스테레오 영상에서의 적응적 영역 가중치 매칭의 효율적 근사화 방법

김 호 영^{a)}, 이 성 원^{a)*}

An Efficient Approximation method of Adaptive Support-Weight Matching in Stereo Images

Ho-Young Kim^{a)} and Seong-Won Lee^{a)*}

요 약

영역기반 스테레오 매칭 분야에서 최근 인간의 시각체계(Human Visual System)에 기반하여 영역내의 밝기값과 거리값에 따라 적응적으로 가중치를 부여하는 적응적 영역 가중치(Adaptive Support-Weight, ASW) 방법이 좋은 매칭 결과를 보이고 있다. 그러나 이 방법은 좋은 매칭 결과에 비해서 많은 연산비용을 필요로 하게 되고, 매칭의 실시간 시스템화에 큰 장애요소로 작용한다. 이에 Bilateral filter 수식으로 근사화 후 Integral Histogram 기법을 적용하여 영역 윈도우의 크기에 상관없이 상수 시간 $O(1)$ 내에 매칭을 수행하는 연구가 진행되었다. 하지만 이 방법은 근사화 과정에서의 원 ASW 수식을 왜곡하기 때문에 매칭 정확도의 손실을 가져오게 된다. 본 논문에서는 적응적 영역 가중치 알고리즘의 매칭 정확도를 유지하면서 적응적 영역 가중치 알고리즘의 계산 비용을 줄이는 알고리즘을 제안한다. 이를 위해 영역내의 픽셀을 그룹화하여 근사화된 매칭을 수행하는 Sub-Block 방법과 영상의 에지 정보에 따라 적응적으로 시차 탐색 범위를 조정하는 방법을 제안한다. 결과적으로 제안된 기법은 기존 방식보다 좋은 매칭 정확도를 유지하면서도 효율적으로 계산 수행 시간을 줄이게 된다.

Abstract

Recently in the area-based stereo matching field, Adaptive Support-Weight (ASW) method that weights matching cost adaptively according to the luminance intensity and the geometric difference shows promising matching performance. However, ASW requires more computational cost than other matching algorithms do and its real-time implementation becomes impractical. By applying Integral Histogram technique after approximating to the Bilateral filter equation, the computational time of ASW can be restricted in constant time regardless of the support window size. However, Integral Histogram technique causes loss of the matching accuracy during approximation process of the original ASW equation. In this paper, we propose a novel algorithm that maintains the ASW algorithm's matching accuracy while reducing the computational costs. In the proposed algorithm, we propose Sub-Block method that groups the pixels within the support area. We also propose the method adjusting the disparity search range depending on edge information. The proposed technique reduces the calculation time efficiently while improving the matching accuracy.

Keyword : Fast Stereo matching, Adaptive-Support Weight, Computer Vision

a) 광운대학교 컴퓨터공학과

Dept. of Computer Engineering, Kwangwoon University

* 교신저자 : 이성원 (swlee@kw.ac.kr)

※ 본 논문은 교육과학기술부의 재원으로 한국연구재단(2011-0003501, 2010-0028178) 및 2011년도 광운대학교 연구년에 의하여 연구되었음.

· 접수일(2011년7월29일), 수정일(1차:2011년9월23일, 2차:10월19일), 게재확정일(2011년10월19일)

1. 서론

스테레오 비전(Stereo vision)은 인간의 여러 시각 기능 중 양안의 시차를 이용하여 공간의 입체를 재구성하는 대뇌에서의 일련의 과정이다. 이와 유사하게 컴퓨터 스테레오 비전(Computer stereo vision)은 한 장면에 대하여 서로 다른 위치에서 획득된 디지털 영상을 분석하여 시차정보를 추정하고, 최종 3D 깊이 정보를 추출하는 과정을 말한다. 이러한 스테레오 비전의 단계는 영상의 획득, 영상 보정(Calibration), 스테레오 매칭(Matching), 시차(Disparity) 추정, 깊이(Depth) 정보 추출 등의 단계^[1]로 이루어진다. 여기에서 매칭 과정은 깊이 정보 추출의 핵심 단계라 할 수 있다.

스테레오 매칭 방법에는 크게 전역적인 방법과 지역적인 방법으로 나눌 수 있다. 전역적인 방법^[2]은 최적화된 목적함수(Objective function, energy function)를 이용하여 전역적인 시차 표면(Disparity surface)을 구한다. 이와는 다르게 지역적인 방법은 지역 윈도우(Local window)를 이용하여 픽셀에 대한 시차를 구한다. 지역적인 방법에는 다시 특징 기반 방법(Feature-based method)^[3]과 영역 기반 방법(Area-based method)^[4]으로 분류할 수 있다.

특징 기반 방법에서 사용되는 정합요소는 영 교차점(Zero crossing), 경계선, 모서리, 마루, 골, 원추 곡선 등이 있다. 이 요소들에 기반한 결과는 정합 점이 정확하고 잡음에 강한 특징을 가지나, 정합 되는 점들이 적으므로 전체 영상의 시차 값을 구하기 위해서는 폐색 모델링(Occlusion modeling)과 시차 연속성(Disparity continuity) 등을 포함하는 어려운 내삽 과정을 필요로 하게 된다.

반면에, 영역 기반 방법은 밝기 정보에 기반하여 정합점을 찾는 방식이다. 이 방법은 영상의 전체 거리 정보를 얻을 수 있다는 장점이 있지만, 다양한 조명환경으로부터의 명도적 왜곡에 따라 민감하게 매칭 정확도가 달라지고, 지역 윈도우가 깊이 불연속적인 영역에 걸쳐 위치하게 되면 전경 평면화(Foreground-fattening) 현상을 얻게 된다. 이를 해결하기 위하여 각 픽셀에 대한 최적의 윈도우를 찾기 위한 적응적 윈도우 방법(Adaptive-window)^{[5][6]}과 다중 윈도우 방법(Multiple-window)^{[7][8]}이 제안되었다. 하지만 임의의 크기와 모양을 갖는 최적의 윈도우를 찾는 것은 NP-hard

문제^[9]에 직면하게 된다.

최근 들어 이를 해결하기 위한 대표적인 방법으로 영역 내의 중심 픽셀과 이웃 픽셀들 간의 컬러의 차이(Photo-metric difference)와 기하학적 차이(Geometric difference)를 라플라시안 모델 가중치를 사용하는 적응적 영역 가중치(Adaptive Support-Weight, 이하 ASW) 방법^[9]이 연구되었다. 그러나 이 방법은 좋은 매칭 결과를 보이지만 사용되는 영역 윈도우의 크기에 따라 기하급수적으로 계산량이 증가하는 단점을 보이고 있다. 이에 기존의 ASW방법의 정확도는 유지하면서도 높은 계산비용을 줄이기 위하여 Integral histogram 기법을 활용하여 좀 더 효율적으로 Bilateral filter 수식에 근사화한 방법^[10]을 ASW 근사화 연산에 사용하는 방법^[11]과 Sub-block 개념을 활용한 방법^[12]들이 연구되었다. 최근에는 좀 더 효율적으로 Bilateral filter를 구현하는 방법이 연구되었다^[13]. 하지만 이러한 접근 방법은 Bilateral filter 수식에 근사화하는 과정에서 원 ASW 수식을 근본적으로 왜곡하게 되어 많은 정확도 손실을 유발하게 된다. 이처럼 원 ASW 수식을 근사화하는 방법들과는 다르게 스테레오 각 영상에 대한 컬러값과 그래디언트(Gradient) 값을 결합(Combine)하여 적절한 필터링 후 폐색 영역에 대한 후처리를 수행하는 연구가 진행되었다^[14].

본 논문에서는 ASW 방법의 근사화된 매칭 방식을 통하여 좋은 매칭 정확도를 유지하면서도 효율적으로 ASW 방법의 많은 계산비용을 줄이는 2단계 매칭기법을 제안하고자 한다. 먼저, Sub-Block 단위의 가중치 계산 및 픽셀 밝기 값들의 평균치를 룩-업 테이블(Look-up table)로 미리 계산해두고 사용하는 방법과 참조 영상의 에지 정보를 기반으로 시차탐색과정에서의 중복탐색 비용을 줄여주는 적응적 시차탐색 방식을 제안한다.

본 논문의 II장에서는 기존의 적응적 영역 가중치(Adaptive Support-Weight, ASW) 방법의 소개와 함께 근사화된 방법으로 높은 계산비용을 줄이려는 Bilateral 접근 방식 및 FBS 방법을 소개하고, III장에서는 제안하는 근사화된 매칭을 수행하는 Sub-block 방식과 적응적 시차탐색 방식을 설명한다. IV장에서는 Middlebury 영상을 가지고 수행한 실험 결과와 함께 기존 알고리즘들과의 비교를 통

하여 각 측면에서의 성능을 검증하고, V장에서는 본 논문의 결론을 맺는다.

II. 기존 연구

1. 적응적 영역 가중치 방법 (Adaptive Support Weight Method)

적응적 영역 가중치 방법(ASW)은 인간의 시각체계(Human visual system)를 기반으로 하여 스테레오 영상의 대응점을 찾는 방식이다. 매칭 비용 계산과정에서 영상내의 기하학적(Geometric) 차이와 광도적(Photometric) 차이를 기반으로 가중치를 다르게 적용하는 방법이다. 여기에서 영역 가중치(Support-weight)는 형태학적 그룹화(Gestalt grouping)를 기반으로 하여 명도적 유사성(Similarity)과 픽셀위치 상의 근접성(Proximity)의 세기를 다음의 식 (1)과 식 (2)와 같이 라플라시안 가중치 함수로 표현할 수 있다.

$$W_I(p, q) = \exp\left(-\frac{\Delta I_{pq}}{\gamma_I}\right) \quad (1)$$

$$W_d(p, q) = \exp\left(-\frac{\Delta D_{pq}}{\gamma_d}\right) \quad (2)$$

여기에서 $W_I(p, q)$ 는 영역 윈도우 내의 중심 픽셀 p 와 임의의 픽셀 q 사이의 밝기값의 차이에 의한 가중치를, $W_d(p, q)$ 는 픽셀 p 와 픽셀 q 사이의 기하학적 거리에 의한 가중치를 나타낸다. 위의 식 (1)과 식 (2)를 이용하면, 대응점간의 비유사도(Dissimilarity)를 나타내는 매칭 비용(Matching cost)을 영역 윈도우(Support window)내의 영역 가중치들을 원 매칭 비용(Raw matching cost)에 적용하고 통합(Combining)함으로써 구할 수 있다. 이 때, 기존의 매칭 비용 계산과는 달리 참조 영상(Reference image)과 대상 영상(Target image)을 모두 고려한다. 만약 참조 영역 윈도우

우만 사용하게 되면, 대상 영역 윈도우가 다른 깊이를 갖는 곳에 위치하는 경우에 매칭 비용 계산에 있어서 문제가 발생한다. 따라서 픽셀들 간의 매칭 비용 계산은 양쪽의 영역 윈도우에 대한 영역 가중치를 통합한다. 즉, 픽셀 p 와 p_d 간의 매칭 비용은 아래 식 (3)과 같이 표현된다.

여기에서 $\bar{q}_d$ 와 $\bar{p}_d$ 는 참조 영역 윈도우 내의 픽셀 p 와 q 에 대하여 시차값 d 에 해당하는 대상 영역 윈도우 내의 픽셀들을 나타낸다. 그림 1은 이를 나타내고 있다.

그림 1 . 시차가 d 일 때 참조 영역 윈도우 (N_p)와 대상 영역 윈도우 (N_{p_d})
Fig. 1. Reference support window (N_p) and target support window (N_{p_d}) when $d_p = d$

다음 식(4)는 절대값을 이용한 픽셀 기반의 원 매칭 비용을 구하는 수식이다. 여기에서 $I_c(q)$ 는 참조 영상에서의 영역 윈도우내의 중심픽셀의 밝기값을, $I_c(\bar{q}_d)$ 는 시차값이 d 로 탐색중인 대상 영상에서의 해당 영역 윈도우내의 중심 픽셀의 밝기값을 나타낸다. 이 값들의 차값의 절대값에 의하여 원 매칭 비용을 계산하게 된다.

$$e(q, \bar{q}_d) = \sum_{c \in r, g, b} |I_c(q) - I_c(\bar{q}_d)| \quad (4)$$

매칭 비용 계산이 끝나면, 모든 픽셀에 대하여 최소값을

$$E(p, \bar{p}_d) = \frac{\sum_{q \in N_p, \bar{q}_d \in N_{p_d}} W_I(p, q) W_d(p, q) W_I(\bar{p}_d, \bar{q}_d) W_d(\bar{p}_d, \bar{q}_d) e(q, \bar{q}_d)}{\sum_{q \in N_p, \bar{q}_d \in N_{p_d}} W_I(p, q) W_d(p, q) W_I(\bar{p}_d, \bar{q}_d) W_d(\bar{p}_d, \bar{q}_d)} \quad (3)$$

선택하여 시차 값을 결정하는 식(5)와 같은 승자독식 (Winner-Takes-All, WTA) 알고리즘을 적용한다.

$$d_p = \arg \min_{d \in S_d} E(p, \bar{p}_d) \quad (5)$$

여기서 $S_d = \{d_{\min}, \dots, d_{\max}\}$ 는 탐색 시차 범위를 나타낸다.

최근 영역기반 매칭 방법의 연구 중에서 ASW 방법이 좋은 매칭 결과를 보이고 있다. 하지만 추정된 깊이 정보가 정확한 만큼 사용되는 매칭비용 함수가 복잡해 되고, 영역 가중치 윈도우의 크기 또한 증가하게 된다. 이에 따라 방문해야 하는 픽셀수가 기하급수적으로 늘어나게 되어 전체적인 매칭비용 계산의 증가를 가져오는 단점을 가지고 있다.

2. IH Bilateral (Integral Histogram Bilateral) 근사화 매칭 방법

Bilateral Filter 수식에 근사화 시키기 위하여 참조 영상에서의 가중치만을 고려하고 거리에 대한 가중치를 균등하게 고려하면 ASW 수식은 다음 식 (6)과 같이 간소화 될 수 있다^[11].

$$E(p, \bar{p}_d) = \frac{\sum_{q \in W_p, q_d \in W_{p_d}} W_I(p, q) e(q, \bar{q}_d)}{\sum_{q \in W_p} W_I(p, q)} \quad (6)$$

Integral Histogram은 Integral Image를 응용한 방법으로

Porikli^[10]가 Bilateral Filter를 상수 시간에 계산하는데 제안한 방법이다. 그림 2는 Integral Histogram의 생성과정을 나타내고, 관련 수식은 식 (7)과 같다.

$$H(x, y) = H(x, y-1) + H(x-1, y) - H(x-1, y-1) \quad (7)$$

각 픽셀 점마다의 히스토그램 빈도를 누적하여 저장하여 두기 때문에 식(8)을 통하여 주어진 영역에서의 영상 값들을 재계산 없이 구할 수 있다.

$$H(x_1 : x_2, y_1 : y_2) = H(x_2, y_2) + H(x_1, y_1) - H(x_1, y_2) - H(x_2, y_1) \quad (8)$$

식 (6)에 Integral Histogram 개념을 적용해보면 식 (9)와 같이 다시 나타낼 수 있다.

$$E(p, \bar{p}_d) = \frac{\sum_i W_I(p, i) h_{(I, \bar{I})}(i)}{\sum_i W_I(p, i)} \quad (9)$$

밝기값의 범위를 256개에서 64개의 bin으로 양자화 (Quantization)하여 구현하면 상당히 줄어든 상수 시간 $O(1)$ 내에 매칭 연산 과정을 수행 할 수 있다.

이와 같이 ASW 방법의 빠른 스테레오 매칭을 위해 Bilateral filter 수식으로 근사화 후 Integral Histogram 기법을 적용하여 영역 윈도우의 크기에 상관없이 상수 시간을

그림 2. Integral Histogram의 생성
Fig. 2. Integral Histogram construction

$O(1)$ 내에 매칭을 수행하는 Bilateral 근사화 매칭 방법의 연구가 진행되었다. 하지만 이 방법은 Bilateral Filter 수식에 근사화 시키기 위하여 참조 영상에서의 가중치만을 고려하고 거리에 대한 가중치를 균등하게 고려하기 때문에 복잡한 매칭비용 수식의 근사화 과정에서 원 ASW 수식을 왜곡하게 되는 단점을 가지고 있다.

3. FBS 근사화 매칭 방법

FBS(Fast Bilateral Stereo) 방법^[12]은 ASW 방법이 Bilateral Filter의 원리와 유사한 점을 근거로 하여 잡음(Noise)에 좀 더 강인하면서도 효율적인 계산을 위한 Cost aggregation strategy을 제안한 방법이다. 간단하면서도 효율적인 서브-블록(Sub-block)화 개념을 적용하여 근사화된

그림 3 . 제안된 서브-블록 개념

Fig. 3. Proposed sub-block concept

매칭을 수행하는 이 방법은 서브-블록의 중심점(Cental

point)에만 기하학적 가중치를 적용함으로써 서브-블록내의 모든 픽셀에 방문하여야 하는 계산비용을 줄일 수 있지만 원 ASW 수식을 왜곡하는 근본적인 문제점을 여전히 가지게 된다. 그림 3은 이 알고리즘의 블록화 기법을 나타내고 있다. 여기서 p_r 은 참조영상에서의 $W \times W$ 크기의 영역 윈도우내의 중심점을 나타내고, p_t 은 대상영상에서의 중심점을 나타낸다. $w \times w$ 는 서브-블록의 크기를 나타내고, $b_r(u,v)$, $b_t(u,v)$ 는 각각 참조영상과 대상영상에서의 블록들을 나타낸다.

식 10은 참조 영상에서의 밝기값에 의한 가중치 계산 수식을 나타내고 있다. 각 서브-블록의 평균 밝기값과 영역 윈도우의 중심점과의 유클리디안 거리에 의한 가중치값을 계산하는 과정이다. 이 수식은 대상 영상에서도 똑같이 적용된다.

$$W_C(I_r(p_r), \bar{I}_r(b_r(u,v))) = \exp\left(-\frac{\|I_r(p_r) - \bar{I}_r(b_r(u,v))\|}{\gamma_c}\right) \quad (10)$$

III. 제안하는 적응적 영역 가중치 매칭 방식

본 논문에서는 적응적 영역 가중치 알고리즘의 매칭비용 수식을 유지하여 좋은 매칭 정확도는 유지하면서 영역 윈도우 크기가 증가함에 따른 ASW 방법의 많은 계산 비용을 효율적으로 줄일 수 있는 알고리즘을 제안하고자 한다. 먼저, 영역내의 픽셀을 서브-블록 단위로 그룹화하여 영역내의 가중치 값들의 평균값을 룩업 테이블로 미리 저장하여

그림 4. 제안하는 알고리즘의 블록 다이어그램

Fig. 4. Block Diagram of the Proposed Algorithm

두고, 근사화된 매칭을 수행하는 Sub-Block 방법과 참조 영상의 각 픽셀 위치에서의 에지 정보를 참조하여 적응적으로 시차 탐색 범위를 조정하는 ADS 매칭 방법을 제안한다.

그림 4는 제안하는 알고리즘의 전체적인 흐름도를 나타낸 것이다. 1 단계에서는 ASW 매칭의 계산량을 줄이기 위한 준비과정으로 영역내의 픽셀을 그룹화하여 근사화된 매칭을 수행하기 위해 3x3 Sub-Block 단위로 영역 가중치와 픽셀 밝기값들의 평균치를 미리 계산하여 룩-업 테이블 메모리로 저장하여 둔다. 또한, 불필요한 시차탐색범위를 줄여서 불필요한 계산량의 감소를 가져오는 ADS 방법을 위하여 참조 정보로 사용될 에지 정보를 미리 검출해 둔다. 2 단계에서는 1단계에서 미리 저장해둔 가중치 값들의 테이블과 검출해둔 에지 정보를 참조하여 적응적으로 시차 탐색 범위를 조정하여 매칭비용을 계산하고 최소값에 대응하는 위치로 시차값을 결정한다. 최종적으로 결정된 시차 값들에 정규화(Normalization) 과정을 통하여 시차지도 영상을 얻게 된다.

1. Sub-Block 매칭 방법

기존의 ASW 방법은 중심 픽셀로부터 영역 윈도우 내부의 모든 픽셀에 방문하여 가중치를 계산하는 과정으로 동작한다. 이는 영역 윈도우의 크기가 증가함에 따라 기하급수적으로 매칭 연산량을 증가시키는 부담을 가져오게 된다. Sub-Block 방식은 영역 윈도우의 크기 증가에 따른 ASW

방법의 많은 계산량을 감소시키기 위한 대안이다.

매칭비용 계산 시에 필요한 영역 윈도우내의 값들을 시차값에 따라 3x3 크기의 서브-블록 단위로 그룹화 하고 그 대표값들을 미리 룩-업 테이블로 계산해 두었다가 매칭비용 계산 시에 사용한다. 이렇게 미리 계산된 값들은 매칭비용 계산 시에 각 시차값에 따른 서브-블록 단위로 줄어든 탐색범위에서 각각 참조영상과 대상영상에 해당되는 평균 가중치값과 평균 원매칭 비용과의 곱의 형태로 간단히 간략화 될 수 있다. 해당 수식은 식 (11)과 같다.

여기서 $\overline{W1}_{(p,q)}(m)$ 은 참조영상에서의 영역 (p,q)에 해당하는 각 서브-블록 m에서의 평균 가중치를 나타내고, 대상영상에서의 $\overline{W2}_{(p_d,q_d)}(m)$ 은 시차값 d에서의 영역 (p_d,q_d) 에 해당되는 각 서브-블록 m에서의 평균 가중치를 나타낸다. 평균 원매칭 비용을 나타내는 $|\overline{L}_{(p,q)}(m) - \overline{L}_{(p_d,q_d)}(m)|$ 항은 평균 밝기값의 차이로 계산됨을 나타낸다.

그림 5는 9x9 Window의 경우 81개의 픽셀을 방문해야 하지만, Sub-Block 개념 적용 후 방문해야 할 픽셀의 개수가 9개의 대표 블록값으로 대체됨을 보여준다.

본 방식은 크게 2가지 과정으로 나눌 수 있다. 첫 번째는 좌우 영상의 모든 픽셀에 대하여 영역 윈도우 크기만큼의 주변 픽셀 가중치를 계산하고 Sub-block 단위로 가중치와 픽셀 밝기값의 평균값을 저장한다. 이때 밝기값의 차이에

$$E(p, p_d) = \frac{\sum_{m \in block} \overline{W1}_{(p,q)}(m) \overline{W2}_{(p_d,q_d)}(m) |\overline{L}_{(p,q)}(m) - \overline{L}_{(p_d,q_d)}(m)|}{\sum_{m \in block} \overline{W1}_{(p,q)}(m) \overline{W2}_{(p_d,q_d)}(m)} \quad (11)$$

그림 5. 9x9 영역 윈도우와 제안된 3x3 Sub-Block 그룹화의 결과

Fig. 5. 9x9 Support Window and result of proposed 3x3 Sub-Block grouping

따른 가중치 $f_s(\Delta c_{pq})$ 와 기하학적 차이에 의한 가중치를 나타내는 $f_p(\Delta g_{pq})$ 의 값은 모든 영역에 대하여 고정적이기 때문에 미리 룩-업 테이블(Look-up table)로 저장하여 연산의 중복을 피한다. 그림 6은 $f_s(\Delta c_{pq})$ 와 $f_p(\Delta g_{pq})$ 의 가중치 그래프를 나타낸다.

2. 에지정보 기반의 적응적 시차탐색 방식 (Adaptive disparity search)

스테레오 매칭 결과의 지표인 Ground-truth 영상의 시차지도 값들이 에지(Edge) 정보의 유무를 기준으로 새로운 콘텐츠 영역을 시작하기 전에는 시차값이 크게 변하지 않는다는 점에 착안하여 참조영상의 콘텐츠에 따른 에지정보를 기준으로 적응적으로 감소된 시차범위를 탐색(감소된 윈도우의 이동)하는 방식이다. 여기서 영역 윈도우의 크기는 변하지 않

는다. 이는 가급적 매칭 정확도의 손실은 가져오지 않으면서도 효율적으로 중복탐색연산을 줄이고자 하는 방법이다.

그림 7은 ADS 방식을 나타낸다. 탐색중인 현재 픽셀위치에서 에지 정보가 발견되지 않은 경우, 기존 콘텐츠 영역에 머무르고 있으므로 그 시차변화는 적을 것이 예상된다. 고로, 전체시차탐색범위로부터 감소된 범위의 시차를 탐색하게 된다. 이와는 다르게 에지정보가 발견된 경우라면, 새로운 콘텐츠의 시작을 나타내고, 그 시차값을 예상할 수 없으므로 기존의 전체시차범위의 탐색을 수행하여 좋은 매칭 정확도를 유지하는 방법이다.

이 방법이 좋은 매칭 정확도를 유지하기 위해서는 정확한 에지정보를 구하는 것이 우선 중요하다. 본 논문에서는 Canny Edge Detection^[15]을 수행하여 매칭 과정 이전에 에지정보를 미리 파악해두었다. 이 방법은 노이즈 제거(가우시안 스무딩), 기울기 계산(크기, 방향), 비최대 억제(Non-

그림 6 . 밝기값과 근접성에 따른 가중치 값의 변화도
Fig. 6. Weight value transition graph by Intensity and proximity

그림 7 . 적응적 시차 탐색 알고리즘
Fig. 7. Adaptive disparity search algorithm

maximum suppression), 에지 연결성 분석(Hysteresis) 등의 과정을 통하여 좋은 에지 정보를 추출하게 된다.

$$\begin{aligned}
 &(\sigma: \text{적응적탐색범위}, T: \text{에지결정 문턱치 값}, d_{(p-1)}: \\
 &\text{이전 픽셀의 시차값}) \\
 &S_{d_1} = \{d_{(p-1)} - \sigma, \dots, d_{(p-1)} + \sigma\} \\
 &S_{d_2} = \{d_{\min}, \dots, d_{\max}\} \\
 &d_p = \begin{cases} \arg \min_{d \in S_{d_2}} E(p, \overline{p_d}), f(x, y) > T \\ \arg \min_{d \in S_{d_1}} E(p, \overline{p_d}), f(x, y) \leq T \end{cases} \quad (12)
 \end{aligned}$$

식 (12)은 에지 정보를 이용하여 적응적인 범위에서 시차 탐색을 수행하는 ADS 방법의 수식이다. 해당 픽셀위치에서의 영상값이 에지 문턱치 값 T 보다 작거나 같으면 이전 픽셀의 시차보다 줄어든 적응적 시차탐색 범위 S_{d_1} 에서 시차탐색을 수행하고, T 값 보다 크면, 원래의 전체 시차탐색 범위 S_{d_2} 에서 시차탐색을 수행한다. d_p 는 이러한 과정으로 얻어진 최종 시차값을 나타낸다.

IV. 실험 결과 및 고찰

1. 사용한 실험 환경

제안된 방식의 실험에는 2.93GHz의 CPU Intel Core i3

와 4.00GB 메모리를 가진 컴퓨터가 사용되었다. 실험 영상으로는 Middlebury^[16]에서 제공하는 Tsukuba(384x288), Venus(434x383), Teddy(450x375), Cones(450x375) 영상들을 사용하였다. 매칭 정확도 평가에는 Ground Truth 영상을 사용하였다. 매칭과정에서의 영역 윈도우 크기는 33x33을 사용하였고, FBS 방법에서는 39x39 윈도우를 사용하였다. Sub-Block의 크기는 3x3, 적응적 시차탐색 방식의 줄어든 탐색 범위 $\alpha = \pm 3$, Canny edge detection의 이중 임계값은 (Threshold_L=10, Threshold_H=50) 실험적으로 구하여 사용하였다. 잘못 매칭된 픽셀 비율을 나타내는 PBMP(Percentage of Bad Matching Pixels)의 임계값으로는 $\delta = 1.0$, 밝기값 가중치 함수의 인자 γ_I 값은 14.0, FBS의 경우엔 23.0으로 실험을 진행하였다. 기하학적 가중치 함수의 인자 값 $\gamma_d = \text{WINDOWSIZE}/2$, FBS의 경우엔 14.0으로 고정하여 실험하였다.

2. 계산 비용 측면 비교

그림 8은 각 매칭방법에서의 각 알고리즘에 따른 계산 수행 시간을 보여주고 있다. 다소 차이가 있지만 ASW방법에서 가장 많은 수행시간이 필요함을 보여주고 있다. 그리고 앞에서 언급한 바와 같이 실험 영상의 크기가 커짐에 따라 매칭을 수행하기위한 계산량이 기하급수적으로 증가 되는 문제점을 잘 보여주고 있다. IH Bilateral 방법의 경우

그림 8. 수행시간 비용 비교
Fig. 8. The comparison of execution time cost

가장 적은 수행시간으로 Tsukuba 영상의 경우 4.12 sec의 빠른 수행시간을 보이고 있다. FBS방법과 Sub-block 방법의 경우 거의 유사한 수행속도를 보이고 있지만, 영상의 크기와 시차탐색 범위가 커지는 Teddy와 Cones에서는 오히려 Sub-block 방식이 더 빠른 수행시간을 보여주고 있다. 이는 매칭에 필요한 연산을 룩-업 테이블 제작 시 미리 계산해두기 때문에 매칭 비용을 계산하는 부분의 복잡도가 상당히 줄어들기 때문이다. 여기에 추가 적용된 ADS 방법은 Sub-block의 결과에서 더 빨라진 수행시간을 보이게 됨으로써 더 효율적인 계산방식임을 확인할 수 있다. 이 비교에서 IH Bilateral 방법이 가장 빠른 수행시간과 영역원도우의 크기에 상관없이 상수시간 매칭의 장점을 보이고 있지만 표 1에서 알 수 있듯이 큰 정확도 손실을 동반하고 있음을 알 수 있다.

3. 메모리 사용량 측면 비교

그림 9는 각 근사화 방법의 룩-업 테이블 제작에 따른 메모리

리 자원 사용량의 비교를 나타내고 있다. Tsukuba 영상의 경우, ASW 방법을 비롯하여 모든 매칭방식에서 기본적으로 384x288 크기의 원 영상 2개(스테레오 영상)와 시차지도 결과영상, Ground-truth 영상을 저장하기 위한 $4 \times (0.1 \text{ MB}) = 0.4 \text{ MB}$ 의 메모리가 사용되었다.

Tsukuba영상을 사용한 IH Bilateral 방법의 경우엔 영상 밝기값들을 저장하기 위해 2 byte 크기의 unsigned short type의 메모리를 사용하였다. 참조 영상에서 영역내의 각 픽셀들과 중심 픽셀간의 차이를 Integral Histogram으로 구성하기 위해서 16개의 시차값과 384x288 영상 픽셀 위치 및 양자화된 64개의 밝기 빈도값들을 저장하기 위해 $2 \times (16 \times 384 \times 288 \times 64) = 216 \text{ MB}$ 의 메모리 자원이 필요로 하게 된다. 또한, 원매칭 비용에 대한 Integral Histogram 테이블을 구성하기 위해 $2 \times (384 \times 288 \times 64) = 13.5 \text{ MB}$ 의 메모리가 추가로 필요로 하였다. Sub-Block 방법에서는 소수점 계산을 위해서 8 byte 크기의 double type의 메모리를 사용하였고, 33x33 윈도우 크기에 대해서는 3x3 Sub-Block으로 그룹화 하여 계산하게 되므로 11x11 크기의 테이블이 필요

그림 9. 메모리 사용량 비교

Fig. 9. The comparison of memory cost

로 하고 각 블록들로 구성된 평균 가중치 값들과 밝기값의 평균값들을 저장하기 위해서 $2 \times (8 \times 11 \times 11) = 1936$ byte가 필요하게 된다. 이러한 구조로 384×288 참조 영상, 대상 영상 2장에 대하여 계산하므로 $2 \times (784 \times 384 \times 288) = 408.3$ MB의 메모리가 사용되었다. 마지막으로 ADS 방법의 경우엔 에지 영상을 추가로 저장하기 위해서 0.1 MB 메모리가 더 사용되었다.

ASW방법의 경우, 가정 적은 메모리를 사용하고 있지만 앞에서 논의한 것 같이 가장 많은 수행시간을 필요로 하게 된다. FBS 방법의 경우엔 39×39 영역 윈도우를 사용하기 때문에 33×33 윈도우를 사용하는 Sub-block 방식보다 더 많은 메모리 자원을 사용하게 된다. 하지만, 다음 논의에서 알 수 있듯이 Sub-block 방식이 정확도 측면에서도 더 우수함을 확인할 수 있다. IH Bilateral 방법의 경우엔 전체 영상의 크기가 커짐에 따라 그 메모리의 사용량도 가장 크게 증가하는 것을 볼 수 있다. 이는 Integral Histogram을 사용하는 단점이다.

4. 매칭 정확도 측면 비교

그림 10은 각 매칭방법들에 대한 매칭 정확도의 오차 수치(객관적 화질평가)를 RMS(Root Mean Squared) 오차와 PBMP 비율로 나타내고 있다. 근사화 과정에 의한 원 ASW 수식의 왜곡이 없는 ASW 방법이 가장 좋은 매칭 정확도를 보여주고 있다. 제안하는 Sub-Block 방법과 ADS 방법이 기존 근사화 방법인 IH Bilateral 방법과 FBS 방법 보다 더 적은 매칭 정확도 손실을 보여주고 있다. 이는 IH Bilateral 방법이 대상 영상은 생략하고 참조 영상에서만 가중치를 부여하여 적용하고 거리에 대한 가중치를 균등하게 두고 있기 때문이다. 또한, FBS 방법의 경우 각 서브-블록의 모든 픽셀에 대하여 기하학적 가중치를 적용하지 않고 중심 점에 대해서만 가중치를 적용하였기 때문이다. 그림 11의 주관적인 화질평가로도 제안하는 방법들이 더 좋은 정확도 유지 성능을 보임을 알 수 있다.

그림 10. 매칭 정확도 측면 비교

Fig. 10. The comparison of matching accuracy

	Tsukuba 384x288	Venus 434x383	Teddy 450x375	Cones 450x375
Ref.
GT
ASW
IH
FBS
Sub-block
ADS

그림 11. 상위부터 아래로: 참조 영상, groundtruth 영상, 각 알고리즘들에 의한 시차 지도

Fig. 11. From top to bottom: reference image, groundtruth (GT), disparity map yielded by each algorithms

5. 계산량, 메모리, 정확도간 Trade-off

표 1은 각 알고리즘의 계산수행시간 측면, 메모리 사용량 측면, 매칭 정확도 측면에 대한 실험 결과를 종합적으로 나타내고 있고, 그림 12는 각 성능평가 측면에서의 변화 추이 관계를 보여주고 있다. 계산수행시간 측면에서는 FBS 방법이 전체 영상의 크기가 커짐에 따라 그 계산수행시간이 갑자기 증가하는 모습을 보이고 있다. 메모리 사용량 측면에서는 IH Bilateral 방법이 Integral Histogram을 사용하면서

매칭 수행속도가 상당히 빨라진 만큼 메모리 사용량이 매우 커졌다. 또한, 39x39 윈도우를 사용하는 FBS 방법이 33x33 윈도우를 사용하는 Sub-block이나 ADS 방법보다 더 많은 메모리를 사용함을 확인할 수 있다. 정확도 유지 측면에서는 IH Bilateral 방법이 가장 손실이 큰 수치를 보이고 있고, 제안하는 Sub-block 방법이 원 ASW 정확도 수치를 가장 잘 보존하고, 다음으로 ADS 방법이 잘 보존함을 확인할 수 있다.

표 1. 각 알고리즘의 계산수행시간, 메모리 사용, 정확도 평가 실험 결과

Table 1. Experimental results of each algorithm's execution time, memory using, accurate evaluation

		Tsukuba (384x288)					Venus (434x383)					Teddy (450x375)					Cones (450x375)				
Diparity		16					20					60					60				
Scale		16					8					4					4				
Algorithm		ASW	IH	FBS	Sub	ADS	ASW	IH	FBS	Sub	ADS	ASW	IH	FBS	Sub	ADS	ASW	IH	FBS	Sub	ADS
계산 수행 시간 (sec)	Edge 검출 (sec)					0.004					0.005					0.006					0.007
	Weight LUT (sec)		1.56 0.08	2.50	4.95			2.97 0.13	3.92	8.20			9.11 0.13	3.90	8.02			9.11 0.13	3.90	8.02	
	Matching Cost (sec)	38.69	2.47	5.81	3.70	2.22	77.75	4.77	11.32	7.41	3.76	209.72	13.23	31.99	20.55	8.00	209.72	13.23	31.99	20.55	9.12
	Total (sec)		4.12	8.31	8.65	7.174		7.87	15.24	15.66	11.965		22.47	35.89	28.57	16.026		22.47	35.89	28.57	17.167
사용된 메모리 자원 (MB)	LUT	$2^*(8^*(3^*(33))) = 0.016$	$2^*(16^*(384^*(2^*(88^*(64)))) = 216$	$2^*(2^*(8^*(13^*(1^*(3^*(384^*(288)))))) = 570.3$	$2^*(2^*(8^*(11^*(1^*(11^*(384^*(288)))))) = 408.3$		$2^*(8^*(3^*(33))) = 0.016$	$2^*(20^*(434^*(3^*(83^*(64)))) = 405.8$	$2^*(2^*(8^*(13^*(1^*(3^*(434^*(383)))))) = 857.3$	$2^*(2^*(8^*(11^*(1^*(11^*(434^*(383)))))) = 613.8$		$2^*(8^*(3^*(33))) = 0.016$	$2^*(60^*(450^*(3^*(75^*(64)))) = 1235.9$	$2^*(2^*(8^*(13^*(1^*(3^*(450^*(375)))))) = 870.3$	$2^*(2^*(8^*(11^*(1^*(11^*(450^*(375)))))) = 623.1$		$2^*(8^*(3^*(33))) = 0.016$	$2^*(60^*(450^*(3^*(75^*(64)))) = 1235.9$	$2^*(2^*(8^*(13^*(1^*(3^*(450^*(375)))))) = 870.3$	$2^*(2^*(8^*(11^*(1^*(11^*(450^*(375)))))) = 623.1$	
		$8^*(256) = 0.002$	$2^*(384^*(288^*(64))) = 13.5$				$8^*(256) = 0.002$	$2^*(434^*(383^*(64))) = 20.2$				$8^*(256) = 0.002$	$2^*(450^*(375^*(64))) = 20.5$				$8^*(256) = 0.002$	$2^*(450^*(375^*(64))) = 20.5$			
	Image data	0.1 x4	0.1 x4	0.1 x4	0.1 x4	0.1 x5	0.1 x4	0.16 x4	0.16 x4	0.16 x4	0.16 x5	0.1 x4	0.16 x4	0.16 x4	0.16 x4	0.16 x5	0.1 x4	0.16 x4	0.16 x4	0.16 x4	0.16 x5
	Total	0.418	229.9	570.7	408.7	408.8	0.418	426.64	857.94	614.44	614.6	0.418	1257.04	870.94	623.74	623.9	0.418	1257.04	870.94	623.74	623.9
정확도 평가	RMS	18.655	29.708	23.154	18.604	19.262	14.173	20.718	18.144	14.831	15.572	33.076	42.126	35.705	32.645	32.279	33.574	41.286	37.261	32.471	33.066
	PBMP (%)	5.643	11.092	6.954	5.352	5.642	4.478	8.119	6.080	5.218	5.469	11.768	21.792	13.490	12.021	11.783	12.010	20.619	13.650	11.715	12.186

그림 12. Trade-off 비교

Fig. 12. The comparison of trade-off

V. 결 론

본 논문에서는 기존 ASW 방법의 좋은 매칭 정확도는 유지하면서도 영역 윈도우의 크기가 증가함에 따라 지수적으로 증가하는 계산비용을 줄이기 위한 Sub-Block과 ADS 방법을 제안하였다. 실험결과를 통하여 각 근사화 방법에 따라 그 실험결과가 다양함을 보여주었다. Tsukuba 영상을 사용한 경우, 기존의 IH Bilateral 방법이 가장 적은 메모리를 사용하면서도 가장 빠른 상수시간 $O(1)$ 매칭수행 능력을 보여주었다. 하지만 추가적으로 매칭 정확도 손실의 측면까지 고려한다면 매칭 정확도 손실이 가장 크다는 것을 알 수 있었다. 또한, FBS 방법은 더 향상된 매칭 정확도 유지와 효율적인 계산 수행 결과를 보여 주었다. 하지만 이 방법들은 근사화하는 과정에서 원 ASW 수식을 왜곡하거나 이론적인 축약을 포함하는 단점을 가지게 된다. 이에 제안하는 Sub-Block방법과 ADS 방법을 통한 매칭과정은 비

교 실험을 통하여 기존 근사화 방법들 보다 매칭 정확도의 손실은 적으면서도 동시에 효율적으로 계산 수행 시간을 감소시키는 것을 확인하였다.

하지만 제안하는 방법은 사용되는 영역 윈도우의 크기가 커짐에 따라 Sub-Block으로 그룹화하여 가중치 값들을 미리 계산해 두는 과정에서 록-업 테이블들이 많이 사용되어야 하기 때문에 기존의 IH Bilateral 방식보다 더 많은 메모리 자원을 요구하게 되었다. 향후 실시간 시스템에서 적용 가능한 매칭 방법의 실현을 위해서는 GPU를 활용한 병렬 처리에 대한 연구와 매칭 정확도의 추가적인 보완을 위한 후처리 방법들이 추가로 연구되어야 할 것이다.

참 고 문 헌

- [1] R. I. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision", Cambridge University Press, 2004.
- [2] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense

- Two-Frame Stereo Correspondence Algorithm,” Int’l J. Computer Vision, vol.47, no.1, pp.7-42, 2002.
- [3] K. Czarnecki, S. Helsén, “Feature-based survey of model transformation approaches,” IBM Systems Journal, vol.45, no.3, pp.621-645, 2006.
- [4] C. Young-Sheng, H. Yi-ping, F. Chiou-Shann, “Fast Block Matching Algorithm Based on the Winner-Update Strategy,” IEEE Trans. Image Processing, vol.10, no.8, pp.1212-1222, Aug. 2001.
- [5] K. Prazdny, “Detection of Binocular Disparities,” Biological Cybernetics, vol.52, no.2, pp.93-99, 1985.
- [6] Y. Xu, D. Wang, T. Feng, and H.-Y. Shum, “Stereo Computation using Radial Adaptive Windows,” In Proc. Int’l Conf. Pattern Recognition, vol.3, pp.595-598, 2002.
- [7] A. Fusiello, V. Roberto, and E. Trucco, “Efficient stereo with multiple windowing,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 858-863, 1997.
- [8] S. B. Kang, R. Szeliski, and C. Jinxjang, “Handling occlusions in dense multi-view stereo,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol.1, pp.103-110, 2001.
- [9] K. J. Yoon and I. S. Kweon, “Adaptive Support-Weight Approach for Correspondence Search”, IEEE Trans. PAMI, 28(4):650-656, 2006.
- [10] F. Porikli, “Constant Time O(1) Bilateral Filtering”, In Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.1-8, 2008.
- [11] 주명호, 강행봉 “빠른 스테레오 매칭을 위한 Bilateral 접근 방법”, 대한전자공학회 논문지, 제46권, 제1호, pp.136-143, 2009.1
- [12] S. Mattoccia, S. Giardino, A. Gambini, "Accurate and efficient cost aggregation strategy for stereo correspondence based on approximated joint bilateral filtering", Asian Conference on Computer Vision (ACCV2009), pp.371-380, 2009.
- [13] Y. Qingxiong, T. Kar-Han, N. Ahuja, "Real time O(1) bilateral filtering", IEEE conf. on CVPR, pp.557-564, 2009.
- [14] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, M. Margrit, "Fast cost-volume filtering for visual correspondence and beyond", IEEE conf. on CVPR, pp.3017-3024, 2011.
- [15] John Canny, “A computational approach to edge detection,” IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-8(6):679 - 698, Nov. 1986.
- [16] <http://vision.middlebury.edu/stereo>

저 자 소 개

김 호 영

- 2011년 2월 : 광운대학교 컴퓨터공학과 (학사)
- 2011년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 (석사과정)
- 주관심분야 : 영상신호처리, 신호처리 SoC, Parking Assistant System

이 성 원

- 1988년 2월 : 서울대학교 제어계측공학과 (학사)
- 1990년 2월 : 서울대학교 제어계측공학과 (석사)
- 2003년 5월 : Univ. of Southern California (박사)
- 1990년 2월 ~ 2004년 2월 : 삼성전자 반도체부문 시스템 LSI 책임연구원
- 2005년 3월 ~ 현재 : 광운대학교 컴퓨터공학과 부교수
- 주관심분야 : 영상신호처리, 신호처리 SoC, 컴퓨터구조