

다수 카메라를 이용한 3차원 복원

김영수*, 박성찬**, 정 홍***
 포항공과대학교 전자전기공학과

3D Reconstruction Using Multi-Camera

YoungSu Kim*, SungChan Park**, Hong Jeong***

Pohang University of Science and technology Electronic and Electrical Engineering

Pohang, Kyungbuk, 790-784, South Korea

E-mail : *zamgyul@postech.ac.kr, **mrzoo@postech.ac.kr, ***hjeong@postech.ac.kr

Abstract

In this paper, we present a 3D reconstruction method using multi-camera. This method is very compact algorithm so that can be implemented easily on small hardware architecture. By using multi-camera, it gives exacter result than existing method and we propose accurate index for each matching nodes to use multi-camera.

I. 서론

인간의 시각시스템은 두개의 눈을 포함하는데 이 두개의 눈을 통해 우리는 사물의 멀고 가까움을 알 수 있다. 한장의 영상은 3차원 공간을 2차원 공간으로 projection한 것으로, 이 과정에서 depth 정보를 잃어버리게 되는 손실이 생긴다. 사물에 대해 서로 다른 위치에서 얻어진 영상이 두장이상 있을때 3차원 공간의 복원이 가능해진다. 이와 같은 3차원 복원은 일반적으로 계산시간이 오래걸리기 때문에 실시간 반응을 유도하는 시스템에는 적용이 쉽지않다. 따라서 고속으로 이러한 작업을 하기 위하여 하드웨어구현이 가능한 간단한 형태의 높은 매칭 효율을 가지는 알고리즘이 필요하다. 본 논문은 기존의 systolic array를 사용하는 방식[1]을 다수의 카메라에 적용하여 3차원 공간을 고속으로 정확하게 복원하는 알고리즘에 대하여 기술하였다.

II. 본론

그림 1과 같이 좌우 두 시각에서 $P(X, Y, Z)$ 를 바라볼 때 depth Z 는 다음과 같이 구할 수 있다.

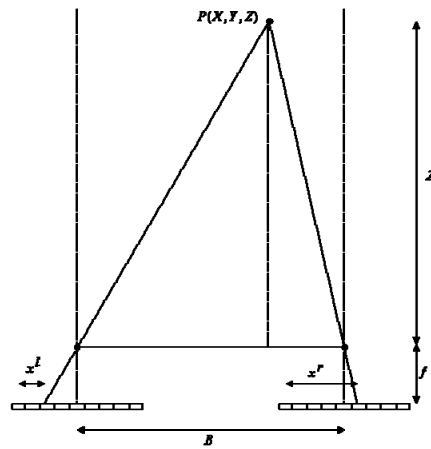


그림 1. 평행광축 시각 기하구조

$$x^r - x^l : f = B : Z, \quad Z = \frac{f \cdot B}{x^r - x^l} = \frac{f \cdot B}{d}$$

여기서 d 는 disparity이고 focal length f 와 base length B 를 알고있다면 disparity로 depth를 대신하여 사물의 깊이를 나타낼 수 있다. [2,3].

본 논문의 알고리즘은 rectified된 N개 디지털영상에서 동일 y 좌표의 한 개 라인을 각각 뽑아내어 그림 2와 같이 정합점의 위치를 유한하게 정의한다.

이 유한하게 정의된 각 정합점에서 정합되는 각

카메라의 픽셀 위치는 다음과 같이 구할 수 있다. N 개의 카메라를 일정한 간격 $\frac{L}{N-1}$ 으로 배치시킬 때, 1번 카메라와 N 번 카메라 사이의 base length는 L 이다.

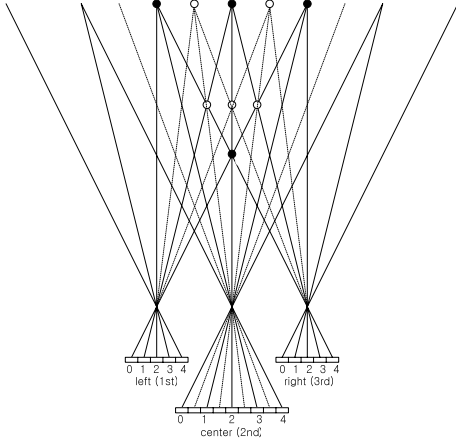


그림 2. Three 카메라에서의 정합점

이때 1번 카메라와 N 번 카메라의 정합점을 기준으로 하여 disparity level d 를 정하고 disparity의 인덱스와 동일한 중앙 카메라의 픽셀 인덱스를 x_c 라 칭한다. 이렇게 하여 d 레벨, x_c 위치에 맞는 i 번째 카메라의 픽셀 인덱스 x_i 는 다음과 같이 구한다.

$$z = \frac{f * L}{x_N - x_1} = \frac{f * L}{d} = \frac{f * L}{x_c - x_i} \left(\frac{N+1}{2} - i \right)$$

$$x_c - x_i = \frac{1}{N-1} \left(\frac{N+1}{2} - i \right) d = \frac{N+1-2i}{2(N-1)} d$$

$$x_i = x_c - \frac{N+1-2i}{2(N-1)} d$$

이와 같이 각 정합점에서 N 카메라의 픽셀 인덱스를 얻은 후 interpolation을 통하여 인덱스에 맞는 픽셀 데이터를 얻고 다음과 같이 각 정합점의 정합신뢰도인 매칭코스트(m)를 계산한다.

$$m = \sum_{i=1}^{N-1} \sum_{h=i+1}^N \left| x_i \left[\frac{t}{2} + \frac{N+1-2i}{2(N-1)} j \right] - x_h \left[\frac{t}{2} + \frac{N+1-2h}{2(N-1)} j \right] \right|$$

그림 3과 같이 각 정합점사이의 이동 가능경로를 $\{-1,0,1\}$ 의 disparity level범위로 제한시키고 viterbi 알고리즘으로 최소 누적 매칭코스트 path를 찾으면 그것이 해당 영상라인의 3차원 복원 결과가 된다.

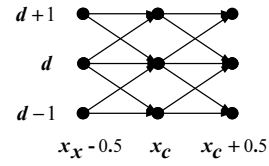


그림 3. viterbi search path

III. 실험 결과

본 알고리즘을 이용한 시뮬레이션 결과 다음과 같은 결과를 얻었다.

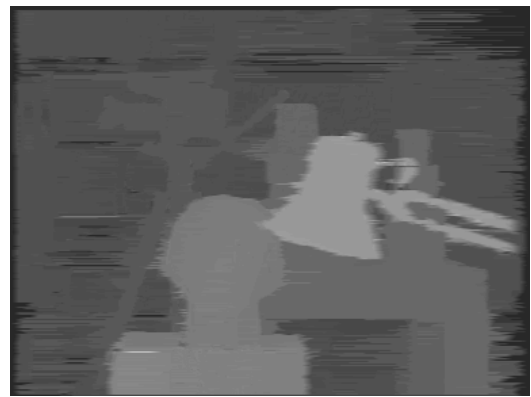


그림 4. 1번, 2번, 3번 영상, 결과영상 (좌측상단부터)

IV. 결론 및 향후 연구 방향

이 알고리즘은 고속 하드웨어 구현에 적합하도록 개발되었으며 다수의 카메라를 사용하여 정합점에 대한 신뢰도가 향상되어 더욱 강인한 결과를 얻을 수 있게 되었다.

참고문헌

- [1] Hong Jeong, Yuns Oh, "A parallel real time implementation of stereo matching", Parallel and Distributed Processing Symposium., Proceedings 15th International 23-27 april 2001 page: 6pp.
- [2] Umesh R. Dhond and J. K. Aggarawal. "Structure from stereo a review" IEEE Transactions on Systems, Man, and Cybernetics, 19(6): pp 553-572, nov/dec 1989.
- [3] Rafael C. Gonzalez and Paul Wintz., "Digital Image Processing." Addison-Wesley, Reading, MA, USA, 1987.