

〈論 文〉

## 다중 스테레오영상을 이용한 대응문제의 해결과 거리오차의 해석

이재웅\* · 이진우\* · 박광일\* · 오준호\*\*

(1994년 10월 31일 접수)

### Solving the Correspondence Problem by Multiple Stereo Image and Error Analysis of Computed Depth

Jae Woong Yi, Jin Woo Lee, Kwang Il Park and Jun Ho Oh

**Key Words:** Stereo(스테레오), Multiple Images(다중 영상), 3-D Recovery(3차원 형상인식)

#### Abstract:

In this paper, we present a multiple-view stereo matching method in case of moving in the direction of optical axis with stereo camera. Also we analyze the obtainable depth precision to show that multiple-view stereo increases the virtual baseline with single-view stereo. This method decides candidate points for correspondence in each image pair and then search for the correct combinations of correspondences among them using the geometrical consistency they must satisfy. Advantages of this method are capability in increasing the accuracy in matching by using the multiple stereo images and less computation due to local processing. This method computes 3-D depth by averaging the depth obtained in each multiple-view stereo. We show that the resulting depth has more precision than depth obtainable by each independent stereo when the position of image feature is uncertain due to image noise. This paper first defines a multiple-view stereo algorithm in case of moving in the direction of optical axis with stereo camera and analyze the obtainable precision of computed depth. Then we represent the effect of removing the incorrect matching candidate and precision enhancement with experimental result.

#### 1. 서 론

스테레오는 2차원 영상으로부터 주위 환경의 3차원 거리정보를 구하는 유용한 방법이다. 스테레오 매칭으로 거리를 구하는 과정은 특징추출, 스테레오 특징 정합, 정합점으로부터 3차원 좌표계산이 이루어진다.<sup>(1)</sup> 이 중 왼쪽영상의 한 점에 대하여 대응

하는 오른쪽영상의 점을 찾아내는 대응문제(correspondence problem)의 해결이 가장 어려운 단계이다.<sup>(2)</sup>

대응문제를 해결하기 위해 사용되는 방법중의 하나는 다중 해상도대응(coarse-to-fine matching) 방법이다.<sup>(3~4)</sup> 잘못된 정합을 줄이기 위하여 낮은 해상도에서 탐색영역을 정하고 높은 해상도에서 정확한 정합을 행한다. 그러나 이 방법으로도 규칙적으로 반복되는 물체 등에 대하여는 항상 정합을 잘 할 수 없다. 다른 접근으로 카메라의 미소운동에

\*한국과학기술원 기계공학과

\*\*정회원, 한국과학기술원 기계공학과

의해 다중 영상을 이용하는 방법이 있다.<sup>(5-7)</sup> 연속되는 영상에 대하여 카메라간격(baseline)이 작기 때문에 영상간의 특징점이 정합이 쉽다는 장점을 가진다. 하지만 이 방법은 영상간의 영상운동량(optical flow)를 쉽게 구하기 위하여 카메라의 미소운동을 가정하였기 때문에 운동로봇과 같이 운동을 하면서 경로계획을 할 경우 카메라의 운동속도를 제한하는 단점이 있다.

본 논문에서는 다중 스테레오 매칭 알고리즘을 이용하여 대응문제를 해결하는 과정과 기존의 스테레오 매칭 알고리즘과 비교하여 얻을 수 있는 거리 정보의 오차를 해석한다. 제안하는 다중 스테레오 매칭 알고리즘은 카메라의 광축방향의 운동에 의해 얻어지는 여러 쌍의 스테레오 영상을 이용한다. 본 논문에서 가정하는 상황은 (1) 주위환경은 고정되어 있으며, (2) 스테레오 관측자의 운동파라미터를 알고 있고, (3) 스테레오 관측자의 운동이 커서 각각의 카메라의 영상면에 대한 영상운동량(optical flow)를 구하기 힘들다는 것이다. 기존 스테레오 영상에 대하여 후보 대응 특징점을 찾고, 카메라의 운동 파라미터로부터 다른 위치에서의 스테레오영상에서 해당하는 점을 가정하고 특징점의 속성이 유지되는가를 확인하는 과정을 통하여 스테레오 매칭이 행하여 진다. 이 방법의 장점은 다중 스테레오영상을 사용함으로써 잘못된 대응점의 개수를 줄일 수 있고 스테레오 매칭과정이 국소적이기 때문에 계산량이 작다는 데 있다. 다중 스테레오 매칭의 결과로부터 거리는 각각의 스테레오 쌍에서 얻어지는 거리의 평균값을 취하여 얻는다. 이렇게 구한 거리는 영상에 잡음이 존재하여 특징점의 위치에 불확실성이 존재할 때 각각의 스테레오 쌍에서 독립적으로 얻을 수 있는 거리보다 더 나은 정확도를 가짐을 보인다.

본 논문에서는 먼저 다중 스테레오 매칭의 원리를 정의하고, 계산된 거리의 오차를 해석하며 잘못된 대응점의 개수를 줄여나가는 과정을 실제 영상에 대해 실험적으로 보인다. 또한 실제 다중 스테레오영상에 대하여 스테레오 매칭을 행하고 거리오차를 구함으로써 제안한 방법의 효용성을 보인다.

## 2. 다중 스테레오 매칭 알고리즘

### 2.1 다중 스테레오비전 시스템의 구성

이 장에서는 다중 스테레오비전 시스템의 구성과

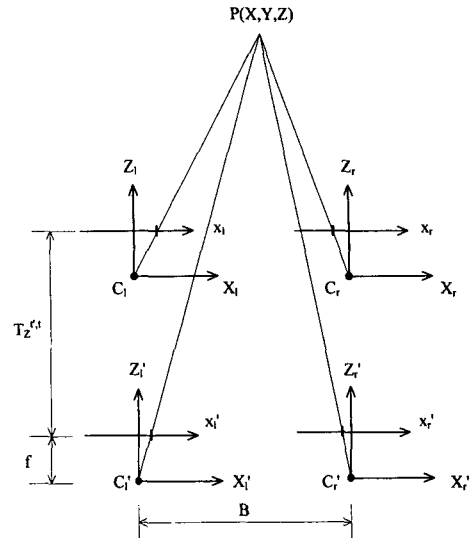


Fig. 1 Configuration of multiple-view stereo

카메라의 운동에 따른 영상점의 위치변화를 구하기로 한다. 다중 스테레오 카메라의 구성은 Fig. 1과 같다. 각각의 카메라에 대하여 pinhole 카메라 모델을 가정하고 있다.  $C_l, C_r$ 은 왼쪽과 오른쪽 카메라의 렌즈의 중심이며 촛점거리는  $f$ 이고 카메라사이의 거리(baseline)는  $B$ 이다. 카메라 좌표계  $X_l Y_l Z_l, X_r Y_r Z_r$ 는 각각 카메라의 렌즈중심  $C_l, C_r$ 에 원점을 가지고 있다. 영상좌표계  $x_l y_l, x_r y_r$ 는 각각의 영상면에 위치하며 영상면의 중심  $O_l, O_r$ 에 원점을 둔다. 왼쪽 카메라의  $x_l, y_l$ 축은 각각  $X_l, Y_l$ 에 평행하며 오른쪽 카메라의 경우도 마찬가지이다.

3차원 공간상의 한 점  $P$ 가 각각의 카메라 좌표계에 대하여  $(X_l, Y_l, Z_l), (X_r, Y_r, Z_r)$ 의 좌표를 가진다고 하자. 점  $P$ 을 왼쪽과 오른쪽 카메라의 영상면에 각각  $p_l, p_r$ 으로 원근투영(perspective projection)되며 그 좌표는 각각  $(x_l, y_l), (x_r, y_r)$ 이다. pinhole 카메라의 가정과 함께 또 하나의 가정은 두 카메라는 평행하며 같은 높이를 가진다는 것이다. 이런 가정으로부터 3차원 공간상의 점  $P$ 의 왼쪽과 오른쪽 카메라의 영상면으로의 투영점은 같은 수평선상에 존재하게 됨을 알 수 있다. (즉  $y_l = y_r$ 이다.) 이상의 스테레오 카메라의 구성에서 두 영상점  $p_l, p_r$ 으로부터 간단히 삼각법으로 점  $P$ 의 좌표를 구할 수 있다.

$$P : (X_l, Y_l, Z_l) = \left[ \frac{Bx_l}{d}, \frac{By_l}{d}, \frac{Bf}{d} \right] \quad (1)$$

점  $P$ 의 좌표는 왼쪽 카메라좌표계에 대하여 나타내었으며 이후에도 계속 3차원 좌표를 나타내는 기준 좌표계로 한다. 식(1)에서  $d(=x_l-x_r)$ 는 왼쪽, 오른쪽 영상면에서 대응하는 영상점간의 거리(disparity)를 나타낸다.

카메라의 운동이 광축방향의 병진운동만으로 이루어질 때 카메라의 운동 전 시간  $t'$ 에서 왼쪽과 오른쪽 카메라좌표계에 대한 점  $P$ 의 좌표를 각각  $(X'_l, Y'_l, Z'_l), (X'_r, Y'_r, Z'_r)$ 이라 하면

$$X'_l = X_l, Y'_l = Y_l, Z'_l = Z_l + T_z \quad (2.1)$$

$$X'_r = X_r, Y'_r = Y_r, Z'_r = Z_r + T_z \quad (2.2)$$

의 관계에 있다.  $(X'_l, Y'_l, Z'_l), (X'_r, Y'_r, Z'_r)$ 에 대응하는 영상면에서 영상점  $p'_l, p'_r$ 의 좌표를 각각  $(x'_l, y'_l), (x'_r, y'_r)$ 라 하면

$$x'_l = \frac{x_l}{1+md}, y'_l = \frac{y_l}{1+md} \quad (3.1)$$

$$x'_r = \frac{x_r}{1+md}, y'_r = \frac{y_r}{1+md} \quad (3.2)$$

의 관계가 있으며,  $m = \frac{T_z}{Bf}$ 이다.

### 2.2 다중 스테레오 매칭 알고리즘

기준 카메라좌표계에서 한 영상점에 대한 3차원 좌표를 구하기 위해서는 대응하는 스테레오카메라의 영상점 사이의 대응관계를 알아야 한다. 즉 왼쪽 카메라의 영상면의 한점  $(x_l, y_l)$ 에 대하여 대응하는 오른쪽 영상면의 한점  $(x_r, y_r)$ 를 결정하여 영상점거리(disparity)  $d(=x_l-x_r)$ 를 구하면 식(1)로부터 3차원 좌표를 구할 수 있다. 하지만 오른쪽 영상면에서 대응 후보점은 일반적으로 epipolar line 상에 여러개 존재하기 때문에 대응문제(correspondence problem)가 발생한다. 본 논문에서는 한쌍의 스테레오영상만을 가지고 스테레오 대응점을 구하는 기존의 방법과는 달리 스테레오 관측자가 운동을 하면서 얻은 다중 스테레오영상으로부터 대응문제를 해결한다.

시간  $t$ 에서 왼쪽 카메라의 영상  $I_t^l$ 을 기준 영상으로 하여 스테레오 매칭을 수행하는 과정을 설명한다. 왼쪽 카메라의 영상  $I_t^l$ 에서 한 특징점  $p_t^l$ 에 대하여 대응하는 오른쪽 카메라의 영상  $I_t^r$ 에서 유사 특징점의 후보집합을 구한다. 시간  $t$ 와  $t'$ 사이의 카메라 운동을  $T, R$ 로써 알고 있으므로 식(3)으로부터 시간  $t'$ 에서 해당하는 특징점의 위치  $p_t^{l'}$ ,

$p_t^{r'}$ 를 가정할 수 있다. 가정된 영상특징 위치  $p_t^{l'}$ ,  $p_t^{r'}$ 에서  $p_t^l, p_t^r$ 에서의 영상특징의 속성이 유지되는가를 확인함으로써 시간  $t$ 에서 왼쪽 카메라의 영상  $I_t^l$ 에서 한 특징점  $p_t^l$ 이 대응하는 오른쪽 카메라의 영상  $I_t^r$ 의 특징점  $p_t^r$ 에 올바르게 대응되는가를 판단할 수 있다.

영상에 잡음이나 양자화(quantization)의 효과로 인하여 영상특징점의 위치에 오차가 발생할 수 있다. 즉 기준시간  $t$ 에서 영상특징점  $p_t^l, p_t^r$ 의 위치에 오차가 발생하게 되고, 이로 인하여 식(1)로부터 구하여 지는 거리에 오차가 발생한다. 거리의 오차는 다중 스테레오영상에 대해 식(3)에 의해 가정되는 영상특징점의 위치에 오차가 발생하게 되어 실제로는 가정된 위치  $p_t^{l'}, p_t^{r'}$ 에 영상특징점이 정확히 위치할 수 없다. 이러한 이유로 인하여 가정된 특징점을 확인할 때 가정점 주위의 적절한 윈도우(window)에 대하여 탐색할 필요가 있다.

다중 스테레오 매칭에 사용되는 스테레오 쌍의 개수가 많을 수록 기준 스테레오영상에 대하여 잘못된 대응관계를 더 줄일 수 있다. 실험에서 실제 영상에 대하여 잘못된 대응관계가 줄어드는 결과를 보인다.

### 3. 거리계산오차 해석

이 장에서는 영상에 잡음이 존재할 때 다중 스테레오 매칭방법으로 구한 거리의 오차를 기존의 스테레오와 비교하여 해석한다. 영상에 잡음이 존재할 때 특징점의 위치에 대한 불확실성은 영상잡음의 분산으로 나타낼 수 있다.<sup>(8)</sup> 영상잡음에 의한 영향으로 영상특징점의 위치에 다음과 같은 불확실성이 존재한다고 하자.

$$\hat{x}_t^i = x_t^i + n_t^i, \hat{x}_t^r = x_t^r + n_t^r \quad (4)$$

$n_t^l, n_t^r$ 는 서로 독립적인 가우시안 백색잡음이라고 하고 그 분산을 각각  $Var(n_t^l), Var(n_t^r)$ 라 하자. 측정된 영상점거리(disparity)  $\hat{d}^t(=\hat{x}_t^l - \hat{x}_t^r)$ 로 계산한 거리를  $\hat{Z}_t^i$ 라 할 때 식(1)에 의한 정확한 거리와의 오차를 구하면

$$\begin{aligned} e_t^z &= Z_t^i - \hat{Z}_t^i = \frac{Bf}{d^t} - \frac{Bf}{\hat{d}^t} \\ &= Bf \left( \frac{1}{x_t^l - x_t^r} - \frac{1}{\hat{x}_t^l - \hat{x}_t^r} \right) \\ &= Bf \frac{n_t^l - n_t^r}{(x_t^l - x_t^r)^2 + (x_t^l - x_t^r)(n_t^l - n_t^r)} \end{aligned}$$

$$\begin{aligned}
&\approx Bf \frac{n_i^t - n_i^t}{(x_i^t - x_i^t)^2} \\
&= Bf \frac{n_i^t - n_i^t}{(d^t)^2} \\
&= \frac{(Z_i^t)^2}{Bf} (n_i^t - n_i^t) \text{ if } |n_i^t - n_i^t| \ll |x_i^t - x_i^t| \quad (5)
\end{aligned}$$

과 같다. 거리오차의 기대값  $E[e_z^t]$ 과 분산  $Var[e_z^t]$ 은 식(5)에서 다음과 같이 구할 수 있다.

$$E[e_z^t] = 0 \quad (6.1)$$

$$Var[e_z^t] = \frac{(Z_i^t)^4}{(Bf)^2} (Var(n_i^t) + Var(n_i^t)) \quad (6.2)$$

식(4)에서 거리가 일정하고 계단에지일 경우 특징점 위치의 불확실성  $Var(n_i^t)$ ,  $Var(n_i^t)$ 의 값은 같게 되며 그 값을  $\sigma_x^2$ 으로 나타내면 식(6.2)은 다음과 같다.

$$Var[e_z^t] = 2 \frac{(Z_i^t)^4}{(Bf)^2} \sigma_x^2 \quad (6.2)'$$

위의 식(6.2)에서 보면 계산한 거리의 불확실성은 영상잡음의 불확실성에 비례하며 거리에 따라 커지고, 카메라간격(baseline)이 커질수록 작아짐을 알 수 있다. 거리오차의 불확실성을 줄이기 위하여 카메라의 간격을 크게 하면 영상점거리(disparity)의 탐색범위가 넓어지기 때문에 정합은 더욱 어려워지게 되고 결국 잘못된 대응점이 발생할 가능성은 증가한다.

다음으로 시간  $t'$ 에서의 스테레오 쌍에 대한 거리오차를 위와 같은 방법으로 구해보면

$$E[e_z^{t'}] = 0 \quad (7.1)$$

$$Var[e_z^{t'}] = 2 \frac{(Z_i^{t'})^4}{(Bf)^2} \sigma_x^2 = 2 \frac{(Z_i^{t'} + T_z^{t-i,t'})^4}{(Bf)^2} \sigma_x^2 \quad (7.2)$$

위의 식에서 볼 때 시간  $t'$ 에서의 스테레오 쌍에 대한 거리오차의 불확실도는 시간  $t$ 에서의 거리오차의 불확실도보다 크다는 것을 알 수 있으며, 이것은 카메라와 물체와의 거리가 더 크기 때문이다. 이상에서 볼 때 카메라가 광축방향으로 전진함에 따라 거리오차는 줄어드는 것을 알 수 있다. 본 논문에서는 한 쌍의 스테레오영상에 대하여 시간적으로 독립적으로 구하지 않고 운동전의 다중 스테레오영상들에서 구한 거리와 현재 기준 스테레오 영상에서 구한 거리의 평균값을 취하는 방법을 사용한다. 즉

$$\hat{Z}_i^{t,N} = \frac{1}{N} \sum_{i=0}^{N-1} \left( \frac{Bf}{d^{t-i}} - T_z^{t-i,t} \right) \quad (8)$$

의 방법으로 현재 기준시간  $t$ 에서의 거리를 추정한다.  $N$ 은 다중 스테레오 매칭에 사용된 스테레오영상 쌍의 개수이고  $T_z^{t-i,t}$ 은 시간  $t-i$ 에서 시간  $t$  사이에 광축방향의 운동이다. 식(8)로 구한 거리의 오차에 대한 불확실도를 계산하자.

$$\begin{aligned}
e_z^{t,N} &= Z_i^{t,N} - \hat{Z}_i^{t,N} \\
&= \frac{1}{N} \sum_{i=0}^{N-1} \left( \frac{Bf}{d^{t-i}} - \frac{Bf}{\hat{d}^{t-i}} \right) \\
&= \frac{1}{N} \sum_{i=0}^{N-1} \frac{(Z_i^t + T_z^{t-i,t})^2}{Bf} (n_i^{t-i} - n_i^{t-i}) \quad (9)
\end{aligned}$$

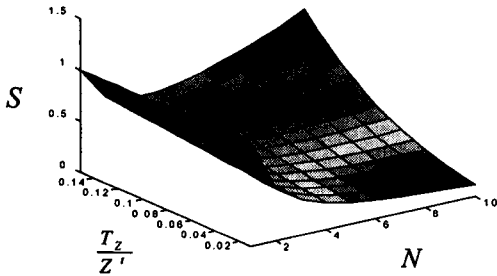
시간  $t$ 에서 시간  $(N-1)$ 사이에서 영상특징점의 위치오차  $n_i^{t-i}$ ,  $n_i^{t-i}$  ( $i=0 \dots N-1$ )는 서로 독립적이므로 식(9)으로 부터 거리오차를 구하면

$$\begin{aligned}
Var[e_z^{t,N}] &= \frac{1}{N} \sum_{i=0}^{N-1} Var[e_z^{t-i}] \\
&= \frac{1}{N^2} \sum_{i=0}^{N-1} 2 \frac{(Z_i^t + T_z^{t-i,t})^4}{(Bf)^2} \sigma_x^2 \\
&= Var[e_z^t] \left[ \frac{1}{N} \sum_{i=0}^{N-1} \left( 1 + \frac{T_z^{t-i,t}}{Z_i^t} \right)^4 \right] \\
&= Var[e_z^t] S \quad (10)
\end{aligned}$$

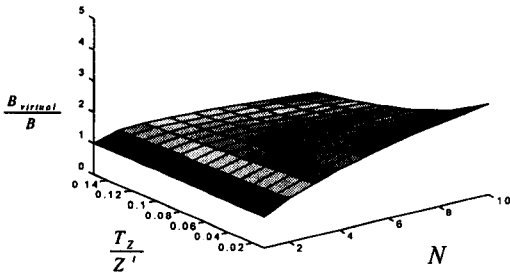
한쌍의 스테레오영상으로부터 구한 거리와  $N$ 쌍의 스테레오영상으로부터 구한 거리의 불확실도의 비를 나타내는  $S (= \frac{1}{N^2} \sum_{i=0}^{N-1} (1 + \frac{T_z^{t-i,t}}{Z_i^t})^4)$ 의 값이 1보다 작아지는  $T_z^{t-i,t}$ 의 값을  $N=2$ 일 경우에 대하여 구하여 보면  $\frac{T_z^{t-i,t}}{Z_i^t} < \sqrt[4]{3}$ 이다. 스테레오에 의해서 구할 수 있는 오차의 정확도는 카메라간격(baseline)의 제곱에 반비례하므로, 한쌍의 스테레오의 경우에 대한 거리의 불확실도와 같은 거리 불확실도를 가지는 다중 스테레오의 가상적인 카메라간격(baseline)을 구해보자. 식(6.2)과 식(10)으로부터 다중 스테레오가 단일 영상스테레오(single view stereo)에 상당하는 카메라간격은

$$B_{virtual} = \frac{B}{\sqrt{S}} \quad (11)$$

으로 구해지며 카메라의 광축방향의 운동피라미터  $T_z^{t-i,t}$ 의 값이  $S < 1$ 의 조건을 만족하면 카메라간격의 값을 늘리는 효과를 가진다. 다시 말하면 단일 영상 스테레오에 비하여 작은 카메라간격(baseline)을 가지고도 카메라의 운동에 의한 다중 스테레오영상을 매칭에 사용함으로써 같은 거리 정밀도를 낼 수 있다. 작은 카메라간격(baseline)을 가지기 때문에 탐색해야 할 영상점 거리(disparity)의 범위가 작아지기 때문에 스테레오 정합이



(a) Ratio of variance of depth error



(b) Ratio of virtual baseline

**Fig. 2** Comparison between multiple-view stereo ( $N > 1$ ) and single-view stereo ( $N = 1$ )

더욱 쉬워지며 잘못된 대응점의 개수를 줄일 수 있다.

Fig. 2(a)에 식 (10)의 다중 스테레오와 단일 스테레오 매칭에 의한 거리오차의 분산의 비율  $S$ 를 도시하였다. 카메라의 운동파라미터의 값은 각 단계에서 일정한 값을 가진다고 할 때, 사용되는 스테레오 쌍이 개수  $N$ 의 값이 클수록 거리오차의 분산은 작아짐을 알 수 있다. 또한 거리에 비하여 카메라의 운동파라미터의 값이 너무 크면 오히려 거리오차의 불확실도는 증가하므로 적절한 카메라의 운동량을 선정할 필요가 있다. Fig. 2(b)는 위의 결과로부터 가상적인 단일 스테레오에 대한 카메라 간격 (baseline)을 카메라 파라미터와  $N$ 에 대하여 도시하였다.

### 5. 실험

이 장에서는 일정한 거리를 가지는 평면 물체에 대하여, 다중 스테레오를 사용함으로써 잘못된 대응점의 개수가 줄어드는 과정을 보이고 다중 스테레오 매칭과 한쌍의 스테레오 매칭으로 얻을 수 있는 거리의 정밀도를 구하여 비교한다. Fig. 3은 사

용되는 기준 스테레오영상과 특징점으로서 에지를 구한 결과이다. 에지는 영상밝기의 2차 미분량의 영점을 통과하는 점으로 화소단위이하(subpixel)의 해상도로 구하였다. 영상내의 물체는 거리가 2m로 일정하다. Table 1에 왼쪽영상에서 에지의 방향, 크기, 밝기의 조건으로부터 대응하고자 하는 특징점을 결정하고, 세가지의 영상특징의 유사성 조건에 대한 오른쪽 영상에서 후보 대응점의 개수를 나타내었다. 모든 경우에 대하여 에지의 판별조건은 동일하게, 구배 (gradient)의 각도의 최대절대값은 15(deg), 크기의 최소값은 15(밝기변화량/pixel)으로 하였다. 또한 오른쪽 영상의 특징점과 대응되기 위한 유사성의 척도중 밝기 구배 (gradient)의 각도의 차이의 허용 최대절대값은 90(deg)로 일정하며, 구배의 크기의 차이와 밝기차이의 허용 최대값을 변화시켜 가면서 세가지의 조건에 대하여 실험하였다.

#### 5.1 밝기구배차의 허용최대값=5, 밝기차의 허용최대값=50일 경우

$N=1$ 일 경우, 즉 기존의 스테레오 매칭의 경우 왼쪽 영상의 한 특징점에 대한 오른쪽 영상의 평균 후보 대응점의 개수는 4.34개이다. 카메라의 운동을 하면서 스테레오영상의 개수 ( $N$ )를 늘려감에 따라 평균 후보 대응점의 개수는 2.23 ( $N=2$ 일 때), 1.63 ( $N=3$ 일 때), 1.41 ( $N=4$ 일 때)로 점점 줄어든다.

#### 5.2 밝기구배차의 허용최대값=5, 밝기차의 허용최대값=50일 경우

밝기차의 허용최대값을 100으로 하여 1의 경우의 50에 비하여 큰 값이기 때문에 후보 대응점의 개수가 증가할 것으로 기대할 수 있다.  $N=1$ 일 경우, 즉 기존의 스테레오 매칭의 경우 왼쪽 영상의 한 특징점에 대한 오른쪽 영상의 평균 후보 대응점의 개수는 5.05개이다. 카메라가 운동을 하면서 스테레오영상의 개수 ( $N$ )를 늘려감에 따라 평균 후보 대응점의 개수는 2.44 ( $N=2$ 일 때), 1.72 ( $N=3$ 일 때), 1.47 ( $N=4$ 일 때)로 점점 줄어든다. 초기 ( $N=1$ )에 후보 대응점의 개수는 5.05로 1의 경우의 4.34에 비하여 많은 차이를 보이지만 스테레오영상의 수 ( $N$ )를 늘려감에 따라 그 차이는 현격하게 줄어들었다.

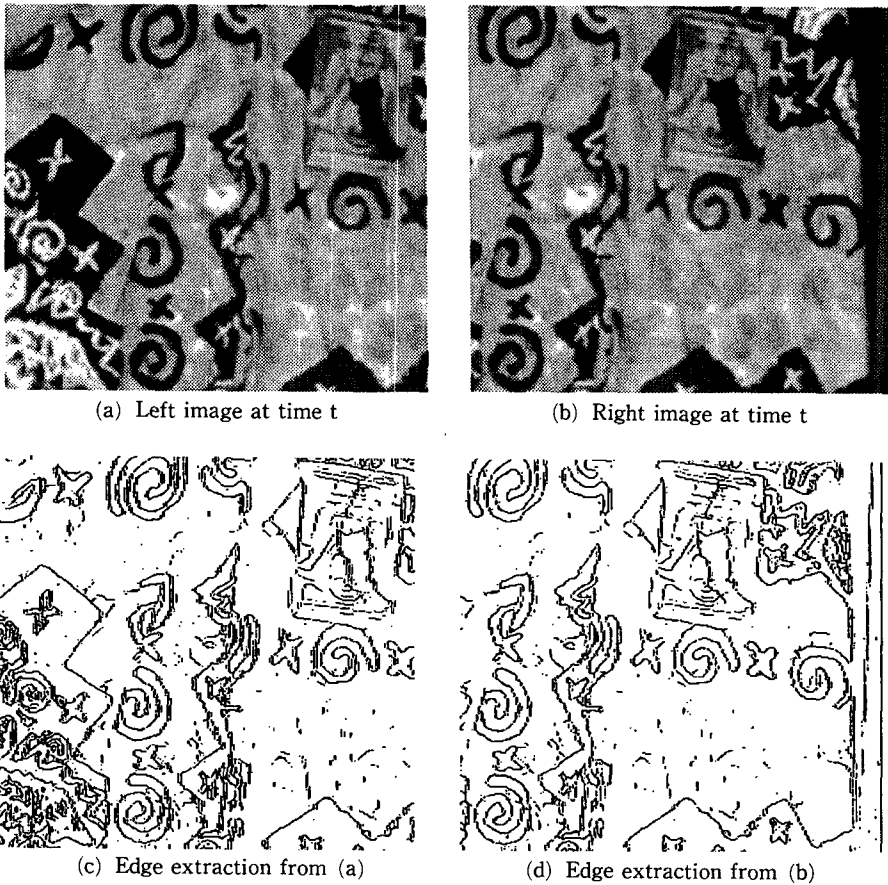


Fig. 3 Stereo images used by multiple-view stereo and edge extraction at time  $t=0$

Table 1 Average number of match candidates in single-view stereo ( $N=1$ ) and multiple-view stereo ( $N=2, 3, 4$ )

Edge condition in left image		Tolerance of feature similarity measure			No. of candidate w. r. t. used stereo pairs( $N$ )			
$ ga _{max}$	$gm _{min}$	$ga(deg)$	$gm$	$I$	1	2	3	4
15	15	90	5	50	4.34	2.23	1.63	1.41
15	15	90	5	100	5.05	2.44	1.72	1.47
15	15	90	10	50	6.76	3.21	2.13	1.76

$ga$  : gradient angle,  $gm$  : gradient magnitude,  $I$  : Intensity of edge

5.3 밝기구배차의 허용최대값=10, 밝기차의 허용최대값=50일 경우

밝기구배차의 허용최대값을 10으로 하여 1의 경우의 5에 비하여 큰 값이기 때문에 후보 대응점의 개수가 증가할 것으로 기대할 수 있다.  $N=1$ 일 경

우, 즉 기존의 스테레오 매칭의 경우 왼쪽 영상의 한 특징점에 대한 오른쪽 영상의 평균 후보 대응점의 개수는 6.76개이다. 카메라가 운동을 하면서 스테레오영상의 개수( $N$ )를 늘려감에 따라 평균 후보 대응점의 개수는 3.21( $N=2$ 일 때), 1.76( $N=4$ 일

**Table 2** Variance of depth error in single-view stereo( $N=1$ ) and multiple-view stereo ( $N=2, 3, 4$ )

No. of stereo pairs( $N$ )	Variance of depth error
1	122.0
2	96.3
3	82.9
4	60.8

때)로 점점 줄어든다. 초기( $N=1$ )에 후보 대응점의 개수는 6.76개로 1)의 경우의 4.34에 비하여 많은 차이를 보이지만 스테레오영상의 수( $N$ )를 늘려감에 따라 그 차이는 현격하게 줄어들었다.

위의 실험에서 후보 대응점의 개수는 스테레오영상의 개수를 증가함에 따라 줄어들음을 알 수 있다. 또한 영상잡음의 영향을 고려하기 위하여 영상특징점의 대응에 대한 유사성의 조건을 느슨하게 하였을 경우에 대하여도, 후보 대응점의 개수는  $N$ 의 값을 따라 현격하게 감소하였다. 후보 대응점의 개수가 감소할 수록 최소 대응오차를 가지는 대응점을 선택할 확률이 커지므로 잘못된 대응점의 수를 줄일 수 있다.

다음으로 기준 스테레오영상의 전체에 대하여 구한 거리오차의 분산을 Table 2에 스테레오 매칭에 사용한 스테레오영상의 쌍의 개수에 대하여 나타내었다. Table 2의 거리오차의 분산은 영상에 분산이 16인 가우시안 백색잡음을 가한 결과이다.  $N=4$ 인 다중 스테레오에서 단일 스테레오( $N=1$ )의 경우에 비하여 그 값이 절반가량으로 줄어든다. 이상의 실험을 통하여 제안한 스테레오 매칭방법은 거리측정의 정밀도와 잘못된 대응점을 제거하는 정확도가 단일 스테레오에 비하여 향상됨을 보였다.

## 6. 결 론

본 논문에서는 이동로봇과 같이 광축방향으로 운동하는 스테레오카메라의 경우에 대하여 다중 스테레오 쌍을 이용하여 거리를 구하는 방법을 제안하였다. 이 방법은 기존의 이동스테레오(motion stereo)방법과는 달리 관측자의 운동방향과 일치하는 카메라운동으로부터 얻은 다중 스테레오영상을 사용한다. 또한 카메라의 운동이 커서 시간적인 영

상운동량(optical flow)을 얻기 힘든 상황을 가정하고 있다. 다중 스테레오 매칭을 통하여 거리를 구하기 위해서 각 시간대에서 얻은 거리의 평균을 취하였으며, 이렇게 구한 거리의 오차에 대한 불확실도(분산)을 구하여 단일 스테레오의 경우와 비교하였다. 또한 단일 스테레오(single view stereo)에 비하여 작은 카메라간격(baseline)을 가지고도 카메라의 운동에 의한 다중 스테레오영상을 매칭에 사용함으로써 같은 거리정밀도를 낼 수 있음을 해석적으로 보이고 실험으로 입증하였다. 결과적으로 작은 카메라간격을 가지기 때문에 탐색해야 할 영삼점거리의 범위가 작아지기 때문에 스테레오 정합이 더욱 쉬워지며 잘못된 대응점의 개수를 줄일 수 있음을 실험으로 보였다. 단일 스테레오의 경우 왼쪽 영상의 한점에 대한 후보 대응특징점의 개수는 4.34개였으나, 카메라가 운동을 함에 따라  $N=4$ 일 때 1.41개로 상당히 줄어들었다. 후보 대응점의 개수가 감소할 수록 최소 대응오차를 가지는 대응점을 선택할 확률이 커지므로 잘못된 대응점의 수를 줄일 수 있다.

## 참고문헌

- (1) Stephen T. Barnard and Martin A. Fischler, 1982 "Computational Stereo," *Computing Survey*, Dec., Vol. 14, No. 4, pp. 553~572.
- (2) Umesh R. Dhond and Aggarwal, J. K. 1989 "Structure from Kstereo:a Review," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-19, Nov., No. 6 pp. 148~1510.
- (3) Grimson, W. E. L., 1985 "Computational Experiment With a Feature-based Stereo Alogorithm," 1985, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Jan., Vol. PAMI-7, No. 1, pp 17~34.
- (4) Marr, D. and Poggio, T., 1979, "A Theory of Human Stereo Vision," *Proc. Roy. Soc. (London)*, pp. 301~328, Vol. B
- (5) Masatoshi Okutomi, and Takeo Kande, 1993 "A Multiple-Baseline Stereo," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-15, Apr., No. 4, pp. 353~363.
- (6) Bolles R. C., Baker H. H. and Marimont D. H., 1987 "Epipolar-Plane Image Analysis: an

- Approach to Determining Structure from Motion," *Int. J. Computer Vision*, Vol. 1, No. 1.
- (7) Matthies, L., Szeliske, R. and Kanade, T., 1989, "Kalman Filter-Based Algorithms for Estimation Depth from Image Sequences," *Int. J. Computer Vision*, Vol. 3, pp. 209~236.
- (8) Canny J., 1986 "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Nov., Vol. PAMI-8, No. 6 pp. 679~698.