

# 임의 시점 영상 합성을 위한 깊이 정보 추출 기법

박 남 준<sup>\*\*\*</sup>, 이 제 호<sup>\*</sup>, 권 용 무<sup>\*</sup>, 박 상 회<sup>\*\*</sup>

## Depth Information Extraction Technique for Arbitrary Viewpoint Image Synthesis

Nam-Jun Park<sup>\*\*\*</sup>, Je-Ho Lee<sup>\*</sup>, Yong-Moo Kwon<sup>\*</sup>, Sang-Hui Park<sup>\*\*</sup>

\* 한국과학기술연구원, \*\* 연세대학교 전기공학과  
e-mail : pnj@cherry.kist.re.kr Fax : 02) 958-5769

### 요 약

본 논문에서는 임의 시점 영상 합성을 위한 깊이 정보 추출에 관한 방법을 제안한다. 깊이 정보의 추출을 위한 방법으로서 기존의 MBS (Multiple-Baseline Stereo) 방법의 깊이 맵의 경계선 연장(boundary overreach) 문제를 감소시키며 처리 시간을 개선하는 방법으로서 계층적 방법인 MR-MBS(Multi-Resolution MBS) 방법을 제시한다. 또한 MBS 방법에서 고려하지 않고 있는 폐색 영역에 대한 적절한 처리 방법을 제시한다.

### I. 서 론

최근 대두되고 있는 3D-TV는 기존의 2D-TV에 현실감과 실재감을 높여 주어 사물이 실제 존재하는 것과 같은 효과를 주고자 한다. 이를 구현하기 위해서 핵심적으로 필요한 기술이 다시점 영상 합성 기술이다[1].

카메라로부터 획득한 영상을 사용하여 원하고자 하는 위치에 있는 영상을 합성해 내는 기술은 여러 가지가 제시되어 왔으나 크게 컴퓨터 그래픽스 쪽에서의 접근과 컴퓨터 비전 쪽에서의 접근으로 나누어 볼 수 있다. 먼저 컴퓨터 그래픽스 쪽의 접근으로는 대표적으로 모핑이론을 적용한 View Interpolation 방식[2], Plenoptic 모델링을 통한 방식[3] 등이 제안되고 있다.

한편 컴퓨터 비전 쪽에서의 접근은 깊이 정보를 이용한 방식으로서 일반적으로 이를 구하기 위해 서로 다른 두 영상간의 대응점을 찾는 스테레오 정합을 사용한다. 즉 정합을 통해 추출한 대응점의 위치 좌표로부터 카메라로부터 물체까지의 깊이 정보를 추출한다. 이렇게 구한 깊이 정보를 사용하여 임의 위치에 있는 영상에 대한 대응점의 위치 좌표를 계산한다. 전자의 접근 방법에 후자의 방법을 접목시키는 연구도 많이 제시되고 있으나 기본적으로는 후자에 근거한 방법들이 대부분을 이루고 있다.

일반적인 깊이 정보의 추출과는 달리 영상 합성에 필

요한 깊이 정보의 추출 방식은 합성하고자 하는 영상의 모든 화소에 대한 깊이 정보를 구하여야 하므로 정확하고 조밀한 깊이 정보의 추출이 필요하다. 또한 자연스러운 동영상 지향하고 있으므로 처리 속도의 비중도 중요하다 할 수 있다. 그리고 영상 획득을 위한 카메라의 배치 문제에 있어서도 자유로운 카메라의 배치를 갖는 알고리즘을 필요로 한다.

따라서 본 논문에서는 보다 정확한 깊이 정보를 추출하며 기존의 방식에 비해 처리 속도를 개선하고, 자유로운 카메라 배치 하에서도 깊이 정보 추출이 가능한 방식으로 MR-MBS(Multi-Resolution Multiple-Baseline Stereo) 방식을 제안한다. 또한 폐색 영역의 정합에 있어서도 기존의 방식과는 달리 개선된 방법으로 적용적 폐색 영역 처리 방법론을 제시한다.

### II. 연구의 배경 및 관련 연구

먼저 깊이 정보 추출에 있어 고려해야할 것으로 정확한 깊이 정보의 추출을 위한 대응점 탐색의 어려움과 영상 획득을 위한 카메라 배치의 문제이다. 이러한 문제를 해결하기 위한 방법론은 많이 제시되었지만 그 대부분이 계산 시간의 상당한 증가를 가져오게 된다. 그러나 동영상을 지향하고 있는 영상 처리에 있어서는 계산 시간 또한 소홀히 할 수 없는 문제이다.

대응점을 구하는 데에 있어 장애물은 텍스처가 적은 (low texture) 영역, 반복된 패턴(repeated pattern)이 있는 영역, 폐색(occlusion) 영역, 모양이 변한 영역에 대한 대응점 탐색이 어렵다는 것이다. 이러한 문제를 해결하기 위하여 스테레오 정합은 크게 영역 기반 방식(area-based method)과 특징 기반 방식(feature-based method)으로 연구가 진행되었으나 영역 기반 방식이 주로 이용되고 있다. 그러나, 영역 기반 방식에 있어서 정합창의 크기에 따라서 깊이 맵의 경계선부분이 연장되는(boundary overreach, 이하 BO) 문제와 잡음에 대한 문제가 존재한다[4].

또한 카메라 배치의 제약은 일반적으로 스테레오 정합에 계산의 간소화를 위해 사용하는 에피폴라 구속조건으로 인해 발생한다. 몇몇의 알고리즘의 경우 카메라를 동일 직선 상에 평행 광축으로 배치하여 기존 영상에서 하나의 선상에 있는 모든 점이 다른 영상에서도 하나의 선상에 존재하게 되는 성질을 사용하고 있다. 이를 위해서는 모든 카메라가 동일 직선 상에 평행 광축으로 배치되어야 한다. 그러나 다시점 영상 기술에서는 카메라의 배치가 이와 같이 특정한 배치에 국한되지 않으므로 이러한 구속조건을 사용하는 알고리즘에는 상당한 카메라의 제약이 따른다.

앞서 기술한 깊이 정보 추출 시 문제점들을 해결하기 위해 많은 연구가 진행되어 왔는데 그 중 단순히 두 영상이 아닌 여러 영상을 사용하여 화질 상의 문제를 극복하며 카메라 배치의 자유로움을 갖도록 한 방식으로 MBS (Multiple-Baseline Stereo) 방식[5]을 들 수 있다. MBS 방식은 여러 개의 영상을 스테레오 쌍으로 보아 각각 정합 에리 곡선(SSD)을 구한 뒤 이를 식 (1)을 통하여 동일한 깊이 정보를 갖는 값들을 더하여 Sum of SSD(SSSD)를 비교하여 가장 작은 값을 갖는 깊이 정보를 선택하는 방식으로 단순히 두 장의 영상만을 사용하는 방법보다 정확한 정합을 시도하였다.

$$\frac{d}{BF} = \frac{1}{z} = \xi \quad \text{식(1)}$$

이 방식은 특히 여러 개의 영상을 사용함으로써 텍스처가 적은 부분과 반복되는 패턴이 있는 부분의 정합에 있어 효과적이며 카메라 배치 면에서도 앞서 설명한 카메라 배치의 제약을 가져오는 알고리즘을 사용하지 않으므로 인해 임의의 배치가 가능한 방식이다. 그러나 정합창의 고정으로 인하여 깊이 맵의 BO 문제를 갖고 있고, 또한 폐색 영역에 대한 적절한 처리 방법이 제시되지 않고 있다. 계산량의 측면에서도 두 장의 영상만을 처리하는 스테레오 정합 방식에 비해 MBS 방식은 영상 pair의 수만큼 많은 처리 시간이 요구되어지고 있다.

깊이 맵의 BO 문제를 계층적 깊이 정보 추출 알고리즘을 사용하여 극복하는 Modified MBS 방법[6]이 제시되었다. 또한 Modified MBS 방법은 기존의 MBS 방법이 고

려하지 못하였던 폐색 영역에 대한 처리법을 제시하였다. 그러나 Modified MBS 방법은 처리 시간이 매우 크게 상승되는 문제점을 갖고 있고 폐색 영역 처리법도 정보의 손실과 임의의 배치 상황에 대해서는 폐색 영역 처리의 정확도가 감소되는 문제를 갖고 있다.

### III. 제안하는 방식

본 논문에서는 첫째로 정확한 깊이 정보의 추출과 빠른 처리 시간에 초점을 맞추어 MBS 방식을 변형한 MR-MBS 방식을 제안한다. 기존의 MBS 방식의 문제점이었던 깊이 맵의 BO 문제와 많은 처리 시간의 문제점에 대한 개선책으로 계층적 방식을 제안한다. 일반적인 계층적 방식에서는 영상의 불연속영역에 대해 잘못된 정합이 발생하지만 이에 대한 처리 방법을 제시한다. 그리고 Modified MBS 방식에서 제안한 폐색 영역 처리에서의 정보 손실 문제점을 개선하고 일반적인 카메라 배치의 상황에도 적용할 수 있도록 개선하였다.

그림 1은 제안하는 시스템을 도식하였다. 즉 모든 입력 영상을 가우시안 피라미드로 쌓고 각 단계별로 MBS 방식을 통해 정합을 하며 폐색 영역에 대한 처리를 한다. 최종적으로 구한 깊이 정보에 OVP(Observable Viewpoint) 방법[7]을 적용하여 임의 시점의 영상을 합성하였다.

#### 1. MR-MBS 방식

앞서 설명한 MBS 방법은 고정된 정합창을 사용함으로써 깊이 맵의 BO 문제가 발생한다. 이러한 문제를 보완하면서 보다 처리 속도를 개선시키는 방법으로 가우시안 피라미드 다해상도 방식을 적용하였다. 최저해상도의 단계에서는 앞서 설명한 MBS 방법을 그대로 사용하여 깊이 정보를 구하고, 그 다음 단계부터는 부노드의 깊이 정보를 2배한 뒤 부분 탐색을 한다. 이는 정합창의 크기를 고정하여도 저해상도에서는 큰 정합창을 사용하고 고해상도에서는 작은 정합창을 사용하는 것과 같은 효과를 나타내게 된다. 따라서 작은 정합창을 사용함에도 불구하고 잡음에 대한 민감도를 낮출 수 있으며 깊이 맵의 BO 문제도 동시에 해결이 가능하게 된다. 그리고 작은 정합창의 사용과 탐색의 범위를 줄이는 것은 처리 시간의 개선을 가져온다.

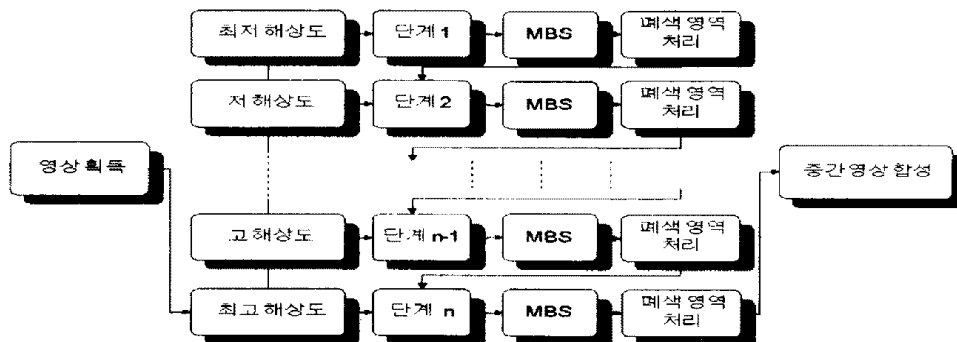


그림 1. 전체 시스템 흐름도

일반적인 다해상도 방법과는 달리 불연속영역에 대한 처리로 그림 2와 같이 부노드 이웃의 깊이 정보를 후보로 하여 모두 부분 탐색을 하여 정합 에러 곡선(SSD)의 분산(Variance of SSD)이 가장 작은 것을 선택하여 이 때의 깊이 정보를 자노드의 깊이 정보로 한다.

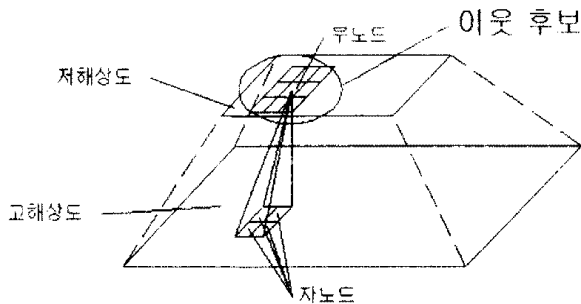


그림 2. 피라미드에서 부노드의 이웃 후보

## 2. 적응적 폐색 영역 처리

본 논문에서는 Modified MBS 방식에서 제안한 폐색 영역 처리 문제를 개선한 적응적 폐색 영역 처리 방법론을 제시한다. Modified MBS 방식은 정합 에러 곡선(SSD)에서 최대, 최소를 제외한 나머지만을 취하여 SSSD를 계산함으로써 폐색 영역을 처리하였으나 이러한 방법은 가지고 있는 영상이 5개 이하로 적은 경우 상당한 정보의 손실을 가져오게 된다. 적응적 폐색 영역 처리 방법은 Modified MBS 방식과는 달리 폐색 영역을 먼저 검출해내어 폐색 영역에 대해서만 처리를 함으로써 정보의 손실을 줄인다. 즉 영상의 모든 화소에 대하여 MR-MBS 방법을 적용하여 SSD의 분산(VSSD)을 구하여 문턱값보다 클 경우 그 화소는 폐색 영역에 속한다고 판단한다.

일반적으로 폐색 영역은 카메라의 배치에 밀접한 영향을 받게 된다. 만약 카메라가 수평으로 배치되어 있을 경우는 거리가 가까운 물체의 왼쪽과 오른쪽 옆부분에 대해 폐색 영역이 나타난다. 이러한 카메라들의 위치 관계는 캘리브레이션[8]으로 구할 수 있다. 먼저 기준(reference) 영상에 대한 각 보조(supporting) 영상들의 카메라 중심 좌표를 구한 뒤, 각 카메라 중심점과 기준 영상에 중심점을 잇는 직선이 이루는 각도에 따라 분류한다. 분류는 그림 3과 같이 영상의 개수에 따라 두 그룹 또는 네 그룹으로 나누고 그 경계는 각 보조 영상들의 각도 간격이 넓은 순으로 선택한다. 이와 같이 그룹을 나눈 뒤 각 그룹별로 MR-MBS 정합 과정을 수행하여 각각의 VSSD를 구하여 가장 작은 그룹의 깊이 정보를 화소의 깊이 정보로 취한다.

## IV. 실험 결과

본 실험에서는 MBS 방식, Modified MBS 방식, 제안하는 방식을 서로 비교 평가하였다. 영상은 일본 쓰쿠바

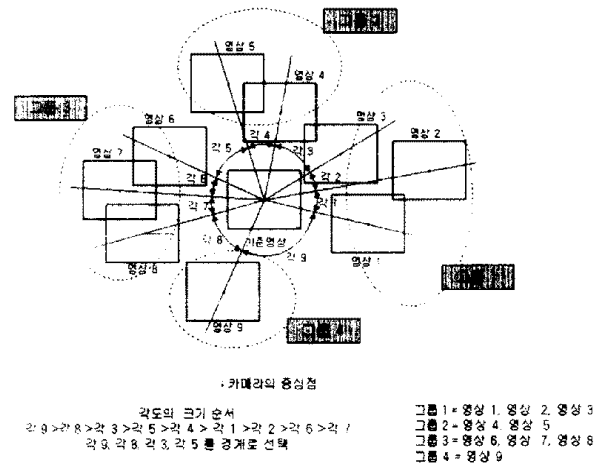


그림 3. 영상들의 위치에 따른 Grouping

대학의 다시점 영상 데이터 베이스[9] 중 Santa 영상을 사용하였다. 이 영상은 동일 직선 상에 평행 광축으로 배치된 카메라로부터 획득된  $320 \times 240$  크기를 갖는 9장의 영상 중에 1, 3, 5, 7, 9번째 영상 5개를 사용하였다. 영상들 간의 baseline의 길이는 40mm이고 초점거리는 10mm이다.

### 1. 화질 및 처리 시간 평가

그림 4는 입력 영상중 세 번째(9장의 영상 중 5번째)의 영상을 기준 영상으로 하였을 경우 구하여진 깊이 맵을 나타내며 기존의 알고리즘과 제안하는 알고리즘으로 각각 구한 결과이다. 이 때 제안하는 알고리즘은  $3 \times 3$ 의 크기를 갖는 정합창을 사용하였다. 그리고 화질 비교를 수치적으로 하기 위하여 OVP 방법[7]을 통하여 9장의 영상 중 4번째 영상을 합성하여 원 영상과의 PSNR값을 구하였다. 표 1의 화질 항목은 바로 PSNR 값을 나타낸다. 표 1의 처리 시간 항목은 Sun UltraSPARC 1 Workstation(Solaris 2.5)의 환경 하에서 실험을 한 경우에 소요되는 시간을 나타낸다.

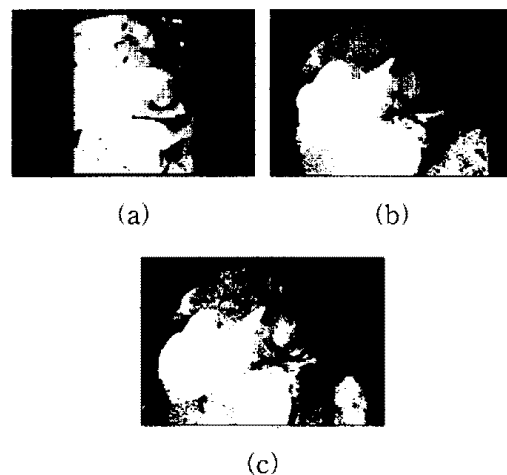


그림 4. 깊이 맵 결과 영상 (a) MBS (b) Modified MBS (c) 제안하는 방법

표 1. 알고리즘의 성능 평가표

	MBS	Modified MBS	제안하는 방식
화질	23.082 dB	29.743 dB	31.900 dB
처리 시간	3분 6초	28분 35초	21초

또한 BO 현상의 비교를 위해 각 정합 알고리즘에 동일한 폐색 영역 처리 방법을 사용한 결과를 비교한 결과 MBS 방식에서는 경계선 부분이 원영상에 비해 배경 쪽으로 4 화소 늘어난 반면 Modified MBS 방식과 제안하는 방식에서는 1 화소 정도도 거의 늘어나지 않았다.

표 2는 Modified MBS 방식에서 제안한 폐색 영역 처리법과 본 논문에서 제안한 적응적 폐색 영역 처리 방법의 비교를 나타내며 이를 위하여 기본적인 정합 알고리즘인 MBS 방식과 Modified MBS 그리고 MR-MBS 방식에 두 가지 방식의 폐색 영역 처리법을 각각 취하였을 경우 합성 영상의 PSNR을 구하여 비교하였다. 표 2의 폐색 영역 처리 1은 Modified MBS 방식에서 제안한 폐색 영역 처리법을 나타내며 폐색 영역 처리 2는 제안하는 방식인 적응적 폐색 영역 처리 방법을 나타낸다.

표 2. 폐색 영역 처리법 성능 평가표

	MBS	Modified MBS	MR-MBS
폐색영역처리 1	25.678 dB	29.743 dB	26.136 dB
폐색영역처리 2	31.444 dB	32.629 dB	31.900 dB

표 1과 표 2를 통하여 본 논문에서 제안하는 방식의 화질이 우수함을 볼 수 있다. 그러나 동일한 폐색 영역 처리법을 가하였을 경우에는 본 논문에서 제안하는 MR-MBS 방식의 화질이 Modified MBS 방식보다 약간 떨어지는 것을 볼 수 있다. 이는 정합창의 크기에 변화를 주며 정합 하는 방식이 정합창의 크기는 고정하고 영상의 크기에 변화를 주며 정합 하는 방식보다 정합의 정확도가 높음을 나타낸다. 반면 처리 시간에 있어서는 본 논문에서는 제안하는 방법이 MBS와 Modified MBS 방법에 비해 상당히 개선됨을 볼 수 있다. 또한 제안하는 적응적 폐색 영역 처리 방법 또한 기존 폐색 영역 처리 방법에 비해 많은 화질 개선 효과가 나타나는 것을 볼 수 있다.

## V. 결론

MR-MBS가 계층적 방법을 사용함으로써 가져오는 효과는 크게 두 가지이다. 먼저 깊이 맵의 BO 문제에 대한 해결책이 되며 이를 Modified MBS 방식에서는 계층적 window 정합을 통해 극복한 반면 본 연구에서는 다해상도 방식을 적용하여 극복하였다. 또한 정합창의 크기를 감소시키고 탐색 영역을 줄여 처리 시간의 개선을 가져왔다.

계층적 방법을 사용함에 있어서 일반적인 계층적 방법이 불연속영역에 대해 잘못된 정합을 하는 문제를 본 논문에서는 9 후보 선택 방법을 통해 감소시켰다.

또한 폐색 영역에 대한 처리 역시 MBS에서는 전혀 고려하지 않고, Modified MBS 방법에서도 대안을 제시하고 있으나 이 또한 정보의 손실 등의 문제점을 갖고 있다. 반면에 본 논문에서는 폐색 영역을 먼저 찾아낸 뒤 찾은 영역에 대해서만 카메라 위치 관계를 고려하여 폐색 영역 처리를 수행함에 의해 정보의 손실을 줄이며 정확한 깊이 맵을 구할 수 있었다.

앞으로 실제 카메라를 사용하여 임의의 위치에 카메라를 배치한 환경에서의 영상을 구하는 작업이 남아 있고 또한 고속 하드웨어를 통한 처리 시간의 개선에 대하여 좀더 많은 연구가 이루어 져야 한다.

## 참고문헌

- [1] D. B. Diner and D. H. Fender, "Human engineering in stereoscopic viewing devices," *Plenum Press*, 1993.
- [2] S. E. Chen and L. Williams, "View interpolation for image synthesis," *Computer Graphics Proceedings, Annual Conferences Series*, pp. 279-288, 1993.
- [3] L. McMillan and G. Bishop, "Plenoptic modeling: an image-based rendering system," *Computer Graphics Proceedings, Annual Conferences Series*, pp. 39-46, 1995.
- [4] U. Dhond and J. Aggarwal, "Structure from stereo: a review," *IEEE Trans. System, Man, and Cybernetics*, Vol. 19, No. 6, pp. 1489-1510, Nov./Dec. 1989.
- [5] M. Okutomi and T. Kanade, "A multiple-baseline stereo," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 4, pp. 353-363, 1993.
- [6] 박종일, 이노우에 세이키, "여러 대의 카메라를 이용한 계층적 깊이정보 추출 알고리즘," 방송공학회지, 제2권
- [7] Jong-II Park and Seiki Inoue, "Arbitrary view generation from multiple cameras," *Proc. IEEE ICIP'97 to be appeared*
- [8] R. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE J. Robotics and Automation*, Vol. RA-3, No. 4, pp.323-344, Aug. 1987.
- [9] Y. Ohta, et al., "Multiview image database for computer vision research : CD-ROM." Feb. 1995.