

적응적 가중치를 이용한 스테레오 정합 기법

윤희주* · 황영철* · 차의영*

*부산대학교 컴퓨터공학과

Adaptive weight approach for stereo matching

Hee-Joo Yoon* · Young-Chul Hwang* · Eui-Young Cha*

^{*}Dept. of computer engineering, Pusan National University

E-mail : kidjulia@gmail.com

요 약

본 논문에서는 스테레오 영상의 대응점을 찾기 위한 영역 기반 스테레오 정합 기법을 제안한다. 영역 기반 스테레오 정합의 주된 문제점은 윈도우 크기에 따라 다른 결과를 초래한다는 것이다. 지금까지 대부분의 영역 기반 정합 기법은 윈도우의 크기를 반복적으로 갱신하는 방법을 사용하였으나, 이는 초기 시차(disparity)에 매우 민감하며 계산 비용도 많이 든다. 이러한 문제를 해결하기 위해, 본 논문에서는 스테레오 영상의 특징 정보를 이용하여 가중치를 생성하고, 각 영상의 대응점을 탐색하여 정합한다. 먼저, 평행하게 설치된 두 대의 카메라로부터 획득된 영상에 대한 에지를 검출하여 특징점을 추출한다. 이를 이용하여 두 영상간의 상관관계를 구하여 가중치 함수를 생성하고, 각 영상에 대한 가중치를 적용한 후, 기준영상에 대한 대응점을 찾아 정합한다. 제안된 방법의 성능을 평가하기 위하여 ground truth가 존재하는 다양한 스테레오 영상을 이용하여 실험하였으며, 실험결과 다양한 영상에서도 적응적인 가중치를 생성함으로써 향상된 결과를 보였다.

ABSTRACT

We present a area-based method for stereo matching using varying weights. A central problem in a area-based stereo matching is different result from selecting a window size. Most of the previous window-based methods iteratively update windows. However, the iterative methods very sensitive the initial disparity estimation and are computationally expensive. To resolve this problem, we proposed a new function to assign weights to pixels using features. To begin with, we extract features in a given stereo images based on edge. We adjust the weights of the pixels in a given window based on correlation of the stereo images. Then, we match pixels in a given window between the reference and target images of a stereo pair. The proposed method is compared to existing matching strategies using both synthetic and real images. The experimental results show the improved accuracy of the proposed method.

키워드

Stereo vision, Stereo matching, Weight, Weighted matching, Disparity

1. 서 론

스테레오 비전 시스템(stereo vision system)은 양안을 이용해 원근감을 느끼는 인간의 시각체계를 모방한 것으로, 두 대 이상의 카메라를 이용하여 획득한 2차원 영상 정보로부터 3차원 정보를 획득하는 것이다[1][2]. 스테레오 비전 시스템에서

가장 중요하면서도 시간을 요하는 일은 스테레오 정합(stereo matching) 문제를 해결하는 것이다.

스테레오 정합은 스테레오 영상 간의 대응점을 찾아서 이 점들 간의 시차(disparity)를 구하는 과정으로 3차원 깊이 정보를 획득하는데 있어서 가장 중요한 과정이라 할 수 있다. 스테레오 정합 문제를 해결하기 위한 방법으로는 크게 특징기반

방법(feature based approach)과 영역기반 방법(area based approach)이 있다.

특징 기반의 스테레오 정합은 스테레오 영상의 특징을 추출하여 각 영상의 탐색영역 내에서 정합 후보 점들을 비교하는 정합 방법으로, 특징 추출에 사용되는 특징으로는 에지(edge), 분할된 경계선(line segment), 선소(edge segment), 윤곽선(contour), 모서리(corner) 등이 있다. 특징 기반 방법은 처리속도를 향상시킬 수 있는 반면, 영상 전체 영역에 대한 정확한 시차 획득이 불가능하다는 단점이 있어 깊이 정보를 획득하기 어렵다.

영역 기반의 스테레오 정합은 상관관계 기반 스테레오 정합(correlation based stereo matching)이라고도 하며, 정합을 결정하기 위해서 스테레오 영상의 정합 윈도우의 밝기 정보 변화를 이용하여 정합을 결정하므로 잡영(noise)에 약한 면이 있으나, 전체 시차 정보를 획득할 수 있다[3].

본 논문에서는 스테레오 영상의 대응점을 찾아 3차원 깊이 정보를 획득하기 위해 적응적 가중치를 적용한 고속의 영역기반 스테레오 정합 기법을 제안하였다.

II. 전체 시스템 개요

그림 1은 본 논문에서 제안하는 전체 시스템의 구조를 도식화한 것이다.

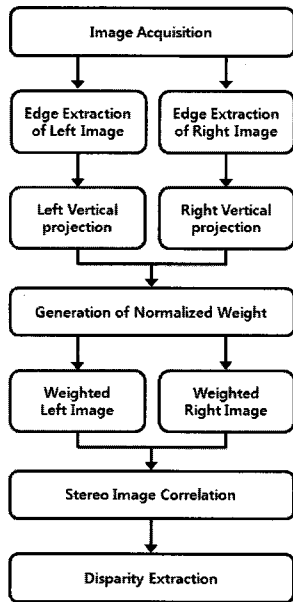


그림 1 시스템 구조도

먼저, 평행하게 설치된 두 대의 카메라를 이용하여 영상을 획득한다. 특징 추출을 위해 각 영상

에 대한 에지를 추출한 후, 수직 투영(vertical edge)을 수행한다. 각 투영 영상에 대한 상관관계를 구하여 가중치 행렬을 생성한다. 각 영상에 대한 가중치를 적용한 후, 기준영상에 대한 대응점을 찾아 정합하여 3차원 깊이 정보를 획득한다.

III. 적응적 가중치(weight) 생성

본 논문에서 스테레오 영상의 시차(disparity)를 구하기 위해 사용하게 될 영역 기반 정합 방법은 윈도우 크기에 민감하므로 이를 보완하기 위하여 각 영상에 가중치를 적용한다. 가중치를 이용한 스테레오 정합 방법에 대한 많은 연구가 진행되고 있으며, 일반적으로 반복적으로 가중치를 생성하는 방법을 채택하고 있다. 이러한 반복적 가중치 생성 방법의 경우, 초기 가중치에 민감하며 계산 시간이 오래 걸리는 단점을 가지고 있다[4]. 이를 보완하기 위하여 본 논문에서는 각 영상의 특징정보를 이용하여 가중치 행렬을 생성하는 방법을 제안하였다.

3.1 특징 추출

가중치 행렬을 생성하기 위하여 각 영상의 에지 영상을 획득한 후, 투영(projection)을 수행한다. 평행식으로 설치된 스테레오 비전 시스템은 카메라 교정(camera calibration)이 정교하지 못하더라도 수직방향의 시차보다 수평방향의 시차가 더 크다. 그러므로 에지 영상은 수직성분이 비교적 잘 추출되는 소벨(Sobel) 에지 검출 방법을 사용하고, 추출된 에지 영상은 $E(x, y)$ 로 나타낸다. 획득된 에지 영상에 대한 수평방향의 시차를 찾기 위해서 수직 방향의 에지 영상을 투영하여 축적된 영상 $P(x)$ 를 다음 식 (1)과 같이 구한다.

$$P(x) = \sum_y E(x, y) \quad \dots (1)$$

식 (1)을 이용하여 스테레오 영상에 대한 각각의 투영 성분 $P_L(x)$ 과 $P_R(x)$ 을 획득한다.

3.2 가중치 행렬 생성

두 영상에서 유사성이 높은 대응점을 찾기 위해 SAD(Sum of Differences), SSD(Sum of Squared Differences), MAD(Mean of Absolute Difference), NCC(Normalized Cross Correlation) 등의 비용함수를 사용한다. 본 논문에서는 고속의 가중치 행렬을 생성하기 위하여 SAD 비용 함수를 이용하여 각 투영영상의 상관관계를 계산한다. 비용함수는 일반적으로 정방형의 윈도우(window)를 사용하여 상관관계를 구한다. 그러나 본 논문에서는 에지영상의 수직 투영 성분을 이용하여 상관

관계를 구하므로 2차원의 영상정보가 아니라 1차원의 정보를 이용한다. 이러한 점을 고려하여 SAD 비용함수를 다음 식 (2)와 같이 수정하여 사용한다.

$$w_{MSAD}(x) = \sum_{x \in W} |P_R(x) - P_L(x)| \quad \dots (2)$$

SAD 비용함수는 SSD, MAD와 더불어 최소화 비용 함수(minimizing cost function)로, 결과 값은 정합 오차를 나타내며, 그 값이 작을수록 정합 점일 가능성이 높아진다. 식(2)의 비용함수를 이용하여 얻은 정합 오차로 행렬을 생성한 후, 정규화(normalize) 과정을 거쳐 가중치 행렬을 생성한다.

IV. 영역 기반 스테레오 정합

영역 기반의 스테레오 정합은 정방향의 윈도우를 사용하며, 물체의 경계가 확장되는 문제점이 있다. 창 크기가 작으면 비교적 정확한 물체의 경계를 찾을 수 있으나 잡영이 많이 생기게 되며, 창 크기가 크면 잡영은 줄일 수 있으나 물체의 경계가 확장되는 문제점이 있어 스테레오 정합을 통한 3차원 깊이 정보를 추출하기에는 문제점이 많이 존재한다[5]. 이를 보완하기 위해 본 논문에서는 각 영상에 대한 적응적 가중치를 적용한 후, 기준 영상에 대한 대응점을 찾아 정합한다.

$$f_{SSD}(x, y, d) = \sum_{x, y \in W} \{I_R(x, y) - I_L(x + d, y)\}^2 \quad \dots (3)$$

논문에서는 각 영상의 대응점을 찾아 정합을 수행하기 위해 식 (3)과 같이 SSD 비용함수를 이용하여 상관관계를 구한다. I_R 과 I_L 은 스테레오 입력 영상을 나타내며 W 는 입력 픽셀 (x, y) 을 에워싸는 윈도우를 나타낸다.

스테레오 영상은 평행하게 설치되어 있는 두 대 이상의 카메라를 통해 획득한 영상이므로 동일한 위치의 물체라 하더라도 외부 광원에 의해 밝기 차가 생길 수 있다. 이러한 밝기차가 오정합을 유발할 수 있으므로 본 논문에서는 입력 영상에 히스토그램 스트레칭(histogram stretching) 기법을 적용하여 영상을 개선한다.

$$I_R'(x, y) = \frac{I_R(x, y) - I_{\min}}{I_{\max} - I_{\min}} \times 255 \quad \dots (4)$$

$$I_L'(x, y) = \frac{I_L(x, y) - I_{\min}}{I_{\max} - I_{\min}} \times 255 \quad \dots (5)$$

식 (4), (5)는 각 영상의 히스토그램 스트레칭 과정을 수식화한 것이다. I_{\max} 와 I_{\min} 값은 영상

내의 픽셀에서 가장 높은 밝기 값과 가장 낮은 밝기 값의 크기를 나타낸 것이다.

$$\hat{I}_R(x, y) = I_R'(x, y) * w_R \quad \dots (6)$$

$$\hat{I}_L(x, y) = I_L'(x, y) * w_L \quad \dots (7)$$

$$f_{WSM}(x, y, d) = \sum_{x, y \in W} \{\hat{I}_R(x, y) - \hat{I}_L(x + d, y)\}^2 \quad \dots (8)$$

식 (6), (7)의 \hat{I}_R 과 \hat{I}_L 은 개선된 입력영상에 각각의 가중치를 적용한 영상이다. 각 영상의 대응점 찾기 위해 본 논문에서 제안한 방법을 식 (8)과 같이 나타낼 수 있다.

V. 실험 결과

본 논문에서 제안된 방법은 Pentium4 3.0GHz, 1GB RAM의 PC환경에서 Visual C++.NET으로 구현하였으며, 성능 평가를 위해 ground truth가 존재하는 다양한 스테레오 표준 영상을 이용하였다.

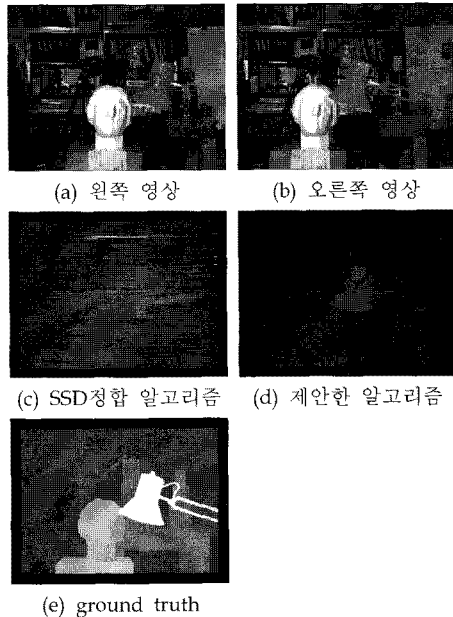


그림 2 시차지도(disparity map)

그림 2는 Tsukuba 대학에서 획득한 스테레오 영상의 일반적인 영역 기반 정합 알고리즘인 SSD 정합 알고리즘을 통해 획득한 시차지도와 제안한 방법으로 획득한 시차 지도를 나타내고 있다. 그

림 2의 (c), (d) 모두 35×35 크기의 동일한 윈도우를 사용했으며, 제안한 알고리즘의 가중치 행렬 생성과정에 사용된 윈도우는 빠른 가중치 생성을 위해 3×3 크기의 윈도우를 사용했다. 그림 2에서 보이는 바와 같이, ground truth와 비교했을 때 제안한 알고리즘이 시차지도의 깊이 정보를 비교적 더 잘 표현한 것을 알 수 있다.

VI. 결론

본 논문에서는 스테레오 영상으로부터 대응점을 찾아 정합하기 위해 적응적 가중치를 적용한 영역 기반 스테레오 정합 기법을 제안하였다. 스테레오 영상으로부터 에지 기반의 특징 영상을 추출한 다음, 수직 투영하여 1차원 행렬을 만든 후, 비유함수를 이용하여 가중치 행렬을 생성하였다. 이를 밝기 값이 개선된 입력 영상에 적용한 후 영역 기반의 스테레오 정합을 수행하였다.

제안한 방법에서는 가중치를 적용하여 영역 기반 정합이 윈도우 크기에 민감한 단점을 보완할 수 있었으며, 반복적으로 가중치를 생성한 방법에 비해 빠르게 가중치를 생성할 수 있었다. 또한 그림 2와 같이 비교적 깊이 정보가 잘 표현된 시차 지도를 생성함을 알 수 있다.

참고문헌

- [1] Heiko Hirschmuller, "Improvements in Real-Time Correlation-Based Stereo Vision", IEEE Workshop on Stereo and Multi-Baseline Vision, Page(s) : 141-148, December 2001
- [2] Clark F. Olson, "Maximum-Likelihood Image Matching," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, Issue 6, Page(s) : 853-857, June 2002
- [3] YongSheng Chen, YiPing Hung, "Fast Block Matching Algorithm Based on the Winner-Update Strategy", IEEE Transactions on Image Processing, Vol. 10, Issue 8, Page(s) : 1212-1222, August 2001
- [4] Yihua Xu, Dongsheung Wang, "Stereo Computation using Radial Adaptive Windows", IEEE Pattern Recognition Proceedings 16th International Conference, Vol. 3, Page(s) : 595-598, August 2002
- [5] T. Kanade, M. Okutomi, "A stereo matching algorithm with an adaptive window : theory and experiment", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.16, Issue 9, Page(s) : 920-932, September 1999