

# 스테레오 내시경 영상의 깊이정보추출 알고리즘 개발

이상학\*, 김정훈\*, 황도식\*, 송철규\*\*, 이영목\*\*, 김원기\*\*, 이명호\*

\*연세대학교 공학대학 전기공학과

\*\*삼성생명과학연구소 임상의공학센터

## Development of Algorithm for Depth Extraction in Stereo Endoscopic Image

S. H. Lee\*, J. H. Kim\*, D. S. Hwang\*, C. G. Song\*\*, Y. M. Lee\*\*, W. K. Kim\*\*, M. H. Lee\*

\*Department of Electrical Engineering, Yonsei University

\*\*Biomedical Engineering Center, Samsung Biomedical Research Institute

### ABSTRACT

This paper presents the development of depth extraction algorithm for the 3D Endoscopic Data using a stereo matching method and depth calculation. The purpose of other's algorithms is to reconstruct 3D object surface and make depth map, but a one of this paper is to measure exact depth information on the base of [cm] from camera to object. For this, we carried out camera calibration.

### 1. 서론

최근에 내시경을 이용한 환자의 장기관찰 및 무혈수술에 대한 관심이 고조되면서 두 개의 카메라를 장착하여 양안 시차효과를 이용, 3차원 효과를 발생시키는 3차원 내시경의 사용빈도가 임상에서 증가되고있는 추세이다. 따라서 본 논문에서는 이러한 3차원 내시경을 이용한 수술 및, 관찰에 도움을 줄 수 있도록 3차원데이터의 깊이(depth)를 연산, 표시할 수 있는 알고리즘을 개발하였다. 기존의 깊이추출 알고리즘의 목적은 3차원 영상의 재구성에 집중되어 있었으나 본 논문에서는 카메라로부터 대상체까지 깊이를 실세계에서 사용하는 단위[cm]를 포함하여 정확하게 추출하는 것이다.

두 개의 렌즈를 이용한 스테레오(stereo)영상에서 깊이(depth)를 추출하기 위해서는 좌/우 영상의 대응점을 찾는 정합(matching)이 가장 중요한 작업이며 특징점(feature)의 선택이 상당히 중요하다. 주어진 두 개의 영상에서 대응점을 찾는 일치(correspondence)과정은 좌/우 두개의 영상에 나타나고 있는 두 개의 점이나 특징점사이에서 수행된다. 그리고 일반적으로 정합할 각 픽셀의 밝기(intensity)와 같은 특징점들의 휘도 레벨이 작고 밀도가 높은 경우에는 영역정합(area-based matching)이라고 부르는 정합방법을 이용하고, 반면에 경계(edge)나 코너(corner)와 같이 일정 간격을 두고 나타나고 있는 좀더 추상적인 특징에 대해서는 특징점정합(feature-based matching)이라고

부르는 정합 기법을 사용한다.

본 논문에서는 상기 언급한 두 가지 기법을 모두 적용하였으며 대응쌍을 이용한 시차(disparity)를 이용하여 기하학적인 방법에 근거한 깊이정보와 카메라보정을 통해 단위성분(scale factor)를 추출하였다. 영상은 삼성의료원에서 제공한 내시경영상과 3D 내시경카메라를 이용해 직접 획득한 영상을 사용하였다.

### 2. 카메라보정과 특징점추출

#### 2.1 카메라보정

공간상의 한 정점이 카메라의 영상면에 투영되는 관계를 알아내는 것이 카메라보정이라고 할 수 있다. 본 논문에서는 Richard Wolf 3D카메라를 보정하였다. 카메라보정을 위해 이미 알고있는 실세계 좌표와 대응되는 몇 개의 2차원 영상의 좌표로부터 우회적인 방법으로 행렬의 파라미터를 구하는 것으로 일반적인 행렬의 형태는 아래와 같다.

$$[C] = [A][X]$$

$$\begin{bmatrix} c_{h1} \\ c_{h2} \\ c_{hm} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.1)$$

여기서,  $x = \frac{c_{h1}}{c_{hm}}$ ,  $y = \frac{c_{h2}}{c_{hm}}$  이다. 식(2.1)를 정리하면 식(2.2)가 생성된다.

$$\begin{aligned} a_{11}X + a_{12}Y + a_{13}Z - a_{41}xX - a_{42}xY - a_{43}xZ - a_{44}x + a_{14} &= 0 \\ a_{21}X + a_{22}Y + a_{23}Z - a_{41}yX - a_{42}yY - a_{43}yZ - a_{44}y + a_{24} &= 0 \end{aligned} \quad (2.2)$$

즉, 실세계 좌표 1개와 2차원 영상좌표 1개의 대응쌍에 대해 위와같은 두 개의 식이 성립한다. 따라서 행렬의 파라미터 개수 12개를 구하기 위해서는 실세계 좌표와 2차원 영상좌표의 대응쌍 6개를 알면 필요로하는 12개의 파라미터들을 구할 수 있다. 그러나 일반적으로 homogeneous 좌표계에서  $a_{44}$ 의 값을 "1"로 두기 때문에 실제적인 파라미터의 개수는 11이다.

본 실험을 통해 추출한 카메라 파라미터의 행렬 A는 식(2.3)이다.

$$A = \begin{bmatrix} -1160.9 & -74.1859 & 129.969 & -12208.9 \\ 17.9678 & 1077.64 & -186.809 & -1291.76 \\ 0.00492613 & -0.294325 & -0.553661 & 1 \end{bmatrix} \quad (2.3)$$

2.2 특징점추출

가장 널리 사용되는 특징점으로는 밝기의 불연속이 일어나는 경계부분이다. 특히 이 점들은 대부분의 장면정보를 포함하기 때문에 상당히 유용하다고 할 수 있다. 본 논문에서는 1차원 특징인 경계보다는 2차원적인 특징을 가지고 있는 코너(corner)를 검출하였는데 이것은 경계에 비해 상대적으로 수는 작지만 포함하고 있는 정보의 양은 유사하고 정합에도 상당한 이점을 가지고 있기 때문이다. 본 논문에서는 코너를 추출하기 위해 수잔코너검출기(SUSAN corner detector)를 사용하였다.

3. 입체정합(stereo matching)

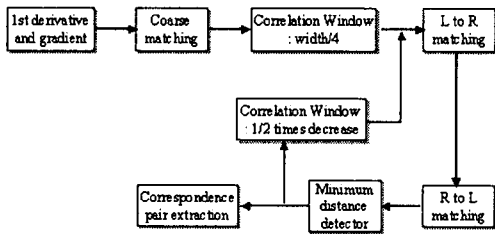


그림 3.1 입체정합과정의 블록선도  
Fig. 3.1 Block diagram of stereo matching process

3.1 미분값과 방향성을 이용한 정합

좌/우 영상에서 추출된 특징점들에 대해 영역 상관도를 이용하여 정합을 시도하기엔 특징점의 수가 지나치게 많을 수 있다. 그런 이유에서 1차적으로 특징점들의 미분값과 방향성만으로 정합을 한 후에 그 결과들만으로 영역정합을 수행하는 것이 이점이 있다. 그림 3.2에서 왼쪽영상의 한 특징점  $m_1$ 에 대해서 우측 영상의 각 특징점들을 조사하는데 좌/우 영상의 극상선(epipolar line)은 수평방향으로 일치한다고 가정한다. 그리고 오른쪽 영상에서 수평방향으로는  $m_1$ 과 같은 위치에서 최대시차(MD : Maximum Disparity)를 고려하고 수직방향으로는 카메라시스템 구성중의 오류를 감안하여  $\pm k$ 만큼의 searching 윈도우내에서 대응점을 찾으려 한다. 즉

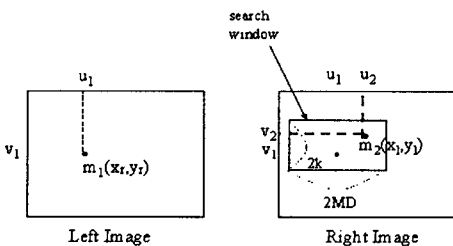


그림 3.2 가정합을 위한 searching 윈도우  
Fig. 3.2 Searching window for coarse matching

searching 윈도우는  $2MD \times 2k$ 의 크기를 가진다.

좌측영상의 특징점  $m_1$ 에 대해 우측영상의 searching 윈도우내의 특징점들을 조사한다.  $m_1$ 의 1차미분값  $M_1$ 과 gradient  $\alpha_1$ 의 값과 우측영상의 searching 윈도우내의 특징점들의  $M_k$ 와  $\alpha_k$ 의 값을 비교하여 일정 임계치(threshold)내의 오차를 가지는 우측영상의 특징점들(R)을 추출한다.

$$R = \{ m_r(x_r, y_r) \mid m_1(x_l) - MD \leq x_r \leq m_1(x_l) + MD, \\ m_1(y_l) - k \leq y_r \leq m_1(y_l) + k \} \\ |M_{x_r} - M_{x_l}| \leq \epsilon, \quad |\alpha_{x_r} - \alpha_{x_l}| \leq \beta \quad (3.1)$$

3.2 상관도를 이용한 정합

본 연구에서 좌/우 영상의 특징점에 대한 영역의 유사성을 조사하는 방법으로 영역상관도(correlation)를 이용하였다. 상관도는 특징점을 중심으로 일정 크기의 윈도우를 씌우 평균값과 표준편차(standard deviation)를 측정하고 각각의 윈도우에서 평균값에 대한 분산(variation)을 구하여 상관도에 대한 결과(score)를 산출한 후 이 값의 크기에 따라 상관도의 정도를 판단하는 것이다.

좌측영상의 특징점  $m_1$ 과 우측영상의 searching 윈도우내의 특징점  $m_2$ 에 대한 상관계수 score( $m_1, m_2$ )는 식(3.2)이다.

$$Score(m_1, m_2) = \frac{\sum_{i=-n}^n \sum_{j=-m}^m [I_1(u_1 + i, v_1 + j) - \overline{I_1(u_1, v_1)}] \times [I_2(u_2 + i, v_2 + j) - \overline{I_2(u_2, v_2)}]}{(2n+1)(2m+1)\sqrt{\sigma^2(I_1) \times \sigma^2(I_2)}} \quad (3.2)$$

$$\overline{I_k(u, v)} = \frac{\sum_{i=-n}^n \sum_{j=-m}^m I_k(u+i, v+j)}{(2n+1)(2m+1)} \\ \sigma(I_k) = \sqrt{\frac{\sum_{i=-n}^n \sum_{j=-m}^m [I_k(u, v) - \overline{I_k(u, v)}]^2}{(2n+1)(2m+1)}} \quad (3.3)$$

상관도(correlation) 윈도우의 크기  $n, m$ 은 크기가 클수록 영역의 유사성에 대해서 좋은 판단을 할 수 있지만 연산시간이 길어지는 단점이 있고 상관도 윈도우의 크기가 작을수록 연산시간이 단축되지만 영역의 정보를 많이 포함하고 있지않기 때문에 영역의 유사성의 판단에 오류가 발생할 수도 있다. 이번의 실험에서는 최소  $n=m=7$ 의 크기를 가지는 윈도우를 이용하여 상관도를 측정하였다.

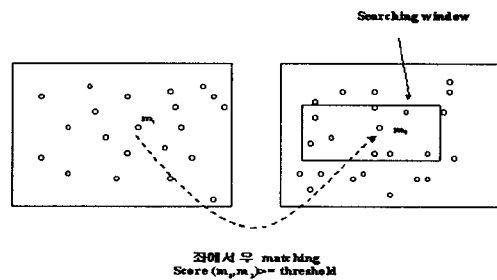


그림 3.3 좌측영상의 특징점( $m_1$ )에 대한 우측영상내의 대응쌍( $m_2$ )  
Fig. 3.3 Candidate match in right image for feature point in left image

그림 3.3에서 score의 임계치(threshold)를 윈도우의 크기에 따라 0.75에서 0.95정도로 변화시켜 가면서

좌측영상의 특징점( $m_1$ )에 대해 임계치 보다 큰 우측영상의 특징점( $m_2$ )을 찾아  $m_1$ 의 대응후보쌍으로 정한다.

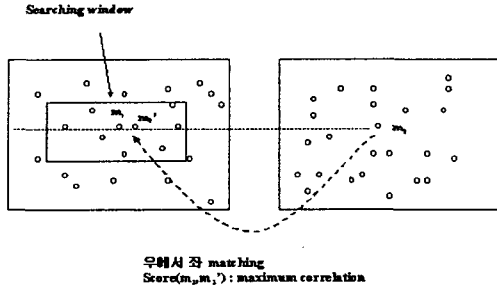


그림 3.4 우측영상의 특징점( $m_2$ )에 대한 좌측영상내의 대응점 ( $m_2'$ )과 원 특징점( $m_1$ )의 비교  
Fig. 3.4 Comparison of  $m_1$  and  $m_2'$  for  $m_2$

대응후보쌍( $m_2$ )에 대해서 우측영상의 특징점( $m_2$ )을 기준으로, 좌측영상의 같은 searching 윈도우에서 특징점들에 대한 상호 상관도를 조사하여 가장 큰 상관도를 갖는 것을 찾도록한다. 그리고 이것을 우측영상의 특징점( $m_2$ )에 대한 좌측영상의 대응후보쌍( $m_2'$ )으로 판단한다. 우측영상의 특징점( $m_2$ )에 대한 좌측영상의 정합된 대응후보( $m_2'$ )의 위치가 최초 좌측영상의 특징점( $m_1$ ) 위치와 비교하여 거리오차가 임계치 이하가 되면 좌/우 대응후보쌍(Left/Right candidate correspondence pair : L/R candidate match)으로 인식한다.

$$|m_1 - m_2| = \text{distance} = \begin{cases} \text{L/R candidate match} & \text{if distance} \leq \text{threshold} \\ \text{ignored} & \text{if distance} > \text{threshold} \end{cases} \quad (3.4)$$

R의 다른  $m_k$ 에 대해 위의 과정을 다시 반복하여  $m_1$ 에 대한 좌/우 대응후보쌍(L/R candidate match)들의 집합을 C라고 하면,

$$C = \{ (m_1, m_k) \mid |m_1 - m_k| = \text{distance} \leq \text{threshold}, \text{score}(m_1, m_k) \geq \text{threshold} \text{ and } \text{score}(m_k, m_1) = \text{max correlation} \} \quad (3.5)$$

집합 C의 좌/우 대응후보쌍에서 최소의 distance를 갖는  $m_k$ 를  $m_1$ 에 대한 최종 대응쌍으로 판단한다.

### 3.3 가변윈도우를 이용한 정합

상관도 윈도우의 크기가 클수록 영역의 많은 정보를 포함하기 때문에 영역의 유사성에 대한 진단을 정확하게 할 수 있다. 그래서 본 연구에서는 영상의 수평크기 1/4만큼의 윈도우에서부터 반씩 윈도우의 크기를 줄여나가면서 최소  $n = m = 7$ 까지 윈도우의 크기를 변화시켰다. 처음의 상관도 윈도우의 크기에서 좌/우 정합의 score에 대한 임계치는 실험적으로 0.85정도였다. 이런 윈도우에서 앞서의 진단과정을 거친 좌/우 대응쌍의 정확도는 상당히 높고 영상내에서 주요 대상체가 놓여있는 가운데 부분에서 대응쌍이 형성된다.

저차 상관도 윈도우를 이용한 좌/우 대응후보쌍은 고차로 진행될 때 상관도 조사를 위한 특징점에서 제외한다. 이것은 저차 상관도 조사에서는 상관도 윈도우의 크기가 크고 정확하기 때문에 여기서

대응후보쌍으로 선택된 특징점들은 제외하여야 고차 상관도 조사과정에서 일어날 수 있는 정합의 오류를 줄일 수 있다. 그리고 2차 윈도우를 통한 좌/우 상관도 정합에서 score의 임계치는 0.75였고 윈도우의 크기가 작아지는 3차 이상에서는 score를 0.75에서 일정값을 증가시켰다. 이는 윈도우의 크기가 작을수록 유사성을 나타내는 score의 크기가 커야만 하기 때문이다.

## 4. 깊이정보 추출

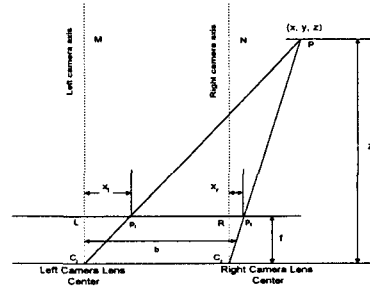


그림 4.1 정합에 대한 기하학적 구성도  
Fig. 4.1 Geometry for stereo matching

그림 4.1에서 삼각형  $PMC_1$ 과  $pLC_1$ 의 비례와  $PNC_1$ 과  $pRC_1$ 의 비례를 이용하면 식(4.1)의 깊이를 얻는다.

$$z = \frac{bf}{(x'_i - x_r)} \quad (4.1)$$

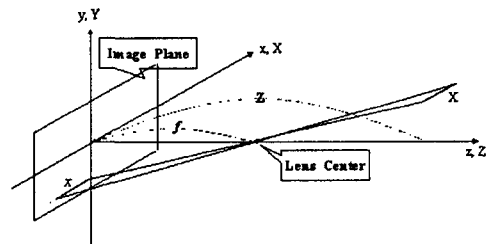


그림 4.2 단위성분 추출을 위한 기하학적 구성도  
Fig. 4.2 Geometry for extracting scaling factor

식(4.1)의 분모는 픽셀단위이므로 초점거리(f)를 픽셀단위로 변환해야한다. 본 논문에서 사용한 Richar Wolf 3D카메라의 초점거리는 27mm인데 그림 4.2의 비례관계를 이용하면, 식(4.3)이 성립한다.

$$X : x = Z + a : f \quad (4.2)$$

$$f(\text{pixel}) = \frac{x(\text{pixel})(Z + a)(\text{mm})}{X(\text{mm})} \quad (4.3)$$

식(4.3)의  $a$ 는 렌즈의 외형때문에 발생한 오차를 보정하기 위한 성분이다. 초점거리  $f$ 는 항상 일정하므로 그림 2.1의 각각 좌표에 대해 평균거리  $X$ 와 평균픽셀거리  $x$ , 실제 깊이  $Z$ 를 각각 대입하여 연립방정식을 풀면  $f$ 는 2156.632[pixel]이고  $a$ 는 0.778이다.

5. 실험결과 및 결론

동물복강의 내시경 영상(A)와 3D 카메라를 이용하여 직접 획득한 영상(B)에 대한 좌측영상의 특징점 추출결과를 그림 5.1에 도시하였다.

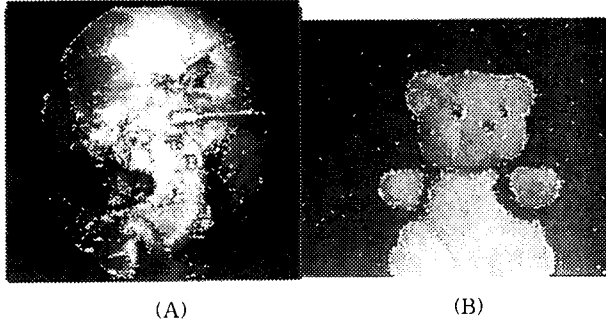


그림 5.1 특징점 추출결과  
Fig. 5.1 Results of feature extraction

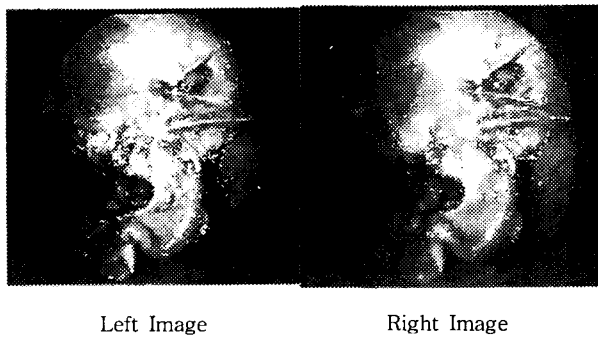


그림 5.2 입체정합 후의 결과  
Fig. 5.2 Results of stereo matching

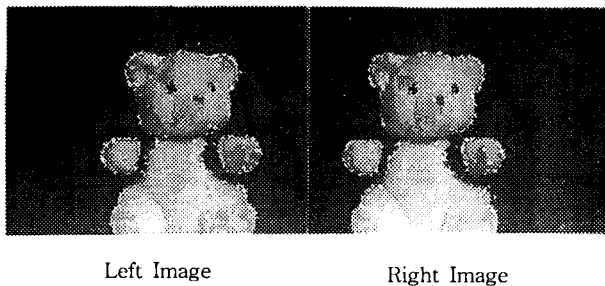


그림 5.3 입체정합 후의 결과  
Fig. 5.3 Results of stereo matching

그림 5.2는 그림 5.1(A)에 대한 입체정합 후의 대응쌍들을 나타내고 있다. (A)의 경우는 직접 실험을 통해 얻은 것이 아니므로 단위성분을 찾지 못했다. 그림 5.2(A)의 대응쌍 중에서 대표할 수 있는 몇 부위의 깊이에 대한 것을 그림 5.4에 도시하였다. 그리고 그림 5.3은 그림 5.1(B)에 대해 입체정합 후의 대응쌍들을 나타내고 있다. 이것의 대응쌍들에 대해서 실측깊이를 구할 수 있었는데 표 5.1에 카메라로부터 곰인형의 각 부위까지 실제거리와 입체정합 결과를 통해 얻은 깊이를 비교하였다.

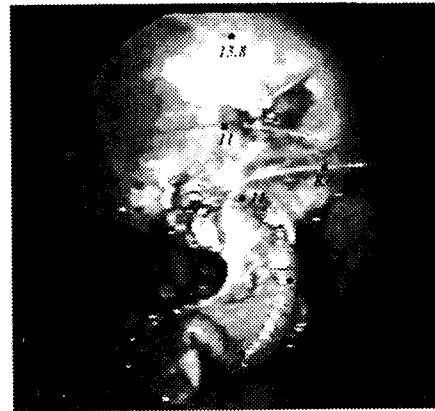


그림 5.4 (A)에 대한 깊이표시  
Fig. 5.4 Depth for sample (A)

표 5.1 (B)의 실제깊이와 본 논문에서 제안한 정합 후 깊이와의 비교

Table 5.1 Comparison of real depth and the depth after proposed matching processing

	시 차 (disparity)	본 논문에서 제안한 입체정합 후의 깊이[cm]	실제깊이 [cm]
코	107	43.6	44.5
왼쪽 눈	105	44.4	45.2
오른쪽 눈	104	44.8	45.5
오른손	102	45.7	45
왼손	104	44.8	44.5
오른쪽 귀	100	46.7	47
오른쪽 발	105	44.4	44.4

표 5.1에 나타난 실제깊이와 제안된 깊이와의 오차는 대략 1.52% 정도로 상당히 정확한 깊이를 찾아냈음을 나타내고 있다. 이상의 결과들에서 볼 수 있듯이 본 논문에서는 기존의 3차원 재구성을 위한 입체정합과는 달리 대상체까지의 정확한 깊이를 검출하였다. 그리고 내시경영상에 적합한 특징점 추출과 각종 입체치를 설정하였고 실제 3D 내시경 카메라의 보정을 통해 카메라모델의 파라미터를 추출하였다.

이 연구는 1996년도 보건복지부 G7 의료공학기술 개발사업의 연구지원에 의해 수행되었음.  
(# HMP-96-G-1-5)

6. 참고문헌

[1] Zhengyon ZHANG. "A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry", INRIA Sophia Antipolis, pp.3-10, 1994  
 [2] Steven Douglas Cochran. "Surface Description from Binocular Stereo", Volume1, Faculty of the Graduate School University of Southern California, pp.1-19, 1990.  
 [3] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Addison-Wesley Publishing Company, pp.51-71, 1993