

스테레오 카메라의 미소 병진운동을 이용한 3차원 거리추출 알고리즘

박 광 일,* 이 재 웅,* 오 준 호*

3-D Range Finding Algorithm Using Small Translational Movement of Stereo Camera

Kwang-Il Park,* Jae-Woong Yi,* Jun-Ho Oh*

ABSTRACT

In this paper, we propose a 3-D range finding method for situation that stereo camera has small translational motion. Binocular stereo generally tends to produce stereo correspondence errors and needs huge amount of computation. The former drawback is because the additional constraints to regularize the correspondence problem are not always true for every scene. The latter drawback is because they use either correlation or optimization to find correct disparity. We present a method which overcomes these drawbacks by moving the stereo camera actively. The method utilizes a motion parallax acquired by a monocular motion stereo to restrict the search range of binocular disparity. Using only the uniqueness of disparity makes it possible to find reliable binocular disparity. Experimental results with real scene are presented to demonstrate the effectiveness of this method.

Key words : stereo matching(스테레오 매칭), motion stereo(모션 스테레오), 3-D structure estimation(3차원 거리 추정), correspondence problem(대응문제)

1. 서 론

이동 로봇의 경로계획이나 장애물 회피등의 작업을 할 경우 관측자로부터 주위 물체들 간의 거리를 알아내는 것은 필수적이다. Laser나 초음파등과 같은 능동장치를 사용하지 않고 CCD 카메라만을 사용하여 얻은 영상으로부터 거리를 추출하는 방법에는 스테레오 매칭과 모션 스테레오(motion stereo) 두가지가 있다. 스테레오

매칭은 상대적으로 고정된 두대의 카메라로부터 독립된 두장의 영상을 받아들인 후 영상내의 각점을 서로 일치시켜 카메라의 상대적인 기하학적관계로부터 삼각법에 의해 거리를 구한다.⁽¹⁻⁵⁾ 모션 스테레오(motion stereo) 방법은 현대의 카메라가 움직이며 연속적으로 영상을 받아들여 각 점의 변위를 구하여 변위에 반비례하는 거리를 구하는 방법이다.⁽⁶⁻¹²⁾ 관찰하는 물체가 고정되어 있을 경우 스테레오 매칭은 모션 스테레오에 포

* 한국과학기술원 기계공학과

함될 수 있지만, 일반적으로 모션 스테레오는 미세변위 운동을, 스테레오 매칭은 상대적으로 큰 변위의 운동을 다룬다고 생각할 수 있다.

스테레오 매칭과 모션 스테레오에서 가장 중요하고 어려운 문제는 한 영상에서의 한점이 다른 영상에서 어디에 해당하는가를 찾아내는 대응문제 (correspondence problem)이다.⁽⁴⁾ 이 대응문제가 해결되면 카메라사이의 기하학적인 관계와 대응점들의 상대위치등을 사용하여 거리는 쉽게 구해진다. 스테레오 비전에서는 일반적으로 각각의 영상으로부터 특징을 추출하여 이 특징을 정합하여 대응문제를 해결한다. 영상에서 특징점을 추출하는 과정은 대개 경험적인 조건에 의해 구별되어지는 경우가 많으며,⁽²⁻⁴⁾ 탐색범위내에 유사한 특징이 많은 경우 유일한 매칭을 찾아내기가 어렵다. 한편 모션스테레오 방법은 카메라의 운동이 작다는 가정하에 국부적인 영상의 밝기 변화로부터 영상점의 변위를 계산함으로써 대응문제를 해결한다.^(6,9,10) 그러나 카메라의 미소운동을 가정하기 때문에 변위의 계산은 간단하게 되지만, 대응의 결과로부터 구한 물체까지의 거리는 스테레오 매칭에 비하여 더 큰 오차를 가지게 된다.

영상면에서 발생하는 영상변위 (optical flow)와 카메라 운동식으로부터 거리를 구하는 모션 스테레오에 관한 연구는 다음과 같다. Tistarelli는 영상의 경계 (contour)부분에서의 운동을 고려한 알고리즘을 내놓았으며,⁽⁶⁾ 모션 스테레오 (motion stereo)에서 구한 영상변위 (optical flow)가 많은 오차를 갖고 있으므로 이를 적분하여 깊이를 측정하는데 이때 칼만필터 (Kalman filter)를 이용하여 오차를 줄이려는 노력등이 있었다.⁽¹³⁾ 또한 운동의 시간축을 이용하여 한 에피폴라 라인 (epipolar line)을 시간축으로 영상의 평면을 만들어 이 평면 상의 에지를 구하여 정확한 운동을 구하려는 노력이 Marimont⁽⁸⁾에 의해 행하여졌으며, Kanade⁽⁷⁾는 스테레오 카메라의 간격 (baseline)을 바꾸어 가면서 각 카메라 간격 (baseline)에 대한 영상밝기값의 오차를 합산하여 대응문제를 해결하였다. 한편 영상변위 (optical flow)와 스테레오 매칭 (stereo matching)을 결합하려는 노력들이 있어왔는데 3차원의 병진운동에서 영상변위의 여러 제한 조건식과 스테레오 매칭의 식들을 조합하여 깊이 정보를 알아내려는 Miyazaki와 Arimoto에 의해 소개된적이 있다.⁽¹⁵⁾ 또한 이재웅 등은 다중스테레오 영상을 이용하여 스테레오 대응의 가정, 확인 과정을 통하여 대응문제를 해결하였다.⁽¹⁶⁾

본 연구에서는 스테레오 매칭과 모션 스테레오 방법의 장점을 이용하여 물체까지의 거리 정보를 효율적으로 구하는 방법을 제안한다. 본 연구에서 상정하는 상황은 고정된 주위환경에서 스테레오 카메라를 가진 관측자가 평행운동을 하는 경우이다. 운동 전후에 2쌍의 스테레오 영상을 얻을 수 있으며 왼쪽 영상간의 영상변위 (optical flow), 오른쪽 영상간의 영상변위 (optical flow), 그리고 왼쪽과 오른쪽 영상간의 스테레오 매칭으로부터 거리를 구한다. 스테레오 매칭에서 거리와 왼쪽과 오른쪽 영상에서 대응 거리 (disparity)의 관계는 다음과 같다.

$$d = \frac{Bf}{Z}$$

B 와 f 는 각각 스테레오 카메라의 간격 (baseline)과 초점거리이다. 거리 Z 는 대응거리 (disparity) d 에 반비례하므로 정밀한 거리추정을 위해서는 카메라의 간격을 크게하여 대응거리의 값을 크게 하여야 한다. 하지만 이럴 경우 거리 추정의 정밀도는 높아지지만 스테레오 매칭의 탐색범위는 넓어 지므로 정확한 대응을 찾기가 어렵다. 즉 거리 추정의 정밀도와 정확도사이에는 상충관계가 존재한다.⁽⁷⁾ 한편 물체까지의 거리는 영상변위와 카메라 운동식으로부터도 구할 수 있지만 카메라의 운동량이 작기 때문에 구한 거리는 많은 오차를 가지고 있다. 하지만 왼쪽 영상의 한 특징점에 대하여 모션스테레오에 의하여 구한 거리와 그 계산오차로부터 오른쪽 영상에서의 스테레오 매칭의 탐색범위를 결정할 수 있다. 결과적으로 스테레오 매칭에 의하여 정밀한 거리 추정을 행할 수 있다. 이것이 본 논문에서 제안하는 알고리즘의 기본 생각이다.

2장에서는 본 논문에서 사용하는 카메라 모델과 운동 방정식에 대하여 기술하고, 3장에서는 영상변위 계산오차로부터 스테레오 매칭의 탐색범위를 구하는 방법을 제안한다. 4장에서는 실제 영상을 가지고 실험을 행하여 제안한 알고리즘의 효용성을 입증한다. 5장에서는 결론을 보인다.

2. 카메라 모델과 운동방정식

이 장에서는 카메라의 3차원 운동과 영상면에서의 영상변위 (optical flow)의 관계를 보이고 영상변위 (optical flow)의 추정으로부터 거리를 구하는 모션스테레오 (motion stereo)의 방법을 개략적으로 설명한다. 또한

평행카메라의 경우에 대하여 스테레오매칭에 의하여 거리를 구하는 방법을 보인다. 연속된 영상간의 카메라의 운동이 작다고 가정하면 결과적으로 구해지는 영상변위(optical flow)는 카메라의 순간속도의 향으로 근사시킬 수 있다. 카메라의 속도는 병진속도 $\vec{T} = [T_x T_y T_z]^T$ 와 회전속도 $\vec{R} = [R_x R_y R_z]^T$ 로 나눌 수 있다. 카메라 좌표계에서 3차원 공간 상의 한점 $P(X, Y, Z)$ 의 운동은 다음과 같다.

$$\frac{dP}{dt} = -\vec{T} - \vec{R} \times P \quad (1)$$

$P(X, Y, Z)$ 를 촛점거리가 1인 영상면에 원근 투영하면 그 투영점 $p(x, y)$ 은 다음과 같다.

$$x = \frac{X}{Z}, y = \frac{Y}{Z} \quad (2)$$

영상면의 좌표 x 와 y 는 각각 3차원 좌표 X 와 Y 에 평행한다.

식 (1), (2)으로부터 다음과 같은 영상변위(optical flow) 방정식을 얻을 수 있다.

$$\begin{aligned} \begin{bmatrix} u \\ v \end{bmatrix} &= \frac{1}{Z} \begin{bmatrix} -1 & 0 & x \\ 0 & -1 & y \end{bmatrix} \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \\ &+ \begin{bmatrix} xy & -(1+x^2) & y \\ (1+y^2) & -xy & -x \end{bmatrix} \begin{bmatrix} R_x \\ R_y \\ R_z \end{bmatrix} \end{aligned} \quad (3)$$

식 (3)은 거리 Z 를 카메라의 운동 \vec{T}, \vec{R} 그리고, 영상면에서의 영상변위(optical flow) (u, v) 와 연관시킨다. 영상변위 성분 u, v 는 영상변위의 각각 영상면 좌표계 x, y 에 대한 성분이다. 우리는 앞으로 식 (3)을 이용하여, 알고 있는 카메라의 운동에 대하여 영상변위(optical flow)를 구하여 거리를 추정하도록 한다.

한편 평행스테레오 카메라 구조는 회전운동은 없고 X 방향의 운동이 카메라사이의 간격(baseline) B 에 해당한다. 식 (3)에서 x 방향의 영상변위 u 와 구별하여, 왼쪽 영상과 오른쪽 영상과의 대응점의 에피폴라 라인(epipolar line)상의 위치 차이를 d 라고 하면 거리 Z 는 다음과 같이 표현된다.

$$Z = \frac{B}{d} \quad (4)$$

3. 제안한 알고리즘과 오차 해석

3.1 영상 변위(optical flow) 추정

우리는 왼쪽(또는 오른쪽) 카메라 영상간의 카메라의 움직임을 안다는 가정을 하기로 하였다. 또한 카메라 운동이 병진운동만으로 이루어져 있다면 영상간의 영상변위(optical flow)의 방향이 결정된다. 카메라가 광축에 수직인 방향으로 운동을 하는 경우(lateral motion stereo)는 바로 카메라와 같은 방향으로 영상변위(optical flow)의 방향이 결정되며 광축방향으로 카메라가 운동하는 경우(longitudinal motion stereo)는 카메라가 앞으로 전진 또는 후진 할 경우로서 영상에 영상변위(optical flow)가 한 곳으로 모이거나 한 곳에서 나가는 점이 존재하는데 이를 FOE(Focus of Expansion) 또는 FOC(Focus of Contraction)라 한다. FOE는 촛점거리를 f , 카메라의 움직임을 $\vec{T} = [T_x T_y T_z]^T$ 라 할 때 다음과 같다.

$$FOE = \left(f \frac{T_x}{T_z}, f \frac{T_y}{T_z} \right) \quad (5)$$

따라서 카메라가 움직인 방향을 고려하여 영상변위의 방향을 정하여 그 방향에 대하여 영상밝기의 차이의 합(SSD)이 최소가 되는 점을 찾거나,⁽⁷⁾ 에지(edge)의 수직인 방향에 대한 성분을 구하여 식 (5)와 조합하여 실제의 영상변위를 구할 수 있다.⁽¹⁷⁾

3.2 영상변위(optical flow)의 추정 오차

영상변위가 정확하다면 식 (3)으로부터 거리를 구할 수 있다. 하지만 영상변위의 추정은 오차를 가지고 있으며 오차의 원인은 밝기값의 잡음과 영상의 이산화의 영향에 의한 화소(pixel) 위치오차이다. 이 절에서는 위의 두가지의 오차 원인에 대하여 알아 본다.

먼저 밝기값의 잡음에 대하여 고찰하여보자. 본 연구에서는 영상변위를 구하는데 있어서 밝기값의 변화가 급격한 곳 즉 에지(edge)에서만 행하며 에지에서 에지 사이의 영상변위를 구한다. 하지만 이 에지의 위치는 영상잡음이 있는 경우 불확실성을 갖게 되어 영상변위의 오차의 원인이 된다. 에지 위치의 불확실성은 에지의 크기, 방향과 영상잡음이 함께 영향을 주는데 이 에지 위치의 불확실성은 바로 영상변위의 불확실성과 관계가 있다. 평행 스테레오 카메라의 경우 매칭은 에피폴라 라인(epipolar line)에서만 이루어지므로 1차원 영상신

호에 대하여 에지 추출과 에지위치의 분산을 구하기로 한다. 에지의 위치는 식 (6)의 가우시안 연산자의 2차 미분을 영상에 컨볼루션(convolution)한 후 영점통과 지점을 찾음으로써 결정할 수 있다.⁽¹⁸⁾

$$F(x) = \frac{d^2 G(x)}{dx^2} \cdot I(x) \quad (6)$$

$G(x)$ 는 가우시안 함수, $I(x)$ 는 밝기값의 함수, $F(x)$ 는 에지의 함수이다. 이 때 영상 I 에 분산 σ_n^2 의 가우시안 백색 잡음이 섞여 있다고 하자. 위의 연산자에 대한 반응을 신호에 의한 반응 F_s 와 잡음에 의한 반응 F_n 으로 나누면 에지위치의 분산은 다음과 같이 구해진다.

$$E[\delta x^2] = \sigma_e^2 = \frac{\sigma_n^2 \cdot E[(F_n(x_0))^2]}{[F_s(x_0)]^2} \quad (7)$$

실제 구현을 할 때는 $E[(F_n(x_0))^2]$ 는 컨볼루션 마스크(convolution mask)의 계수들의 제곱의 합으로 계산될 수 있으며 $[F_s(x_0)]^2$ 는 영점을 통과하는 기울기로 표현될 수 있다.

영상이 깨끗하여 노이즈가 거의 없는 영상이 있다 하더라도 영상점의 위치 오차는 생기기 마련이다. 이 오차는 영상이 이산화 되어 있어 정량화 오차(quantization error)가 생기기 때문인데 한 화소를 1로 보았을 때 이 오차는 0.5이하로 줄지는 않는다. 영상내의 한 화소의 실제 크기를 Δc 라 하고 처음 위치에서 화소까지의 거리를 B' 라 할 때 평균과 분산은 다음 식과 같다.⁽¹⁹⁾

$$E(\beta) = \frac{B'}{\Delta c} - \frac{1}{2}, \quad V(\beta) = \frac{1}{6} \quad (8)$$

여기에서 β 는 이산 1차 좌표계에서 오른쪽에 붙은 좌표 값이 된다. (8)에서 볼 수 있듯이 이산화된 영상의 화소 좌표는 불확실한 좌표가 된다.

영상변위를 구함에 있어서 위의 영상점 위치 오차를 무시할 수 없으며 이 오차는 우리가 제한하는 스테레오 매칭 알고리즘의 범위를 결정하는 근거가 된다.

3.3 스테레오 매칭의 범위 결정

카메라의 운동이 병진운동일 때, 왼쪽 영상의 한 점 (x, y) 에 대하여 영상변위(optical flow)의 추정결과로부터 물체까지의 거리를 식 (3)으로부터 다음과 같이 구할 수 있다.

$$Z = \frac{-T_x + xT_z}{u} \quad (9)$$

식 (9)에서는 카메라의 운동이 X, Z 의 방향만을 고려하였다. 일반적인 운동에 대하여도 어렵지 않게 확장할 수 있다. 앞절에서의 특징점위치의 오차로 인하여 영상 변위 u 는 Δu 의 오차를 가진다고 하자. 이 오차는 식 (9)에 의하여 거리추출오차를 유발한다. 이 때 거리 범위의 최대값과 최소값을 각각 Z_{\max} 와 Z_{\min} 라고 하면 다음과 같이 구할 수 있다.

$$Z_{\min} = \frac{-T_x + xT_z}{u + \Delta u} \quad (10.1)$$

$$Z_{\max} = \frac{-T_x + xT_z}{u - \Delta u} \quad (10.2)$$

위의 식 (10)과 식 (4)로부터 오른쪽 영상에서의 스테레오 매칭의 범위를 다음과 같이 구할 수 있다.

$$d_{\min} = d - \frac{B}{-T_x + xT_z} \Delta u \quad (11.1)$$

$$d_{\max} = d + \frac{B}{-T_x + xT_z} \Delta u \quad (11.2)$$

위의 식 (11)에서 정확한 영상변위로 부터 구한 영상점 거리(disparity) d 는 다음과 같다.

$$d = \frac{B}{-T_x + xT_z} u \quad (12)$$

따라서 오른쪽 영상에서 스테레오 매칭의 범위 $\Delta d (= d_{\max} - d_{\min})$ 와 영상변위 추정오차 Δu 의 관계는 다음과 같다.

$$\frac{\Delta d}{d} = \frac{2\Delta u}{u} \quad (13)$$

대응범위 Δd 는 카메라 간격(baseline)의 크기와 영상 변위의 오차에 비례한 양이 된다.

카메라의 움직임이 커지면 대응범위는 좁아지게 되며 깊이 방향의 움직임의 경우 광축에서 멀리 떨어져 있는 곳일수록 범위는 작다. 이 범위에서 대응점의 후보가 1개일 경우에는 그 점이 바로 유일한 대응 점이되고 2개 이상일 경우 에지(edge)부분의 값을 찾아 후보를 줄여서 대응점을 찾는다. 이 과정을 Fig. 1에 보였다.

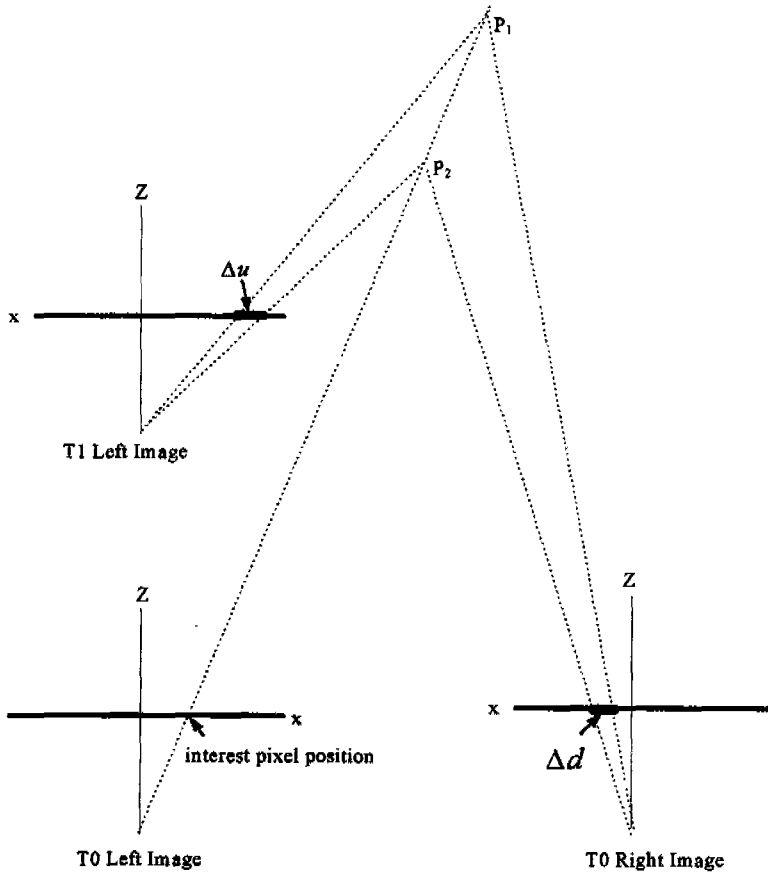


Fig. 1 Stereo matching using optical flow

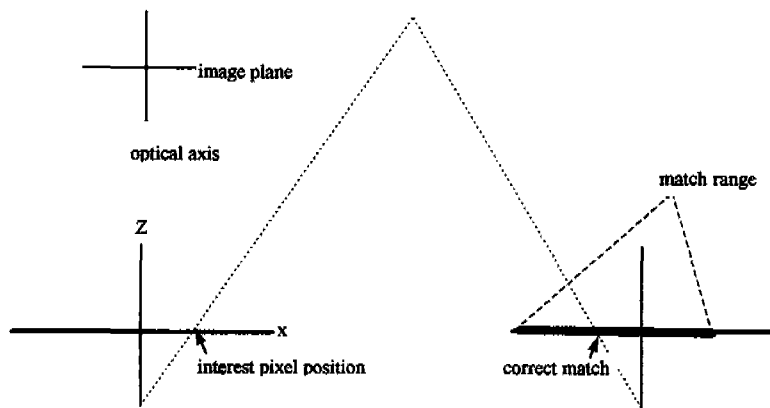
3.4 제안한 알고리즘의 장점과 기존 알고리즘과의 비교

영상변위(optical flow)만을 써서 깊이를 알아내는 모션 스테레오(motion stereo) 방법은 거리 Z 가 영상변위에 반비례한다. 또한 영상변위의 추정을 용이하게 하기 위하여 카메라의 운동을 작게 하므로 영상변위는 작은 값으로 구해진다. 결과적으로 영상변위 추정의 조그만 오차에도 거리는 많은 오차를 갖게 된다. 한편 상대적으로 카메라사이의 변위가 큰 스테레오 매칭에서는 영상점 거리(disparity)가 상대적으로 큰 값을 갖게 되므로 조그만 대응 오차에 대해서 거리 추정은 큰 영향을 받지 않는다. 반면에 스테레오 매칭의 경우 대응관계를 찾아 내기 위해서는 에피폴라 라인(epipolar line)상에 있는 모든 점에 대하여 탐색을 행해야 하므로 계산량이 아주 많아진다. 또한 탐색범위가 크므로 유사한 특징점

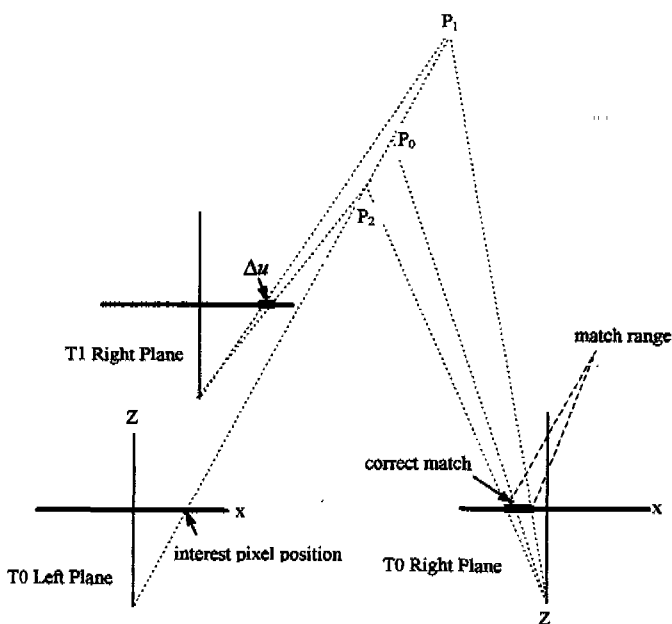
이 존재할 경우 유일한 대응점을 찾아 낼 수 없다. 하지만 본 논문에서 제안한 방법으로 스테레오 매칭의 범위를 구하게 되면 후보 대응점의 탐색범위가 줄어들게 되어 계산량이 작아지는 장점을 가지며, 유사한 특징점이 존재할 확률이 적어지므로 거리추정의 정확성이 높아진다. Fig. 2는 일반적인 스테레오 매칭의 과정과 본 논문에서 제안한 탐색범위의 제한방법에 의하여 스테레오 매칭의 범위가 줄어들어 드는 과정을 보인다.

3.5 두 개의 영상변위(optical flow)를 이용하는 방법

지금까지의 방법은 사실상 세계의 영상에서 모든 작업이 이루어져 왔다. 즉 운동전의 왼쪽, 오른쪽 영상과 운동후의 왼쪽영상을 사용하였다. 하지만 실제로는 네 개의 영상을 가지고 있으므로 나머지 한 영상을 이용하



General stereo matching



Stereo matching by optical flow information

Fig. 2 Comparison with traditional stereo matching

는 방법을 논의해 보자. 카메라가 광축방향의 평행운동을 하고 있을 경우 영상면의 중심근처에서는 영상변위의 값이 아주 작게 계산되면 결과적으로 식 (13)에 의하여 스테레오 매칭의 범위는 아주 큰 값을 가진다. 따라서 영상변위를 이용하여 스테레오 매칭의 범위를 제한하는 효과가 적다. 하지만 왼쪽 영상의 중심부근에 해당하는 영역은 오른쪽 영상에서는 중심에서 상당히

떨어져 있게 되므로 오른쪽 영상에서 거리를 구하여 보상하는 방법을 생각할 수 있다. 스테레오 매칭의 범위 내에 두 개 이상의 대응점이 존재할 때 그 점의 영상변위를 이용하여 잘못된 대응점을 제거하여 보자. 제거하는 과정을 나타낸 그림이 Fig. 3이다.

스테레오 카메라가 초기 시간 T_0 에서 시간 T_1 까지의 병진 운동(T_x, T_z)을 하였다고 하자. 시간 T_0 에서의 왼쪽

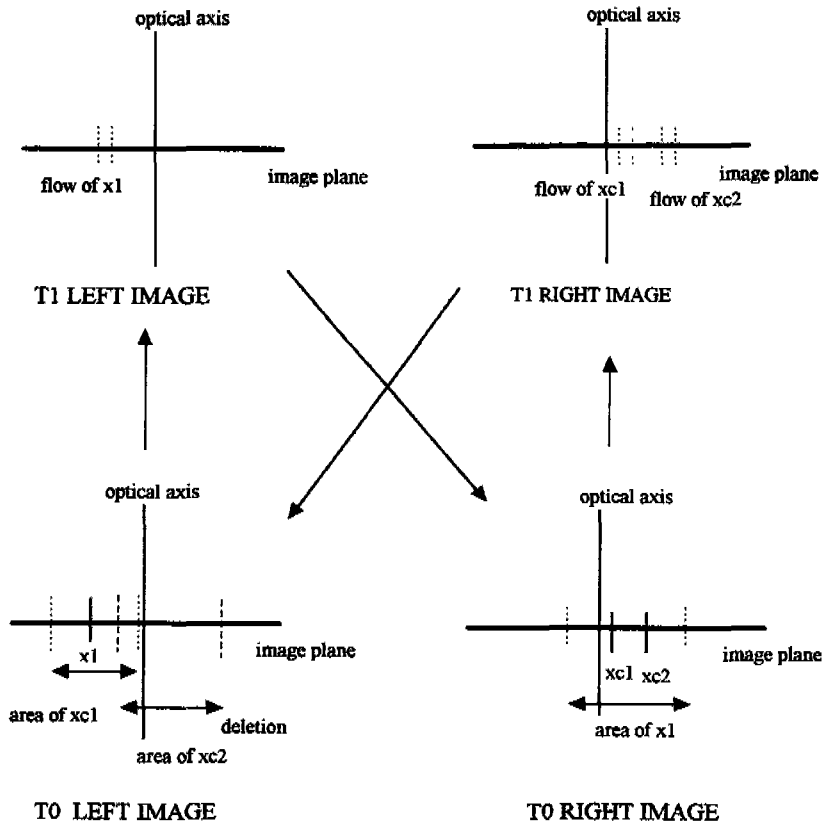


Fig. 3 Stereo matching using 2 pair of stereo images

영상의 한 점 x_{10} 에 대하여 영상변위를 u_1 , 그 오차 범위를 $u_1 \pm \Delta u_1$ 라고 하면 식 (13)에 의하여 오른쪽 영상에서 대응점의 탐색범위를 계산할 수 있다. 후보 대응점이 여러개 존재할 때 그 좌표를 $x_{r0}, x_{r1}, x_{r2}, \dots, x_{rk}, \dots, x_{rn}$ 이라하고, 오른쪽 영상의 이 점들에서 구한 영상변위를 각각 $u_{r0}, u_{r1}, u_{r2}, \dots, u_{rk}, \dots, u_{rn}$ 이라 하자. 오른쪽 영상의 대응 후보점 x_{rk} 에 대하여 왼쪽영상에 대한 유사특징점의 존재 범위는 위의 방법과 동일하게 식 (13)에 의하여 구할 수 있다.

$$d_{lk}|_{\min} = d_{lk} - \frac{B}{-T_x + x_{rk}T_z} \Delta u \quad (14.1)$$

$$d_{lk}|_{\max} = d_{lk} + \frac{B}{-T_x + x_{rk}T_z} \Delta u \quad (14.2)$$

이때 $d_{lk} = \frac{B}{-T_x + x_{rk}T_z} u_{lk}$ 이다.

따라서 오른쪽 영상의 대응 후보점 x_{rk} 에 대하여 $[x_{rk} + d_{lk}|_{\min}, x_{rk} + d_{lk}|_{\max}]$ 사이에 x_{lk} 가 존재하지 않으면 이 점은 올바른 대응점이 아니므로 대응점 후보에서 제거할 수 있다.

4. 실험

영상변위(optical flow)를 구하는데 있어서 윈도우(window)내에서 영상 밝기의 차이의 합(SSD)를 구하여 최소가 되는 점을 대응점으로 함으로써 영상 잡음에 대하여 둔감하게 영상변위를 구하였다. 영상 특징을 예지로 잡았는데 실제영상에 적용할 경우에는 적당한 크기의 윈도우(window)를 선택한 범위에서 블럭(block) 매칭을 하였다. 따라서 실제 영상에 우리가 적용할 때 쓰이게 되는 제한 조건은 다음과 같이 두가지이다.

조건 1 : 영상의 에지(edge)에서만 매칭을 한다. 따라서 에지는 대응되어야 하고 그 크기와 방향도 같아야

한다. 에지를 찾는 방법으로는 소벨 연산자(Sobel operator)를 썼으며 밝기 구배(gradient)의 크기를 gm, 방향(phase)를 gp라고 한다면 gm의 값이 어떤 값 이상일 때 에지로 판단하므로 후보 대응점의 밝기구배 크기 Can_gm , 방향 Can_gp 는 다음과 같다.

$$Can_gm > LIMT_gm, \\ gp - \Delta gp < Can_gp < gp + \Delta gp$$

조건 2 : 일정한 범위에서 상관관계(correlation)가 가장 큰 점을 영상변위로 한다. 이때 블럭 매칭(block matching)을 하였으므로 오차율이 어느값보다 작은 경우에만 해당한다. 즉 윈도우의 폭을 Win_X , 높이를 Win_Y , 한 화소의 밝기변화를 $Inten$ 라 하면 블럭 매칭의 오차는 다음의 조건을 만족한다.

$$Error < Win_X \cdot Win_Y \cdot Inten$$

여기에서 밝기변화 Inten은 한 점이 다른 영상에 대응될 경우에 같은 밝기값이 아니고 약간의 변화(variance)를 가지고 대응되기 때문이며 또한 한 점이 바로 다른 영상의 한 점으로 대응되는 것이 아니고 점 과 점 사이(subpixel)로 대응되기 때문에 생기는 오차가 있기 때문에 고려하였다.

다음은 스테레오 매칭(stereo matching)을 할 때, 탐색범위에서 대응점 후보를 정하는 방법인데 이 때 대

응점을 찾는 방법으로는 영상변위를 찾는 방법과 동일한 방법으로 1) 에지의 밝기구배의 크기와 방향이 일정 범위에 있어야 하며 2) 블럭 매칭(block matching)의 오차가 조건 2를 만족하는지가 제한 조건이 된다. 이때 대응 후보점을 정하는 방법으로는 조건 2를 만족하는 범위내의 점을 모두 대응 후보로 하였다.

4. 1 평행 스테레오 운동(parallel motion stereo)의 경우

먼저 행한 실험은 스테레오 카메라를 X방향(수평방향)으로 이동하여 구한 영상변위를 가지고 범위를 정하여 스테레오 매칭을 하였다. 영상변위는 영상의 에지(edge)부분에서만 계산하였다. 카메라 간격(baseline)은 50.8(mm), $T_x=6.35$ (mm)로 하였다. 물체까지의 평균거리는 2000(mm)이다. Fig. 4는 사용한 4개의 영상의 그림이며 Fig. 5는 왼쪽 영상에 대하여 영상변위를 구한 결과를 도시하였다. Fig. 6은 영상변위만으로부터 거리를 구한 결과이다. 거리를 추출 한 결과는 영상전체영역에 걸쳐서 몇 개의 단계를 가지고 있으며 거리 추출에 제대로 이루어지지 않은 부분이 보인다. Fig. 7은 제한한 알고리즘에 의하여 탐색범위를 정한 후 스테레오 매칭을 한 결과를 나타낸다. 모션 스테레오만을 가지고 구한 깊이 정보는 대강의 깊이정보를 나타내고 있으며 다시 스테레오 매칭을 한 결과는 상대적

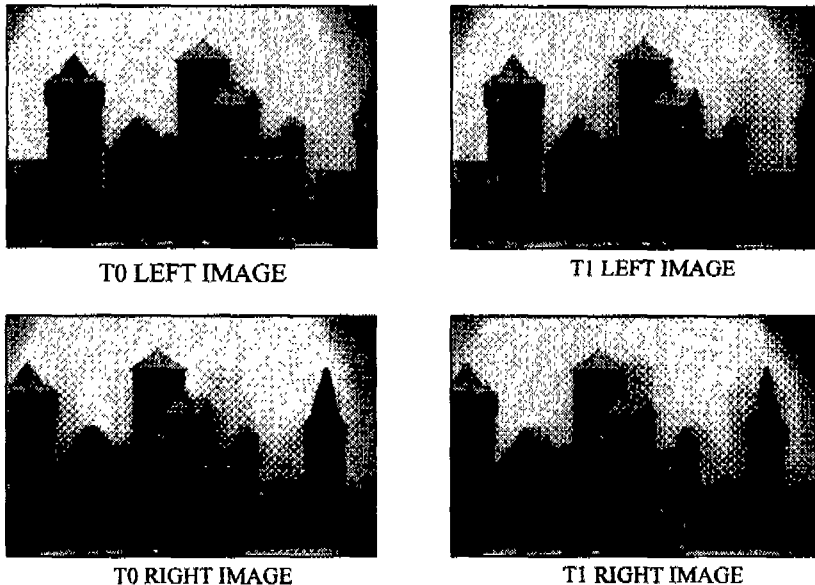


Fig. 4 stereo images at time 0 and 1(parallel motion)

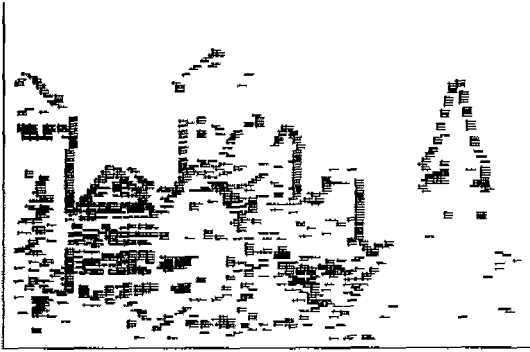


Fig. 5 result of optical flow computation

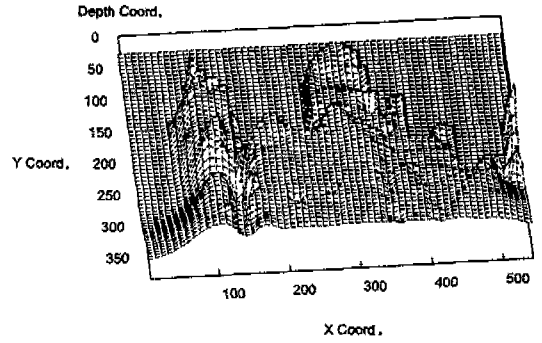


Fig. 7 Depth from stereo matching and optical flow

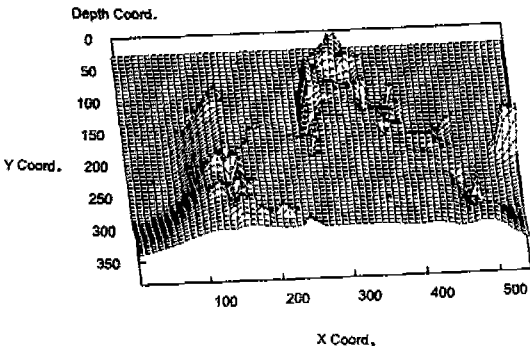
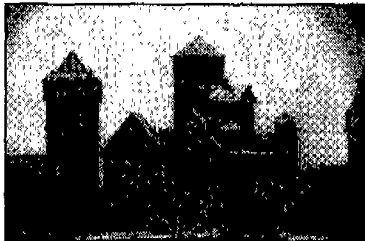


Fig. 6 Depth from optical flow(Fig. 5)

으로 다양한 깊이정보를 갖고 있음을 알 수 있다. 그리고 수평으로 에지(edge)를 가지고 있는 부분은 매칭이 잘 이루어 지지 않음도 알 수 있다.

4. 2 광축 스테레오 운동(axial motion stereo)의 경우

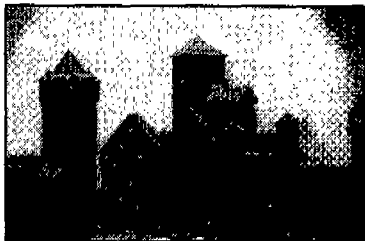
Z축 방향의 운동량 T_z 를 63.5(mm)로 하여 얻은 네 장의 영상을 Fig. 8에 나열 하였다. Fig. 9는 왼쪽 영상에서 특징점들에 대하여 영상변위를 구한 결과이다. Fig. 10은 이 영상변위를 이용한 깊이를 표시한 그림이며 이는 바로 모션 스테레오를 이용한 깊이 정보가 되는 것이다. Fig. 11은 제안한 알고리즘에 의하여 탐색



T0 LEFT IMAGE



T0 RIGHT IMAGE



T1 LEFT IMAGE



T1 RIGHT IMAGE

Fig. 8 Stereo images at time 0 and 1(axial motion)

범위를 정한 후 스테레오 매칭을 한 결과를 나타낸다. 그림에서 알 수 있듯이 FOE부근에 존재하는 특징점에 대해서는 탐색범위가 너무 넓기 때문에 정확한 거리를 구할 수 없음을 알 수 있다.

다음으로 오른쪽 영상에 대하여 영상변위를 구하여 Fig. 12에 나타내었다. 역시 가운데 부분에서는 영상변위가 구해지지 않지만, 왼쪽 영상의 중심부근에 해당하는 영역에서는 영상변위가 구해지는 것을 볼 수 있다. 3.5절에서 제안한 방법으로 먼저 구한 깊이를 보상을 해 줄 수 있으며 그 결과를 Fig. 13에 나타내었다. 영상 중심부근(FOE)에서 깊이 정보를 잘 구할 수 있음을 알 수 있다. 한편 Table 1은 영상 세개를 이용한 경우와 네개를 이용한 경우에 대응점후보의 갯수가 어떻게 변하는지 나타낸 것이다. 탐색 범위내에 존재하는 후보 대응점의 갯수에 대한 특징점의 갯수를 보면 네개의 영상의 경우 그 총계가 세개의 영상의 경우에 비하여 훨씬 작다. 따라서 대응점후보의 평균갯수가 작으므로 대응의 정확도를 높일 수 있다.

5. 결 론

본 논문에서는 스테레오 매칭(stereo matching)을 통하여 고정된 물체까지의 거리를 구할 때, 미소 카메라 병진운동으로부터 얻는 영상변위(optical flow)를 이용하여 스테레오 매칭의 탐색범위를 제한하는 방법을

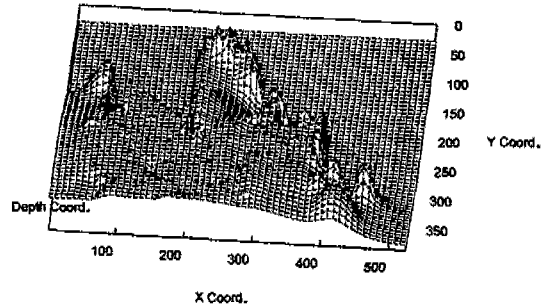


Fig. 11 Depth from stereo matching and left optical flow

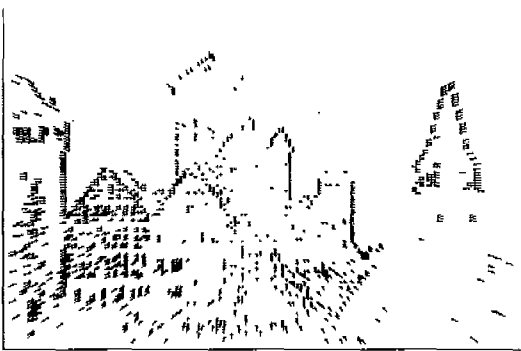


Fig. 9 optical flow at left image

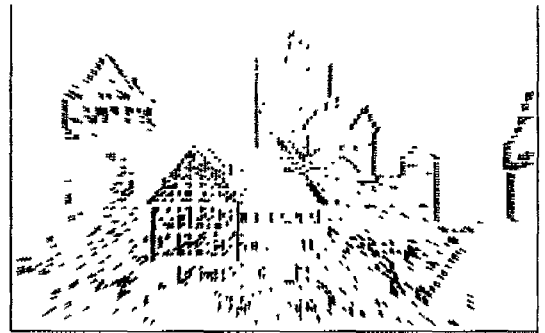


Fig. 12 optical flow at right image

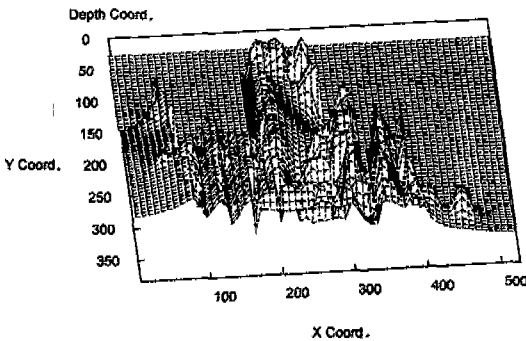


Fig. 10 Depth from optical flow at left image

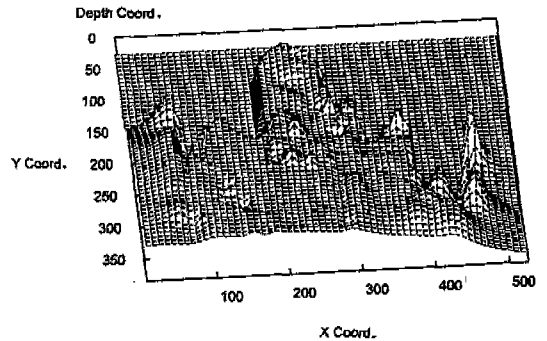


Fig. 13 Depth from stereo matching and left & right optical flow

Table 1 Results of reducing number of candidate mat-ching point(total number of feature= 7339)

No. of match candidate	3 images	4 images
N=1	62.1%	92.8%
N≤2	94.0%	99.4%
N≤3	98.1%	99.8%
N≤4	99.6%	99.9%
N≤5	99.9%	100%
N≤6	99.9%	100%
N≤7	100%	100%

제안하였다. 결과적으로 대응 후보점에 대한 탐색 범위가 줄어들기 때문에 계산량을 줄일 수 있다. 또한 유사 대응 후보점의 수가 작아지기 때문에 대응의 정확도를 높일 수 있으며 이러한 사실을 실험적으로 입증하였다. 영상변위(optical flow)는 이미 알고 있는 카메라의 운동정보를 이용하여 그 방향을 알 수 있다. 또한 스테레오 카메라(stereo camera)가 미소운동을 하면 영상범위가 양쪽 카메라에서 각각 구해지게 되는데 양쪽의 영상범위를 이용하여 대응 후보점을 줄일 수 있는 방법을 제시하였다. 카메라가 광축방향으로 움직일 경우 FOE 부근에서 영상변위(optical flow)값이 매우 작아서 거리를 구하기 힘들어지는데 다른쪽 영상에서 이를 보정한 효과를 실험적으로 입증하였다. 실험 결과 3장(1쌍의 스테레오 영상과 운동 후의 1장의 영상)의 영상을 사용한 결과, 영상에서 전체 특징점의 수가 7339일 경우 대응후보점의 수가 1개로 줄어드는 비율은 62%이었고, 네장의 영상(2쌍의 스테레오 영상)을 사용하였을 때 그 비율은 92.8%로 현격하게 줄어들음을 확인하였다.

본 논문에서는 고정된 배경하에서 카메라의 운동을 이용하여 거리를 구하였다. 카메라의 운동은 외부의 센서를 이용하여 얻을 수 있다고 가정하였다. 하지만 많은 응용에서는 외부의 센서의 도움없이 자유롭게 운동할 수 있는 시스템이 요구된다. 추정된 영상 변위로부터 카메라의 운동파라미터를 구하는 방법을 사용하여 관측자의 운동을 추정함으로써 이러한 문제를 극복할 수 있다.

6. 참고 문헌

1. Umesh R. Dhond and J.K. Aggarwal, "Structure from stereo," IEEE Trans. Sys. Man. & Cyber., Vol. 19, No. 6, Nov 1989.
2. D. Marr and T. Poggio, "A computational theory of human stereo vision," Proc. Royal. Soc. London, Vol. B204. pp. 301-328, 1979.
3. W.E.L. Grimson, "A computational experiments with a feature based stereo algorithm," IEEE Trans. Patt. Anal. & Machine Intell., Vol. 7, pp. 17-34, 1981.
4. Mayhew and Frisby, "Psychological and computational studies towards a theory of human stereopsis," Artificial Intelligence, Vol. 17, pp. 349-385, 1981.
5. Peter Burt and Bela Julesz, "A disparity gradient limits for binocular fusion," SCIENCE, Vol. 208, pp. 615-617, 1980.
6. Daid and Massimo Tistarelli, "Using Camera Motion to Estimation Range for Robot Oarts Manipulation," IEEE Trans. Robotics and Auto., Vol. 6, No. 5, 1990.
7. Masatoshi Okutomi and Takeo Kanade, "A Multiple Baseline Stereo," IEEE Trans. Patt. Anal. & Machine Intell., Vol. 15, No. 4, 1993.
8. Robert C. Bolles, H. Harlyn Baker, And David H. Harimont, "Epipolar-Plane Inage Anaysis: An Approach to Determining Structure from Motion," Int. J. Computer Vision, pp. 7-55, 1987.
9. Enrico Grosso, Giulio Sandini and Massimo Tistarelli, "3-D object Reconstruction Using Stereo and Motion," IEEE Trans. Sys. Man & Cyber., Vol. 19, 1989.
10. Ze-Nian, Li and Danpo Zhang, "Real-time Line-based Motion Stereo," IEEE. International Conference on Robotics and Automaion, pp. 367-372, 1993.

11. John Ens and Ze-Nian Li, "Real-time Motion Stereo," *IEEE. Proceeding of Comp. Vision & Patt. Recog.*, pp.130-135, 1993.
12. Frank Tong and Ze-Nian Li, "Reciprocal-Wedge Transform in Motion Stereo," *IEEE. International Conference on Robotics and Autumation*, pp.1060-1065, 1994.
13. Jeff B. Burl, "A Reduced Order Extended Kalman Filter for Sequantial Images Containing a Object," *IEEE Trans. Image Processing*, July, pp.285-295, 1993.
14. Lingxiao Li and James H. Duncan, "3-D Translational Motion and Structure from Binocular Image Flows," *IEEE Trans. Patt. Anal. & Machine Inttell.*, Vol.15, No.7, pp.657-667, 1993.
15. N.Maru, A.Nishikawa, F.Miyazaki and S.Arimoto, "Actiive Binocular Stereo," *IEEE. Proceeding of Comp. Vision & Patt. Recog.*, pp.724-725, 1993.
16. 이재웅, 이진우, 박광일, 오준호, "다중 스테레오 영상을 이용한 대응문제의 해결," '94 KACC *Computer Vision(1)*, KS-35, Oct 1994
17. Giulio Sandini and Massimo Tistarelli, "Active Tracking Strategy for Monocular Depth Inference Frames," *IEEE Trans. Patt. Anal. & Machine Intell.*, Vol.12, No.1, pp.13-27, 1990.
18. J.F.Canny, "A computational approach to edge detection," *IEEE Trans. Patt. Anal. & Machine Intell.*, Vol.8, No.6, pp.678-698, 1986.
19. Robert M. Haralick and Linda G. Shapirio, *Computer and Robot Vision*, Addison-Wesley, pp.535-544, 1993.