

영상휘도 함수의 온 라인 스테레오 매칭을 이용한 고속 3차원 거리 추정 알고리즘

°권 호열*, 변 증남*, 서 일홍**

*한국과학기술원 전기및전자공학과, **한양대학교 전자공학과

A Fast 3D Depth Estimation Algorithm Using On-line Stereo Matching of Intensity Functionals

H. Y. Kwon*, Z. Bien*, I. H. Suh**

*KAIST, **Hanyang Univ.

Abstract

A fast stereo algorithm for 3D depth estimation is presented. We propose an on-line stereo matching technique, by which the flows of stereo image signals are dynamically controlled to satisfy the matching condition of intensity functionals. The disparity is rapidly estimated from the control of the signal flows and 3D depth is determined from the disparity.

1. 서 론

산업용 로봇의 제어와 이동로봇의 유도문제등에 있어서 실제로 중요한 문제가운데 하나가 로봇 환경에 대한 3차원 정보를 얼마나 빠르고 신뢰성있게 추정할 수 있는가하는 것이다. 이러한 정보추정을 위하여 제안된 거리감지 기법으로는 Simple Triangulation, Structured Lighting, Time-of-Flight Range Finder, Stereo Disparity in Video Images 또는 Image Texture Gradient 를 이용한 것등이 있으며[1,2] 이들 기법중에서 특히 스테레오 비전은 환경에 대한 거리정보를 추출하는 데 있어서 사전 지식을 필요로 하지 않으며 Active Ranging System 에서와 같이 신호발생및 무사장치가 요구되지 않는다는 등의 많은 장점이 있다. 그러나 대부분의 전통적인 스테레오 알고리즘들은 한 쌍의 영상에 대한 특징점 추출과 스테레오 매칭과정에서의 막대한 계산량 때문에 어려움을 겪고 있으며, 이러한 관점에서 볼때 스테레오 비전기법은 실시간 동작이 요구되는 응용분야에 적용하기에는 아직 해결되어야 할 많은 과제를 안고 있는 실정이다.[3]

영상을 이용한 고속 거리추정방법에 대하여 발표된 연구결과들중 중요한 것들은 다음과 같다. Nishihara[4]는 Marr 와 Poggio 의 Zero-Crossing 이론[5]에 기초한 Real-Time Stereo Matcher 를 제안하였는데, 영상처리 를 돕기 위하여 Unstructured Light 기법과 함께 고속의 처리 속도를 갖는 하드웨어인 Pipelined Convolver를 사용하였다. Miller[6]는 스테레오 영상을 Thresholded Binary 영상으로 변환한 뒤 Phase-Locked Loop 를 사용하여 실시간으로 Disparity 를 추정하는 방법을 제안하였다. 또한 Lee, Lim, and Bien[7]은 효율적인 특징점 추출을 위하여 Superimposed Difference Image 의 개념을 도입하여 영상처리 시간을 단축할 수 있음을 보여 주었다.

본 논문에서는 3차원 거리정보 추정을 위한 일종의 고속 스테레오 알고리즘인 온 라인 스테레오 매칭기법이 제시되었다. 스테레오 매칭을 위한 특징값으로는 영상휘도(Intensity) 함수가 사용되었으며 스테레오 영상신호의 흐름이 이들 영상 휘도함수의 매칭조건을 만족하도록 Dynamic 하게 제어된다. Disparity 정보는 영상신호의 흐름에 대한 제어 신호로부터 쉽게 계산될 수 있으며 3차원적인 거리가 이 Disparity 값들로부터 결정된다.

2. 영상휘도 함수의 온 라인 스테레오 매칭

평행하게 배치된 스테레오 비전시스템을 통하여 얻어진 한 쌍의 영상에 대해 생각해보자. $f_r(x,y)$ 를 오른쪽 영상의 Pixel (x,y) 에서의 휘도라고 하고 $f_l(\hat{x},y)$ 를 이에 대응하는 왼쪽 영상의 Pixel (\hat{x},y) 에서의 휘도라고 하자. 이들 대응점 사이에서 일어날 수 있는 Occlusion 과 Reversal 현상등을 다른 연구[4,6,7]에서 처럼 무시하면 이들 두개의 영상은 Pixel

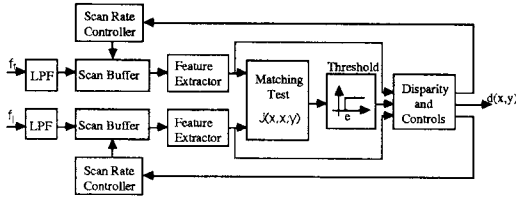


Fig. 1 Block diagram for the proposed depth-estimation system.

(x, y) 로부터 3차원 물체의 한점까지의 거리 $z(x, y)$ 에 따라 Scan Line y 상에서의 Disparity $d(x, y) = \hat{x} - x$ 만큼 위치적으로 왜곡되어 있는 점을 제외하면 매우 흡사한 정보를 갖고 있음을 알 수 있다.

그림 1은 영상위도 함수의 온 라인 스테레오 매칭을 이용한 3차원 거리추정을 위하여 제안된 시스템의 블록다이어그램이다. 영상 신호는 먼저 Noise 의 영향을 줄이기 위하여 Low-Pass Filter 를 통과한다. Scan buffer는 Digital 또는 CCD Analog 방식의 Shift Register 로 구성되며, Scan Rate Controllers 는 이전의 영상 Pixel 에 대한 결과에 따라서 다음 매칭 평가에 사용될 영상 Pixel 을 결정하는 역할을 한다. 한 쌍의 Pixel 에 대한 매칭 평가는 각 Pixel 의 위도함수를 사용하여 Threshold Logic 에서 수행된다.

스테레오 영상으로부터 거리정보를 얻기는 과정은 크게 특징점 추출 (Feature Extraction) 과 스테레오 매칭과정의 두 부분으로 나눌 수 있는데, 실시간 거리정보 추정이 이루어 지려면 이들 두 가지과정이 모두 온 라인 (on-line) 으로 실행되어야 한다. 이러한 온 라인 처리의 구현을 위하여 먼저 특징점 벡터가 영상위도 함수[8]를 사용하여 Analog 혹은 Digital 필터에 의해 $f_r(x, y) = [f_r(x, y), f_r^{(1)}(x, y)]^T$ 및 $f_l(\hat{x}, y) = [f_l(\hat{x}, y), f_l^{(1)}(\hat{x}, y)]^T$ 와 같이 추출된다. 이때 첫첨차 (1) 은 1차 도함수를 의미한다. 이렇게 얻어진 한 쌍의 특징점 벡터에 대하여 온 라인 스테레오 매칭은 다음과 같이 수행된다. 단순한 형상의 물체로만 이루어진 로봇 비전의 환경속에서 Ordering Constraints[3] 를 사용할때, 스테레오 영상 신호의 매칭여부는 각 Scan Line 에 있어서 가장 좌측에 있는 Pixel 들로부터 아래와 같은 영상위도 함수의 매칭오차 기준 $J(x, \hat{x}, y)$ 에 의해 평가되기 시작한다.

$$J(x, \hat{x}, y) = \|A (f_r(x, y) - f_l(\hat{x}, y))\| \quad (1)$$

이 때 $A = \text{diag}(\alpha_1, \alpha_2)$ 은 가중치 대각 행렬이며, $\| \cdot \|$ 은

벡터에 대한 Euclidean norm 을 나타낸다.

초기에 \hat{x} 은 x 와 같은 "0"의 값을 갖는다. 이들 Pixel들에 대해 평가된 $J(x, \hat{x}, y)$ 이 미리 지정된 Threshold 값인 ϵ 보다 작거나 같다면, (x, y) 과 (\hat{x}, y) 의 한 쌍의 Pixel은 매칭이 된 것으로 판정하며, 매칭 평가는 새로운 Pixel 쌍, $(x + \delta x, y)$ 과 $(\hat{x} + \delta x, y)$ 에 대해서 반복적으로 수행된다. 이때 δx 는 영상의 Pixel 크기를 의미한다. 그러나, 만일 매칭 평가의 결과가 ϵ 보다 크면, 그 Pixel 쌍은 매칭되지 않은 것으로 판정한다. 이 때 두 개의 영상신호중 어느 쪽이 빨리 변화했는가에 따라서 두 가지의 매칭되지 않은 상태가 존재한다. 만일 오른쪽 영상신호의 흐름이 왼쪽 영상신호쪽보다 앞선다면 Disparity는 감소하게 된다. 그렇지않고 왼쪽 영상신호의 흐름이 오른쪽 영상신호보다 앞서면 Disparity는 증가하고 있는 것을 의미한다. 이와 같이, 다음 매칭에 사용될 Pixel 쌍들이 Disparity가 증가하는 경우에는 (x, y) 과 $(\hat{x} + \delta x, y)$, 감소하는 경우에는 $(x + \delta x, y)$ 과 (\hat{x}, y) 들로 위에서 살펴 본 영상신호들의 흐름에 대한 관측에 의해 Dynamic하게 결정된다.

3. 거리 정보의 추정

Pixel (x, y) 에 있어서 Disparity $d(x, y)$ 는 매칭 조건을 만족시키도록 하는 Scan 신호의 흐름에 대한 제어신호로부터 다음과 같이 얻어진다.

$$d(x, y) = d(x - \delta x, y) + \delta d(x, y), \quad (2)$$

이 때 Disparity의 증분 $\delta d(x, y)$ 은 위의 매칭 평가과정에서 두 개의 스테레오 영상신호가운데 어느 쪽의 신호가, 몇 Pixel만큼 앞서는 가에 의해 결정된다. 따라서 각 Pixel에 대응하는 거리정보 $z(x, y)$ 는 스테레오 비전 시스템에 있어서 두 개의 카메라사이의 Baseline 이 D 이고 각 카메라 렌즈의 촛점거리가 λ 로 주어지면 아래와 같이 결정될 수 있다.

$$z(x, y) = \frac{D \cdot \lambda}{d(x, y)}, \quad (3)$$

4. 실험 및 결과

실험은 연구실 환경하에서 수행되었다. 스테레오 영상의 입력을 위하여 각각 8 [mm]의 촛점 거리를 갖는 렌즈를 사용한 두 개의 동일한 CMOS 카메라가 20 [cm]의 Baseline을 사이에 두고 평행하게 설치되었다. 이들 CMOS

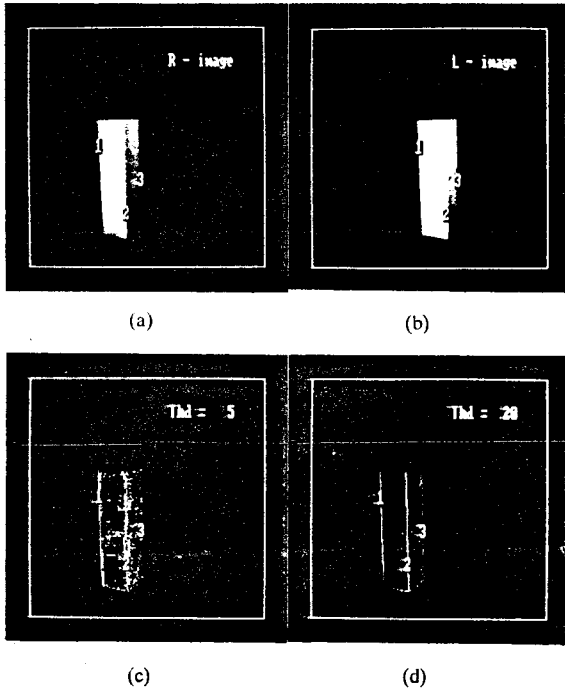


Fig. 2 Experimental results.

카메라에서 얻은 영상은 DT2853 Frame Grabber(Data Translation Inc.)에 의하여 256 제조를 갖는 256 x 256 의 디지털 영상으로 변환된다. 제안된 알고리즘은 Microsoft-C 컴파일러를 사용하여 IBM-PC/AT 시스템상에서 구현되었다.

실험된 결과를 그림 2 에 보였다. 입력된 오른쪽과 왼쪽의 영상은 각각 그림 2a 와 그림 2b 와 같다. 매칭오차 기준을 위해 식 (1)의 가중치들로서 각각 $\alpha_1 = 1$ 과 $\alpha_2 = 0.1$ 을 사용하였으며, 본 논문의 온 라인 매칭 방법을 통하여 얻어진 Disparity Map 이 매칭오차 한계를 $\epsilon = 5$ 과 $\epsilon = 20$ 로 한 경우에 대해서 각각 그림 2c 와 그림 2d 에 보였으며, 이들 Disparity Map 에 있어서 각 특징점들의 밝기가 Disparity 의 절대값에 비례하도록 하였다. 3차원 거리 추정에 있어서 그림 2 의 1, 2 및 3 으로 표시된 점들까지의 실제 거리는 각각 115, 110 and 121 [cm] 인 데 비해서, $\epsilon = 5$ 인 경우는 각각 116.0, 105.2 과 125.7 [cm], $\epsilon = 20$ 인 경우는 각각 116.0, 107.7 과 122.3 [cm] 로 계산되어 양호한 거리추정 능력을 보여주고있다. 제안된 방법에 의하면 보다 엄밀한 매칭 오차 한계를 적용할 때 보다 자세하지만 Noise 의 영향이 커진 거리 정보를 얻게되는 것을 알 수 있다. 알고리즘 수행시간에 있어서 위의 Disparity Map 을 얻는 데

약 3 초가 걸렸으며, 이것은 같은 시험용 영상을 사용하여 약 20 초가 소요된 Lee, Lim, and Bien[7]의 방식과 비교할 때 현저한 시간 단축의 효과가 있음을 알 수 있다.

5. 결 론

실제적인 로봇 응용을 위한 스테레오 Disparity 및 3차원 거리추정을 위한 고속 알고리즘이 제안되었다. 제안된 알고리즘은 영상 회도 함수의 매칭조건을 만족하도록 스테레오 영상신호의 흐름을 Dynamic 하게 제어하는 온 라인 스테레오 매칭 기법에 의한 것이다. 본 논문의 연구가 갖는 특징의 하나는 온 라인 매칭을 사용하여 제조를 갖는 스테레오 영상신호로부터 Disparity 정보를 직접 추출한다는 점이다. 보다 신뢰성이 향상된 Disparity 추정을 하기 위해서는 Noise에 약한 Threshold 방법을 보완하기 위하여 입력 영상이 갖는 Noise의 통계학적 성질을 측정하여 매칭오차 한계인 ϵ 값을 가변하는 Adaptive Thresholding 과 같은 방법도 적용할 수 있다. 또한 보다 고차의 도함수를 포함하는 영상 회도함수를 사용하거나 다양한 대역 특성을 갖는 필터를 사용하게 되면 제안된 알고리즘은 Natural Scene 과 같은 복잡한 영상에 대해서도 적용할 수 있으리라고 전망된다. 이상의 특징과 아울러 알고리즘의 구조상 로봇 응용분야에서 Low Cost 하드웨어로 구성될 수 있는 실시간 거리추정기로서 적용가능하다.

참 고 문 헌

- [1] JARVIS, R. A. : ' A perspective on range finding techniques for computer vision', IEEE Trans PAMI, 1983, PAMI-5, pp. 122 - 139.
- [2] FU, K. S., GONZALEZ, R. C., and LEE, C. S. G. : *Robotics: Control, Sensing, Vision, and Intelligence*. McGraw-Hill, 1987.
- [3] NISHIHARA, H. K. and POGGIO, T. : ' Stereo vision for robotics', *Robotics Research*, edited by BRADY, M. and PAUL, R., MIT press, 1984, pp. 489 - 505.
- [4] NISHIHARA, H. K. : ' Practical real-time imaging stereo matcher', *Optical Engineering*, 1984, 23(5), pp. 536 - 545.
- [5] MARR, D. and POGGIO, T. : ' A computational

- theory of human stereo vision', *Proc. R. Soc. Lond. B*, 1979, **204**, pp. 301 - 328.
- [6] MILLER III, W. T. : ' Video image stereo matching using phase-locked loop technique ', *IEEE Proc. of Int. Conf. on Robotics and Autom.*, 1986, **1**, pp. 112 - 117.
- [7] LEE, J. S., LIM, J., and BIEN, Z. : ' Simple stereo algorithm for 3-D positioning of object points', *IEE Electronics Letter*, 1988, **21**(18), pp. 1139 - 1140.
- [8] KASS, M. : ' A computational framework for the visual correspondence problem', *Proc. ARPA Image Understanding Workshop*, Washington, D. C., 1983.