0 & \text{otherwise} \end{cases} \quad P_r[i,j] = \begin{cases} \max_{z \leq 300} (z_p) & \text{if } \vec{v}[i,j,z] \cdot \vec{p}[i,j,z] < 0 \\ 0 & \text{otherwise} \end{cases} \quad P_d[i,j] = P_r[i,j] - P_f[i,j]$$

식에서 $P_f[i,j]$ 는 가장 가까운 경계면을, $P_r[i,j]$ 는 가장 먼 경계면을, $P_d[i,j]$ 는 두께를 나타낸다. $\vec{v}[i,j,z]$ 은 카메라 촬상면 상의 한 점(i, j)에서 수직인 벡터이고, $\vec{p}[i,j,z]$ 는 촬상면 (i, j)에 대응하는 3차원 물체를 구성하는 한 점에 대한 노말 벡터이며, z_p 는 $\vec{p}[i,j,z]$ 의 z 성분의 크기를 의미한다. 각각의 프로젝션 맵은 특징점을 추출하고자 하는 물체를 미리 정해진 시점에서 렌더링함으로써 획득할 수 있다. 이때, 원근 투영으로 인한 왜곡을 없애기 위하여 직교 투영을 이용한다. 따라서 기준 평면 위의 모든 점은 수직 방향의 시점을 가지게 되고 기준면으로부터의 깊이(z) 값만을 이용하여 프로젝션 맵을 간단히 획득할 수 있다. 깊이 정보는 물체가 렌더링 될 때 그래픽스 하드웨어에서 자동으로 계산된다. 생성된 픽셀들의 깊이 값은 깊이 버퍼에 저장되며, 이 값을 텍스처에 저장하여 프로젝션 맵을 생성한다. 표 1은 제안된 방법이 기존의 방법들과 달리 실시간 동작 인식 환경에 적용할 수 있음을 보이기 위한 비교 결과이다. 실험을 위해, 하나의 좌표계로 정확하게 캘리브레이션 되어있는 8개의 카메라로 촬영된 640x480 크기의 입력 영상을 사용하였다.

표 1. 제안한 방법과 기존의 3차원 특징값을 추출하는 방법의 각 세부 단계별 수행 시간 비교.

특징값 추출 방법	과정 별 수행 시간(ms)		
	실루엣 콘 생성	비주얼 헐 생성(GPU기반)	특징값 추출
세션화 기반 Skeleton 추출	21	370	10^7
3차원 bin-distribution 생성	21	370	10
제안된 방법	21	2~3	

실험에서 3차원 모델은 8개의 다시점 카메라를 통해 입력받은 영상을 이용하여, 300x300x300 크기의 voxel 공간 안에 생성되었으며, 비주얼 헐을 생성하는 단계는 모두 GPU를 이용한다. 표에서 보는 것과 같이 제안하는 방법은 기존 방법과 달리 3차원 모델 생성 과정에서 특징값 추출이 가능하기 때문에, 특징값 추출 시간이 매우 빠르고, 입력 영상에서 실루엣을 추출하는 시간을 포함하여 초당 13~14개의 입력 영상 세트를 처리할 수 있다.

3. 결론

본 논문에서 우리는 명시적인 3차원 물체 재구성 과정을 생략하여 실시간으로 3차원 특징값 추출이 가능한 방법을 제안한다. 제안된 방법은 3차원 물체를 생성하는 대신 HAVH 렌더링 기법을 이용하여 3차원 모델에 대한 3종류(가장 가까운 경계면, 가장 먼 경계면, 두께)의 프로젝션 맵을 생성한다. 특히 우리는 불필요한 과정을 제거한 수정된 HAVH 알고리즘을 사용하기 때문에, 입력 영상 내의 물체에 대한 3차원 정보를 빠르게 계산할 수 있고, 따라서, 실시간 동작 인식 시스템에 적용이 가능하다. 그러나 제안된 방법은 현재 GPU 하드웨어의 제약 때문에 16개 이상의 카메라를 동시에 사용할 수 없고, CPU와 GPU를 동시에 사용할 경우 데이터 전송에 지연이 발생한다. 또한, 제안하는 방법은 카메라 촬상면에서 가장 가까운 경계면과 가장 먼 경계면 사이의 거리만을 계산하기 때문에, 물체 내의 겹쳐진 영역 사이의 거리를 계산하지 못한다. 따라서, 현재 우리는 제안한 방법의 성능을 향상시키기 위해, 메모리 전송 시간 감소와 정확도 향상에 대한 연구를 진행하고 있다.

참고 문헌

[1] M. Li, M. Magnor, H. Seidel, "Hardware-accelerated visual hull reconstruction and rendering," Graphics Interface 2003, pp. 65-71 (2003).