

다양한 GPU 플랫폼에서의 반전역적/전역적 스테레오 정합 병렬화 성능 분석

안드레 이반, 전운배, 박인규

인하대학교 정보통신공학과

{andreivan13@gmail.com, ybjeon0923@gmail.com, pik@inha.ac.kr}

요 약

본 논문에서는 지역적 정합부터 전역적 정합까지 다양한 범위의 스테레오 정합 알고리즘의 조합으로 유연하게 작동하는 GPU 기반의 병렬처리 프레임워크를 제안한다. 본 프레임워크는 안드로이드 OS 가 탑재된 스마트 폰, 리눅스 기반의 Odroid 임베디드 보드, PC 세 플랫폼에 OpenCL 로 구현되어 임베디드 GPU 부터 고성능 GPU 까지 대부분의 GPU 상에서 동작이 가능하다. 각 플랫폼 모두 CPU 와 GPU 중 선택한 프로세서로 프로그램이 동작할 수 있도록 구현되었다. 실험결과를 통해 각 플랫폼에서 다양한 알고리즘 조합에 따른 GPU 실행 시간과 정확도를 제시한다.

1. 서 론

지난 수십 년간 다양한 스테레오 정합 알고리즘들이 광범위하게 연구됐다. 기존 연구[1]에서는 다양한 스테레오 정합 알고리즘의 성능 평가를 제시하였다. 스테레오 정합의 구성은 (1) 정합 비용(matching cost) 계산 (2) 비용 집계(cost aggregation) (3) 변위(disparity) 계산 및 최적화(optimization) (4) 변위 개선(refinement)의 단계로 분류된다. 각 단계에서는 수행하는 기능에 따라 다양한 알고리즘을 포함하고 있으며 정확도와 계산량의 상보관계에 따라 특정 응용분야에 적합한 알고리즘을 선택한다.

스테레오 정합 알고리즘의 가장 큰 문제점은 계산 속도와 정확도 간의 절충(trade-off)을 정하는 것이다. 스테레오 정합 알고리즘은 변위 탐색 시 수많은 반복문을 돌기 때문에 상당한 비용을 요구한다. 또한 비용 집계(cost aggregation)와 최적화 부분은 알고리즘 수행 시간을 늘리고 복잡하게 만든다. 하지만 반복문의 대부분은 GPU 병렬화를 통해 고속으로 알고리즘을 수행할 수 있다. PC 환경에서 GPU를 활용하여 복잡한 스테레오 정합 알고리즘을 최적화한 기존의 연구들이 많이 존재하지만[2, 3], 하드웨어적 제한을 가진 기기에서 복잡한 스테레오 정합을 구현한 연구는 많지 않다.

본 논문에서는 우리는 기존에 존재하던 스테레오 정합 알고리즘의 조합을 기반으로 GPU 병렬화를 통한 고속 수행 결과를 제시한다. 본 논문에서는 모든 과정을 C 언어와 OpenCL 을 이용한 GPGPU 프로그래밍으로 서로 다른 세 GPU 플랫폼에 구현하였다. 기존 연구에서 제안한 알고리즘을 바탕으로 본 연구를 통해 제시하는 것은 아래와 같다.

- 계산 속도와 정확도를 상황에 맞게 고려하여 적용할 수 있는 스테레오 정합 프레임워크.

- 기존 스테레오 정합 알고리즘을 최신 임베디드 보드의 CPU/GPU 에 적용한 성능 분석결과.
- 스마트폰, 임베디드 보드(Odroid XU4) 그리고 PC 세 플랫폼에서 유연하게 적용될 수 있는 멀티 플랫폼 스테레오 정합 프레임워크.

2. 스테레오 정합 프레임워크

본론에서는 기존 연구[1]에서 제안한 다양한 알고리즘의 조합을 적용하여 전처리 과정, 정합 비용 계산, 변위 계산 및 변위 최적 개선 작업의 단계로 이루어진 GPU 병렬처리 기반의 프레임워크를 제안한다. 또한 단계별로 적용되는 알고리즘의 조합을 통해 정확하고 빠른 스테레오 정합 수행 방법을 제안한다. 본 논문에서 제안하는 프레임워크는 스테레오 카메라로부터 직접 촬영한 영상 데이터와 표준 스테레오 영상 데이터 모두에 대하여 성능을 측정하였다. 상기 알고리즘 전체 과정은 C 와 OpenCL 을 이용하여 CPU 와 GPU 에서 모두 구현되었다.

2.1 전처리 과정

전처리 과정에서는 스테레오 카메라의 좌/우 영상의 잡음(noise)을 처리하기 위해 5x5 크기의 가우시안 커널로 전처리 과정을 적용하였다. 가우시안 필터를 적용한 영상은 Gradient 필터 또는 Census 필터를 통해 더 정제될 수 있다. 두 필터는 스테레오 정합시 비용 계산 부분에서 텍스처가 없는 영역의 변위 오류 문제를 해결하는데 용이하다. Census 필터는 텍스처가 없는 영역을 다루는 데 좋지만, 병렬처리에는 용이하지 않으므로 변환에 시간을 많이 소요하게 된다. 따라서 전처리 과정의 시간적 비용을 줄이기 위해 본 연구에서는 Gradient 필터를 기본 전처리 필터로 활용하였다.

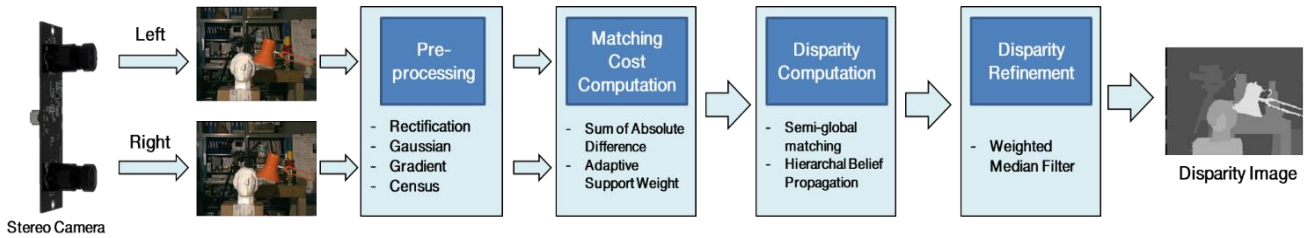


Fig. 1 스테레오 정합 프레임워크와 각 단계

2.2 정합 비용 계산

정합 비용계산을 위해 주로 사용되는 Sum of absolute difference (SAD)와 Adaptive support weight (ASW) [4] 두 가지 알고리즘을 차용하였다. 두 알고리즘 모두 윈도의 폭과 높이에 따라 4 개의 주 반복문을 영상의 너비와 높이에 각각 적용하여 실행한다. 본 연구에서는 GPU 에서 상위 2 개의 반복문을 병렬적으로 수행하도록 하였다.

2.3 변위 계산

변위 지도(disparity map) 계산을 위해 우리는 다양한 최적화 기법에 대해 다루었다. 본 연구에서는 3 가지 방법으로 최적화 기법을 제안하며 이는 지역적 기법(local), 반전역적 기법(semi-global), 전역적 기법(global)으로 구성된다. 지역적 기법에서는 비용 볼륨(cost volume)을 개선하기 위한 부가적인 과정을 적용하지 않고 winner-takes-all (WTA) 알고리즘을 적용하여 최종 변이 지도를 얻게 된다.

반전역적 최적화 기법으로 본 연구에서는 동적 프로그래밍 기반의 SGM[5] 알고리즘을 적용한다. SGM 은 제한조건을 통해 1 차원 화소 배열에서 모호하거나 잘못된 정합 비용 계산에 페널티를 부여한다. 반전역적 기법의 주요 원리는 2 차원 영상에서 다양한 방향의 1 차원 화소 배열의 변위 최적화를 수행하는 것이다. 비용 계산시 경로 최적화에 적합한 동적 프로그래밍은 최소 비용을 찾는 데 유용하다. 또한 SGM 기법은 계산 복잡도와 정확도 사이에서 균형 잡힌 절충점을 제공한다.

전역적 최적화 기법을 적용하면 지역적, 반전역적 기법에서 잡음이 많고(noisy) 텍스처가 없는(textureless) 영역의 부정확한 정합 비용 계산을 줄일 수 있다. 전역적 기법에서는 2 차원 영상 전체에서 최적화를 진행하는 것을 주요 과제로 삼는다. 하지만 2 차원 상의 최적화 작업은 상당한 비용과 복잡도를 요구하기 때문에 실시간으로 정합 비용을 계산하기 어렵다. 본 연구에서는 시간 비용을 줄이기 위해 hierarchical belief propagation (Hier-BP) [6] 알고리즘을 도입하였다. Hier-BP 는 피라미드형으로 구축된 영상 층을 반복적으로 최적화하여 전역적으로 정합 비용을 최적화한다. Hier-BP 는 각 영상 층에서 이웃 화소 간 신뢰 메시지(belief messages) 전파를 통해 최적화를 수행한다. 이러한 메시지 전파는 최

상층에서 최하층까지 N 번의 반복을 거쳐 수행된다. 상위층에서 메시지 전파가 끝난 융합된 화소는 불필요한 계산을 줄여 시간 비용적 이득을 볼 수 있게 한다. 또한 Hier-BP 는 계층적 구조 접근법을 통해 적은 반복으로 전역적 최소 비용(global minimum)에 도달할 수 있다.

2.4 후처리 기법

최종 변이 지도를 개선하기 위해 후처리 과정으로 weighted median filter(WMF) [7] 를 사용하였다. WMF 는 측정된 변이 지도를 기반으로 정합 비용을 개선하는 과정을 수행한다. 측정하려는 변이거리에서 비용은 1 로 주어지고 나머지는 0 으로 초기화한 뒤 각 변이 거리에 가중 필터(weighted filter)가 적용되는 방식이다. 본 연구에서 WMF 적용 시 윤곽선 검출 및 잡음 제거를 위해 bilateral 필터를 사용하였다. 그 외 domain transform filter 와 같은 윤곽선 검출 필터도 있지만 본 연구에서 제안하는 GPU 기반의 프레임워크에서는 복잡도만 증가하므로 제외하였다.

3. 실험 결과

본 실험은 Samsung Galaxy Note 8 (Mali-G71 MP20), Odroid XU4 (Mali-T628 MP4) 그리고 PC (Nvidia GTX 1080TI) 3 종의 다른 플랫폼으로 동일한 프레임워크 내에서 시행되었다. 본 논문에서 제안하는 프레임워크를 평가하기 위해 ground truth 를 포함한 기존 Middlebury 스테레오 데이터 셋[8] 를 활용하였다. 실험 결과의 수치적 지표로서 mean square error (MSE) 와 bad pixel percentage (BP %) 를 사용하였다.

그림 3 및 그림 4 를 통해 본 실험에 대한 정성적 평가를 확인할 수 있다. 결과를 통해 SAD 는 변위의 불연속적 구간에서 취약한 것을 확인할 수 있다. 반면 ASW 알고리즘은 윤곽선 보존에 뛰어나며 텍스처가 없는 연속 구간을 잘 처리하는 것을 확인할 수 있다. 반 전역적 기법의 SGM 과 전역적 기법의 Hier-BP 는 지역적 기법에 비해 잡음과 윤곽선 처리에 뛰어난 것을 확인할 수 있다. 결론적으로 ASW 와 Hier-BP 의 조합을 통해 가장 정확한 변이 지도를 추출할 수 있다.

정합 비용계산과 변위 최적화에 대한 정량적 평가는 Table 1 에서 확인할 수 있으며 SGM 과 Hier-BP 를 이용하여 최적화하였을 때 정확도가 상당히 향

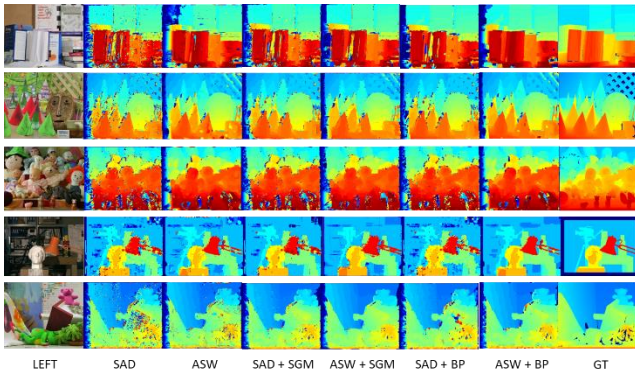


그림.3 제안된 프레임워크에서 MiddleBury 스테레오 데이터 셋을 이용한 정성적 평가. 영상 해상도는 각각 463x370, 450x375, 463x370, 384x288, 450x375.

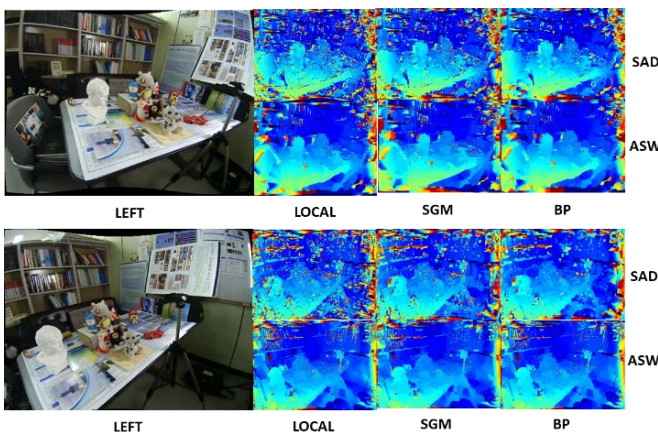


그림. 4 스테레오 카메라(영상 해상도: 640×480)를 이용한 정성적 평가.

상되는 것을 확인할 수 있다. 또한 비용 계산 방식에서 ASW 기법이 SAD 기법보다 높은 정확도를 얻는 것을 확인하였다. 본 실험에서 SAD와 ASW의 탐색 윈도우 크기는 각각 5×5와 25×25를 사용하였다. SGM의 경우 90도 간격의 4방향 탐색을 통해 반전역적 최적화를 수행하였다.

그림 4에서 보인 영상에 대해 세 가지 플랫폼의 GPU에서 측정된 계산 시간은 Table 2에서 확인할 수 있다. 이 때 정합 비용 계산 방식으로는 시간 복잡도가 낮은 SAD 기법으로 통일하여 사용하였다.

4. 결 론

본 논문에서는 고전적인 스테레오 정합 알고리즘부터 최근에 도입된 알고리즘을 CPU와 GPU에서 유연하게 동작하게 하는 OpenCL 기반의 프레임워크를 소개하였다. 또한 상용 스마트폰, 임베디드 보드, PC의 서로 다른 세 플랫폼에서 수행한 실험 결과를 제시하였다. 실험 결과를 통해 제안한 프레임워크가 수행 속도와 변위 지도의 정확도 사이에서 적절한 알고리즘을 선택하는데 유용한 정보를 제공함을 확인할 수 있다.

Table 1. 변위 지도의 정확도에 대한 정량적 평가

Algorithms	Average MSE	Average BP%
SAD	22.88	12.40
ASW	14.99	8.31
SAD+SGM	11.06	6.10
ASW+SGM	7.36	4.36
SAD+HierBP	6.07	5.92
ASW+HierBP	4.56	4.11

Table 2. 세 기기에서 측정된 계산 시간(단위: 초)

Algorithms	Desktop	Note8	Odroid
SGM	0.25	4.54	13.11
BP	0.54	5.01	19.40

감사의 글

이 논문은 2018년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (No.2017-0-00142, 스마트기기를 위한 온디바이스 지능형 정보처리 가속화 SW 플랫폼 기술 개발)

참고 문헌

- [1] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no.1-3, pp. 7-42, 2002.
- [2] Y. K. Choi, Williem, and I. K. Park, "Memory-efficient belief propagation in stereo matching on GPU," In *Proc. APSIPA Annual Summit and Conference*, pp. 1-4, 2012.
- [3] Y. K. Choi and I. K. Park, "Efficient GPU-based graph cuts for stereo matching," In *Proc. IEEE CVPR Workshops*, pp. 642-648, 2013.
- [4] K. J. Yoon and I. S. Kweon, "Adaptive support-weight approach for correspondence search," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.28, no. 4, pp. 650-656, 2006.
- [5] H. Hirschmuller "Stereo processing by semi-global matching and mutual information," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 30(2), 328-341, 2008.
- [6] Q. Yang, L. Wang, R. Yang, S. Wang, M. Liao, and D. Nister, "Real-time global stereo matching using hierarchical belief propagation." In *Proc. BMVC*, vol. 6, pp. 989-998, 2006.
- [7] Z. Ma, K. He, Y. Wei, J. Sun, and E. Wu, "Constant time weighted median filtering for stereo matching and beyond," In *Proc. CVPR*, 2013.
- [8] H. Hirschmüller and D. Scharstein. "Evaluation of cost functions for stereo matching," In *Proc. CVPR*, 2007.