

동적 계획법과 이미지 세그먼트를 이용한 스테레오 정합

동원표^o 정창성

고려대학교 전자컴퓨터공학과

pposajige^o@korea.ac.kr, csjeong@charlie.korea.ac.kr

Stereo matching using dynamic programming and image segments

Won-Pyo Dong^o Chang-Sung Jeong.

Dept. of Electronics Engineering Graduate School, Korea University

요 약

본 논문에서는 동적 계획법(dynamic programming)과 이미지 세그먼트(segment)를 이용한 새로운 스테레오 정합(stereo matching)기법을 제안한다. 일반적으로 동적 계획법(dynamic programming)은 빠르면서도 비교적 정확하고, 조밀(dense)한 disparity map을 얻을 수 있다. 그러나 경계(boundary)근처의 폐색 지역(occlusion region)이나, 텍스처가 적은 모호한 영역에서는 잘못된 결과를 유도할 수 있다. 본 논문에서는 이러한 문제점들을 해결하기 위해, 먼저 이미지를 아주 작은 영역으로 분할(over-segmentation)하고, 이런 작은 영역들이 비슷한 disparity를 가질 것 이라고 가정한다. 다음으로 동적 계획법(dynamic programming)을 통해 정합을 수행한다. 여기서 계산비용(cost)은 기존의 정합원도우 안에서 세그먼트 영역을 적용한 새로운 비용항수를 사용하며, 이 새로운 비용항수를 통해 정확도를 높인다. 마지막으로 동적 계획법을 통하여 얻어진 조밀한 disparity map을 세그먼트 영역들의 시각특성(visibility)과 유사도(similarity)를 이용하여 에러를 찾아내고, 세그먼트 정합을 통해 수정함으로써 정확한 disparity map을 찾아낸다.

1. 서 론

스테레오 정합(stereo matching)은 두 개 또 그 이상의 이미지들 사이의 대응점들(correspondences)찾아내는 컴퓨터 비전의 한 분야이며, 이런 대응점들의 위치차를 보통 disparity라 부른다. 컴퓨터 비전 분야에서 스테레오 정합은 많은 응용을 가지는 근본적인 아주 중요한 문제이다.

스테레오 정합을 위한 많은 알고리즘들이 개발되어 왔으며, 최근에 이런 알고리즘들의 평가와 리뷰를 [1]에서 볼 수 있다. 일반적으로 스테레오 알고리즘은 크게 두 분류로 나눌 수 있다. 첫 번째는 지역기반(local area window-based) 알고리즘이다. 이 방법은 주어진 점(pixel)의 유한한 범위 안에서 윈도우영역안의 밝기(intensity)를 비교하여 대응되는 점을 찾는다. 이런 지역적 방법(local method)은 텍스처가 많은 지역에서는 비교적 정확한 시차를 얻을 수 있으나 텍스처가 적은 지역에서는 부정확한 시차를 만들어내고, disparity가 불연속한 경계(boundary) 근처에서는 disparity가 뭉개지며(blur) 폐색(occlusion)영역에서는 부정확하게 된다. 두 번째는 전역(global) 알고리즘으로, disparity map의 smoothness 가정을 이용하고, 다양한 최소화(minimization)기법을 사용한다. 전역 알고리즘은 비교적 정확하고 조밀한 disparity를 측정할 수 있지만, 계산 비용이 너무 크다는 단점이 있다. 대표적인 방법이 동적 계획법이다. 동적 계획법은 빠르면서도 비교적 정확하기 때문에 많이 사용되어오고 있다. 하지만 동적 계획법

은 스캔라인만을 고려하여 수행되기 때문에 스캔라인이 아닌 수직 성분을 고려해야만하고 많은 논문들에서 이 문제점을 다루고 있다[2, 3]. 그리고 동적 계획법은 보통 유일성(uniqueness)제약과 순차성(ordering)제약을 사용하기 때문에 잘 못된 시차는 연속적인 에러를 일으키게 된다. 따라서 동적 계획법에서의 오류를 줄이기 위해서 다른 제약이 필요하다. 이에 따라 지역적인 방법과 전체적인 방법을 고려하기 위하여 본 논문에서는 segment를 이용한다.

Tao는 컬러 세그먼트에 기반 하는 스테레오 프레임웍(framework)을 제시했다[4]. 이 프레임웍에서, 동차(homogeneous)컬러 세그먼트 영역에서 시차의 변화는 크지 않다고 가정했으며, 이 가정을 바탕으로, 기존 이미지를 시차에 따라 참고 이미지로 변형 시켰을 때, 기존 이미지와 참고 이미지는 정확히 같을 것이라고 생각했다. 또 Bleyer는 이런 가정들을 이용하여 세그먼트들을 layer화시키는 방법을 제안했다[5].

본 논문에서는 동적 계획법과 세그먼트에 기반 하여 지역적인특성과, 전역적인 특성을 결합시키는 새로운 정합 방법을 제안한다. 우리는 동적 계획법에서 사용하는 정합원도우를 변형시켜 세그먼트 영역을 사용해 더 정확한 비용(cost)을 계산하고, 동적 계획법을 통하여 얻어진 조밀한 disparity map을 Tao가 가정한 세그먼트의 특성, 세그먼트 영역간의 시각특성(visibility)과 유사도(similarity)를 이용하여 에러를 찾아내고 수정함으로써 정확한 disparity map을 찾아낸다. 본 논문에서 사용되는 영상들은 계산의 편의를 위해 수평라인으로 교정된 이미지를 사용한다.

2. 제안된 알고리즘

본 논문에서 제안된 알고리즘은 먼저 기준이 되는 이미지를 over-segmentation한다. 다음에는 세그멘테이션 영역을 기반으로 하여 동적 계획법을 이용해 정합을 하여 조밀한 disparity map을 얻는다. 그리고 찾아낸 disparity map을 세그멘테이션 영역별로 평가하고 valid 영역과 invalid영역으로 나눈다. 마지막으로 invalid영역을 warping 연산을 이용해 세그먼트 영역 정합을 하여 invalid영역에 대한 올바른 disparity를 찾아낸다.

2.1 세그멘테이션

본 논문에서 제시한 방법의 가장 근본은 세그먼트이다. 우리는 같은 세그먼트 영역 안에서 시차의 변화는 크지 않고 오직 세그먼트의 경계(boundary)에서만 시차가 크게 변한다고 가정한다. 따라서 위 조건을 만족시키기 위해 영역을 아주 작게 분할시킨다. 영역들의 크기와, 그 영역이 얼마나 경계를 구분할 수 있는가는 실험에 큰 영향을 준다. 본 논문에서는 watershed기법을 사용하여 영역을 분할하였다.

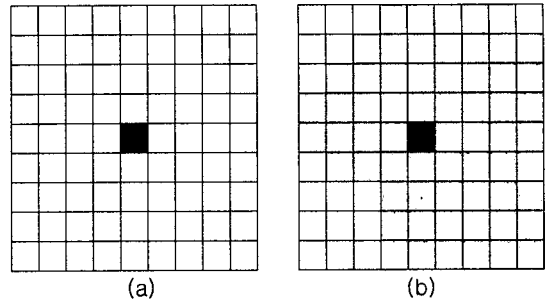
2.2 동적 계획법 비용

이미지를 세그멘테이션하고 동적 계획법을 이용하여 정합을 수행한다. 일반적으로 동적 계획법에서 사용되는 정합비용(matching cost)은 일정크기의 윈도우를 사용하여 계산되는데, 보통 SAD(Sum of Absolute Difference), SSD(Sum of Square Difference), NCC(Normalized Cross Correlation)가 사용된다. 그러나 이런 방법들은 윈도우의 크기에 민감하다는 단점이 있다. 만약 윈도우의 크기가 너무 크면 disparity map이 세밀하지 못하고, 윈도우의 크기가 너무 작으면 많은 노이즈가 생기게 되므로 적당한 윈도우를 설정하는 것이 중요하다. 이에 윈도우의 크기를 다양하게 변형시켜서 최적의 값을 비용으로 사용하는 방법[6]이 제시되었으나 이는 동적 계획법에서 사용하기에는 효율적이 못하다. 최근에는 subsampling 문제를 고려해 샘플링에 둔감한 방법[7]이 사용되기도 한다. 본 논문에서는 고려되는 윈도우 크기를 위에서 계산된 분할된 영역에 기반하여, SAD를 사용하여 비용을 계산한다. 변형된 윈도우를 사용한 SAD의 비용은 아래 식(1)과 같다.

$$\alpha_{x_l, x_r} = \sum_{i, s} \begin{cases} |I(x_l+i, s) - I(x_r+i, s)| & (x_l, s) \in R \\ w_2 |I(x_l, s) - I(x_r, s)| & \text{otherwise} \end{cases} \quad (1)$$

식(1)에서 l, r 은 각각 left 이미지, right 이미지를 나타내고 $I(x, s)$ 는 그 위치에서의 밝기 값을 나타내며, 사용된 윈도우의 크기는 w 이다. s 는 s 번째 스캔라인을 나타내고, R 은 하나의 세그먼트 영역을 나타낸다. 따라서 left 이미지가 기준이미지라고 했을 때 left 이미지위의 점 (x_l, s) 에서(그림1의 윈도우 중심), 이 점이 해당되는 세

그먼트 영역(그림1(b)의 회색영역)을 찾고 윈도우 크기 w 안에서 해당영역에 대해서는 일반적인 차이 값을, 그렇지 않은 경우 중심점의 차이 값으로 계산하게 된다. 그림1(a)에서처럼 윈도우 전체에 대해 이루어지는 기존의 계산법에서 나타날 수 있는 윈도우의 크기 변화에 의해 생기는 영향을 줄일 수 있다.



[그림1] (a)기존의 윈도우 계산영역 (b)제안된 윈도우 계산영역

2.3 세그먼트기반 개선

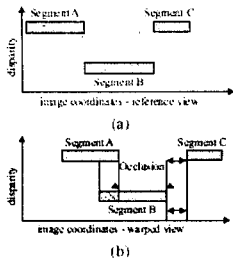
본 논문에서 over-segmentation을 이용하여 각 세그먼트 영역은 비슷한 disparity를 가지며, 오직 영역의 경계에서만 disparity 변화가 커진다고 가정했다. 이 제약에 따라서 우리는 동적 계획법을 통해 얻어진 disparity 이미지를 세그먼트 영역에 따라 평가하고 정확한 정합 점을 찾아내야 한다. 이를 위해 우리는 세그먼트 영역에 대한 분산 값(var)을 계산하고 일정한 임계값(α)에 따라 그 영역의 옳고 그름을 판단한다.

$$\begin{cases} \text{valid} & var < \alpha \\ \text{invalid} & var \geq \alpha \end{cases} \quad (2)$$

일반적으로, 비슷한 disparity를 갖는 영역의 분산 값은 작을 것이고, 그렇지 않은 영역, 즉 disparity변화가 큰 영역은 큰 분산 값을 가질 것이다. 식(2)는 영역별로 이루어지며 이 과정을 통해 우리는 올바른 정합영역과 잘못된 영역을 임계값 α 에 따라 valid영역과 invalid영역으로 구분하고 표시한다. 동적 계획법결과에서 찾은 오류 부분을 그림3(e)에서 볼 수 있다.

invalid영역은 각 세그먼트 영역사이의 밝기 값의 차이를 이용하여 계산되며 이 값을 최소화시키는 disparity를 그 invalid 세그먼트 영역의 값으로 결정한다. Bleyer[5]는 그림2와 같이 Warping 연산을 나타냈다. 우리는 이 연산을 이용하여, 각 영역간의 연산에서 disparity와 occlusion영역을 계산 대상에서 제외시킴으로써 정확도와 속도를 높인다. 이는 일반적으로 occlusion지역이 정합과정에서 많은 오류를 가져오기 때문이다.

그림2(a)는 기준이미지에서의 세그먼트들과 disparity이고, 그림2(b)는 (a)의 세그먼트들을 disparity에 따라



[그림2] Warping operation

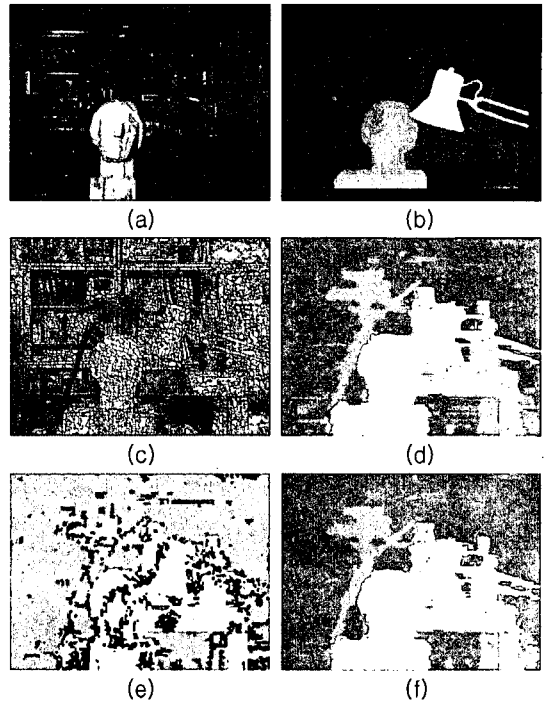
두 번째 이미지로 변환 시켰을 때의 이미지이다. 따라서 계산되는 비용은 그림2(b)에서 occlusion지역을 제외한 부분의 밝기 값의 차들의 합으로 계산된다. 이러한 Warping 연산에 따라 그림2에서 세그먼트B가 invalid 영역이고, 세그먼트A와 세그먼트C를 valid영역이라 하였을 경우, 찾고자 하는 범위를 줄이면서 정확한 결과 값을 얻을 수 있다.

3. 실험 및 결과

제안된 알고리즘을 실험하기 위해 우리는 Tsukba dataset(그림3(a))을 사용했다. 영상의 크기는 384X288 이다. 그림3(b)는 ground truth 이미지이고 그림3(c)는 세그먼테이션결과이다. 그림3(c)를 보면 실험을 위해 아주 작게 over-segmentation되어 있음을 볼 수 있다. 그림3(d)는 제안된 방법을 사용한 동적 계획법결과로 비교적 향상된 결과를 보여주고 있다. 그림3(e)는 세그먼트 영역의 분산을 이용하여 올바른 정합영역을 찾아내고, 올바르지 않은 영역은 검은색으로 나타냈다. 그림3(f)는 세그먼트 영역을 warping 연산기법을 이용한 방법을 사용하여 잘못된 영역의 disparity를 찾아낸 최종 결과이다. 그림3(f)에서 볼 수 있듯이 그림3(d)에서의 여러 부분이 상당히 줄어들었고, 기존의 동적 계획법에 비해서 많이 향상된 disparity image를 얻었다.

4. 결 론

본 논문에서는 동적 계획법과 영역 세그먼테이션을 이용한 새로운 스테레오 정합기법을 제안 하였다. 세그먼트 영역을 기준으로 하는 영역 윈도우를 계산 비용으로 사용하여 정확한 비용을 계산할 수 있도록 하였고, over-segment를 이용하여 앞서 얻은 disparity의 오류를 찾아내고 세그먼트 정합을 이용하여 정확한 disparity map을 찾아내었다. 우리가 제안한 방법으로 동적 계획법에서의 일반적인 스캔라인 사이의 불일치 문제와 occlusion영역에서의 오류에 대해 비교적 향상된 결과를 얻을 수 있었다. 특히, 이번 실험을 통해 비교적 적은 연산을 가지고 정확한 결과를 이끌어 낼 수 있었다. 앞으로 이런 hybrid기법에 대한 연구를 통해 좀 더 정확하면서 빠른 결과를 얻어낼 수 있는 방법에 대한 연구가 계속되어야겠다.



[그림3] (a)Tsukba dataset : Left image (b)Ground truth image (c)Segmentation image (d)Result image using proposed DP (e)Error extraction (f)Final result

참고문헌

- [1] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," IJCV, vol.47, no.1, pp. 7-42, 2002.
- [2] A.F. Bobick and S.S. Intile, "Large occlusion stereo," IJCV, Vol.33, 3, pp. 181-200, 1999
- [3] I.J. Cox, S.L. Hingorani, S.B. Rao, and B.M. Maggs, "A maximum likelihood stereo algorithm," Computer Vision and Image Understanding, Vol. 63, Issue 3, pp.542-567, 1996
- [4] H. Tao, H. S. Sawhney and R. Kumar, "A global matching framework for stereo computation," Proc. Int'l Conf. Computer Vision 2001.
- [5] M. Bleyer, M. Gelautz, "A layered stereo algorithm using image segmentation and global visibility constraints," Int'l Conf. Image Processing 2004.
- [6] T. Kanade, M. Okutomi, "A stereo matching algorithm with an adaptive window: theory and experiment," IEEE Trans. PAMI, Vol. 16, Issue 9, pp. 920-932, 1994.
- [7] S. Birchfield and C. Tomasi, "A pixel dissimilarity measure that is insensitive to image sampling," IEEE Tans. PAMI, Vol. 20, Issue 4, pp. 401-406, 1998