

가우시안 잡음과 계산량을 고려한 하이브리드 센서스 변환

정성환*, 강성진¹

¹한국기술교육대학교 전기전자통신공학과

Hybrid census transform considering gaussian noise and computational complexity

Seong-Hwan Jeong^{1*} and Sung-Jin Kang¹

¹Department of Electronics and Communication Engineering, Korea University of Technology and Education

요 약 스테레오 매칭 중 센서스 변환은 방사 왜곡과 밝기 변화에 강한 특징이 있다. 본 논문은 미니 센서스 변환과 일반화된 센서스 변환을 동시에 이용한 하이브리드 센서스 변환을 제안하였다. 제안한 하이브리드 센서스 변환 방법은 미니 센서스 변환의 적은 계산량과 일반화된 센서스 변환의 잡음에 강인한 특성을 반영하여 설계되었다. 성능을 평가하기 위하여 후처리 과정까지 포함하여 스테레오 매칭을 수행하였다. 그 결과 하이브리드 센서스 변환은 일반화된 센서스 변환과 성능이 비슷하였고, 계산량은 미니 센서스 변환과 일반화된 센서스의 중간값을 갖는다.

Abstract Census transform is one of the stereo vision methods which is robust to radiometric distortion and illuminance change. This paper proposes a hybrid census transform using the mini census transform and the generalized census transform concurrently. This method uses simplicity of mini census transform and noise feature of generalized census transform together. This paper performed stereo matching containing post processing to evaluate each methods. The result shows that hybrid census transform has similar performance to generalized census transform and mean value of calculation complexity between mini census transform and generalized census transform.

Key Words : Census transform, Generalized census, Hybrid census, Mini census, Stereo matching

1. 서론

1.1 연구배경

스테레오 매칭은 양안시차를 이용해 깊이 정보를 얻는 방식으로 두 개의 카메라를 이용해 깊이 정보를 얻는다. Hirshmuller *et al.*은 스테레오 매칭 알고리즘을 비교한 실험을 하였다[1]. 비교 실험 결과 센서스 변환은 원도가 간단하면서도 낮은 스테레오 매칭 오류와 방사 왜곡(radiometric distortion)에 강인한 결과를 나타내었다. 센서스 변환은 픽셀 밝기 값이 아닌 픽셀 간의 관계를 이용해 비트열로 만들어 밝기 변화에 강한 결과를 나타낸다[2].

Chang *et al.*은 계산량을 줄이면서 성능을 높인 미니 센서스 변환(mini census transform)을 제시하였다[3]. 이는 하드웨어 구현에 적합한 방식으로 원도 전체가 아닌 6개의 픽셀을 중심 픽셀과 비교하였다. 결과는 기존 센서스 변환 방법보다 낮은 에러율을 나타낸다. 최근 진행된 연구에서는 가우시안 잡음에 대해 강건한 변환인 일반화된 센서스 변환을 제시하였다[4].

본 연구는 미니 센서스 변환과 일반화된 센서스 변환을 사용한 하이브리드 센서스 변환을 제안하였다. 하이브리드 센서스 변환은 계산이 간단하면서도 잡음에 영향을 덜 받게 하였다. 미니 센서스, 일반화된 센서스, 하이브리드 센서스를 이용한 스테레오 매칭을 비교해 본 결과 이

*Corresponding Author : Seong-Hwan Jeong(Korea University of Technology and Education.)

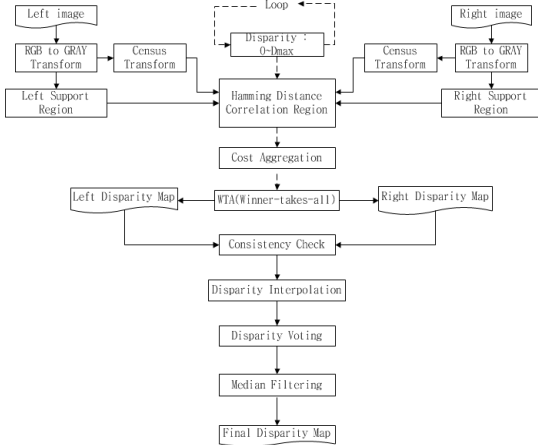
Tel: +82-041-560-1292 email: sbright@koreatech.ac.kr

Received May 20, 2013

Revised July 8, 2013

Accepted August 7, 2013

미지에 잡음이 발생하였을 시, 제안한 방법은 일반화된 센서스와 비슷한 에러율을 나타낸다. 그리고 이미지 크기에 따른 명령어 횟수는 두 방법의 중간값을 가졌다.



[Fig. 1] Block diagram of stereo matching process

논문 구성은 기존에 연구된 스테레오 매칭 기술과 제안한 방법에 대해서 설명하고, 하이브리드 센서스 변환 원도에 따른 결과를 비교하였다. 그리고 미니 센서스 변환, 일반화된 센서스 변환, 하이브리드 센서스 변환을 사용한 스테레오 매칭 결과를 통해 제안한 하이브리드 센서스 변환의 계산량과 잡음에 대한 특성을 검증하였다.

2. 스테레오 매칭 과정

스테레오 매칭은 크게 전처리, 매칭, 후처리 과정으로 나뉜다. 전처리 과정은 이미지의 잡음을 줄이거나, 매칭 알고리즘을 위한 이미지 변환을 한다. 매칭 과정은 실제로 좌, 우 이미지를 비교해가며 대응점을 찾고, 후처리 과정은 매칭 결과의 에러를 줄여준다. 실험의 전체적인 구성은 Fig. 1과 같다. 이 과정은 미니 센서스 변환[3]과 cross-based local stereo[5] 방식을 동시에 이용하여 FPGA로 구현한 연구를 참고하였다[6][7].

2.1 전처리 과정

본 연구에서 사용한 전처리 과정은 RGB to GRAY 변환, 센서스 변환, 그리고 보조 영역 계산이다. RGB to GRAY 변환은 RGB 이미지를 밝기 이미지로 변환해 밝기 성분에 대해서만 이미지 처리를 한다. GRAY값은 식(1)을 이용하여 계산하였다.

$$I_{gray} = I_{red} * 0.299 + I_{green} * 0.587 + I_{blue} * 0.114 \quad \text{식(1)}$$

센서스 변환 과정에서는 미니 센서스[3], 일반화된 센서스[4], 그리고 제안한 하이브리드 센서스 변환을 각각 사용하였다. 센서스 변환 결과는 비트열로 $C(p)$ 이다. p 는 픽셀의 좌표를 의미한다. 자세한 방법은 3절에서 다루었다.

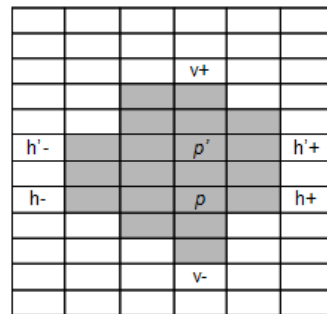
보조 영역은 매칭 과정의 상관 영역 선정과 후처리 과정의 변위 투표에서 사용되는 영역이다. 본 연구는 variable-cross 방식[6]을 이용한다. 이 방법은 cross-based local stereo[5]에서 RGB 값이 아닌 GRAY 값을 이용한다. 보조 영역은 픽셀의 상하좌우 방향별로 비슷한 밝기 값을 가진 영역의 길이를 통해 결정한다. 식(2)은 방향별 최대 픽셀 거리 r^* 을 구하는 식으로, 중심 픽셀 p 와 1부터 최대 L_{max} 까지 떨어진 픽셀 p_i 을 비교하여 가장 큰 값을 선택한다.

$$r^* = \max_{r \in [1, L_{max}]} \left(r \prod_{i \in [1, r]} \delta(p, p_i) \right) \quad \text{식(2)}$$

Π 기호는 곱을 나타내고, $\delta(p, p_i)$ 는 두 픽셀의 밝기를 비교하는 식으로 비교 방법은 식(3)에 나타내었다. τ 는 두 픽셀의 밝기 유사도를 나타낸다.

$$\delta(p_1, p_2) = \begin{cases} 1, & \text{if } |I(p_1) - I(p_2)| \leq \tau \\ 0, & \text{otherwise} \end{cases} \quad \text{식(3)}$$

식(2)을 이용해 구한 픽셀 길이는 Fig. 2와 같이 픽셀 p 의 상하 길이 내에서 다른 픽셀(p')의 좌우 거리를 이용해 보조 영역을 이룬다. v-, v+, h-, h+는 각각 픽셀 p 의 상하좌우 보조영역 길이를 나타낸다.



[Fig. 2] Support region of p

2.2 매칭 과정

매칭 과정은 해밍 거리 계산과 상관 영역 선정, 비용 종합, 최소 비용 선정이 있다. 이 과정은 좌, 우 이미지를 픽셀별로 비교하면서 변위를 찾아내는 단계이다. 각 단계는 0부터 최대 변위($d_{\max} - 1$) 범위까지 좌표를 옮기면서 수행한다.

센서스 변환 결과는 비트열이기 때문에 해밍 거리로 센서스 변환한 좌, 우 이미지의 픽셀들 정보를 비교한다. 해밍 거리는 각 변위별로 매칭 비용을 나타낸다[2]. 다음 식(4)은 해밍 거리를 계산한다.

$$Cost(p,d) = hamming(C_L(p), C_R(p,d)) \quad \text{식(4)}$$

p 는 픽셀의 위치, d 는 변위이다. C_L 과 C_R 은 각각 왼쪽, 오른쪽 이미지의 센서스 변환 값이다. 매칭 비용은 C_R 의 p 픽셀에서 변위 d 만큼 왼쪽으로 이동하면서 C_L 과 비교하여 나타낸다.

상관 영역 선정은 좌, 우 이미지 각각 계산된 보조 영역 중 겹치는 영역을 구하는 단계이다. 이 영역을 통해 물체 경계에 대한 영향을 줄인다.

비용 종합 단계는 변위별로 계산한 매칭 비용을 유일하게 만드는 과정이다. 상관 영역을 통해 인접한 픽셀들의 영역 내에 있는 모든 비용(해밍 거리)을 더한 후, 픽셀 개수로 나눈다. 본 연구에서는 계산 결과를 재사용하여 계산량을 줄이고 결과에 큰 영향을 미치지 않는, 5개의 행을 고정된 비용 종합을 사용한다[6][7]. 식(5)은 비용 종합을 계산하는 식이다. $Count(p,d)$ 는 픽셀 p , 변위 d 에서 5개 행을 고정된 상관 영역의 픽셀 수를 나타낸다.

$$Agg(p,d) = \frac{\sum_{i=1}^{Count(p,d)} Cost(p_i,d)}{Count(p,d)} \quad \text{식(5)}$$

최소 비용 선정은 모든 변위 범위에서 비용 종합한 결과들 중 가장 낮은 비용을 찾아 해당 변위를 선택한다. 그 후 선정된 각 픽셀별 변위들을 통해 초기 변위맵을 만든다. 식(6)은 초기 변위맵을 생성하는 식이다. $d_{\max} - 1$ 은 설정한 최대 변위이고, \arg_d 는 뒤에 오는 식을 만족하는 변수 d 를 의미한다.

$$dispInit(p) = \arg_d \min_{d \in [0, d_{\max} - 1]} Agg(p,d) \quad \text{식(6)}$$

2.3 후처리 과정

후처리 과정은 초기 변위맵을 보완하는 단계로 일관성 확인, 변위 보간, 변위 투표, 그리고 메디안 필터링이 있다.

왼쪽, 오른쪽 초기 변위맵의 동일한 픽셀은 변위만큼 떨어져 있으므로 일관성 확인 단계에서는 변위를 적용한 좌표를 이용해 잘못 매칭된 결과들을 찾아낸다. 일관성 확인으로 찾은 잘못 매칭된 픽셀들은 변위 보간 단계에서 가까이 있는 유효한 픽셀의 변위로 바꾼다. 식(7)은 일관성 확인 식이다.

$$validMap(p) = \begin{cases} 1, & \text{if } dispInit_L(p) = dispInit_R(p') \\ 0, & \text{if } dispInit_L(p) \neq dispInit_R(p') \end{cases} \quad \text{식(7)}$$

$validMap(p)$ 은 일관성 확인으로 찾은 유효한 픽셀을 1로, 잘못 매칭된 픽셀을 0으로 표시한다. $dispInit_L(p)$, $dispInit_R(p)$ 은 각각 왼쪽, 오른쪽 초기 변위맵의 픽셀 p 의 변위를 나타낸다. p' 은 p 좌표에서 변위 d 만큼 이동한 좌표 $(x-d, y)$ 이다. 이때 사용한 변위 d 는 $dispInit_L(p)$ 이다.

변위 투표 단계는 전처리 단계에서 계산된 보조 영역을 사용하여 기존 변위를 해당 영역 내에서 가장 빈도수가 높은 변위로 바꾼다. 이는 보조 영역을 계산할 때, 비슷한 이미지 값을 가진 부분은 같은 영역(물체)이므로 변위가 같다고 가정하기 때문이다. 식(8)은 변위 투표에 대한 식이다.

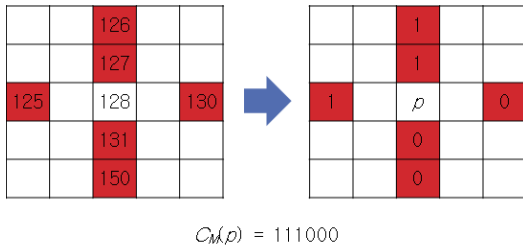
$$dispVote(p) = \arg_d \max_{d \in [0, d_{\max} - 1]} Hist(p,d) \quad \text{식(8)}$$

$Hist(p,d)$ 은 픽셀 p 의 보조 영역에서 변위 d 에 대한 히스토그램을 나타낸다. 마지막으로 변위 투표한 변위맵을 3×3 윈도우의 메디안 필터를 적용해 부분적으로 생김 에러를 감소시킨다.

3. 사용된 센서스 변환들

3.1 미니 센서스 변환

기존 밝기 기반 스테레오 매칭 방법들은 물체의 경계면에서 다른 밝기의 픽셀들(다른 물체의 영역)에 대해 영향을 받는다. 따라서 밝기에 대한 영향을 줄이기 위해 밝기 값을 변환 윈도우의 중심 픽셀과 주변 픽셀 간의 밝기 비교 결과로 바꾼 센서스 변환이 제안되었다[2].



[Fig. 3] Mini census transform example

하지만 기존의 방식은 고정된 윈도우 내 모든 픽셀들을 비교하기 때문에 윈도우 크기가 커질수록 비교할 대상과 결과를 저장할 비트 개수가 많아지게 된다. 이에 미니 센서스 변환(mini census transform, MCT)은 계산량을 줄이기 위해 5×5 변환 윈도우에서 실험으로 매칭 결과가 좋은 6개의 고정된 픽셀만 이용해 센서스 변환을 하였다. 식(9)은 미니 센서스 변환식이다[4].

$$C_M(p) = \otimes_{p' \in \hat{M}(p)} \xi(p, p') \quad \text{식(9)}$$

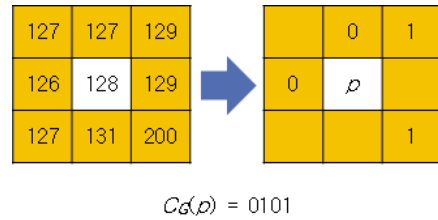
$$\xi(p, p') = \begin{cases} 1, & \text{if } I(p') < I(p) \\ 0, & \text{otherwise} \end{cases} \quad \text{식(10)}$$

식(9)에서 사용한 $C_M(p)$ 은 p 픽셀을 주변 픽셀들과 비교하여 밝기가 작은 것은 1, 큰 것은 0으로 하여 결과를 비트열로 나타낸 것이다. 기호 \otimes 은 계산 결과를 일렬로 나열하여 비트열로 나타내는 의미이다. $\hat{M}(p)$ 은 5×5 윈도우 내에서 픽셀 p 를 중심으로 한 6개의 픽셀 좌표 집합이다. 단, 중심 픽셀은 제외한다. p' 은 p 픽셀과 비교하는 대상이다. ξ 함수는 두 픽셀의 밝기를 비교하여 1 또는 0인 2진수를 반환한다[4]. I 는 해당 픽셀의 이미지 밝기 값이다.

Fig. 3는 미니 센서스 변환 예제로 중심 픽셀과 비교하는 6개의 픽셀들을 나타낸다.

3.2 일반화된 센서스 변환

기존 센서스 변환의 장점은 밝기 변화와 방사 왜곡에 강하고, 계산이 간단하다는 점이다[1]. 이는 픽셀 밝기를 매칭에 직접 이용하는 대신 대상 픽셀과 주변 픽셀과의 밝기 차이를 픽셀 정보로 사용하였기 때문이다. 하지만 이 방법은 잡음에 취약하다는 단점이 있다. 그 이유는 미니 센서스 변환이 중심 픽셀에 의존적이어서 중심 픽셀에 잡음이 생기면 해당 픽셀의 변환 전체에 영향을 주기 때문이다[4].



[Fig. 4] 3×3 generalized census transform example

일반화된 센서스 변환(generalized census transform, GCT)은 변환 윈도우의 중심 픽셀을 사용하지 않는다. 대신 변환 윈도우의 중심을 기준으로 대칭 관계에 있는 두 픽셀을 비교한다. 이는 곧 주변 픽셀들끼리 밝기를 비교하게 되면서, 중심 픽셀에 대한 의존도를 없앤다. 게다가 미니 센서스 변환과 같은 비트열 크기라도 비교하는 픽셀들은 2배로 많아지게 된다. 식(11)은 픽셀 p 에 대한 일반화된 센서스 변환 식이다[4].

$$C_G(p) = \otimes_{1 \leq i \leq n} \xi(p + c_i, p + c'_i) \quad \text{식(11)}$$

여기서 c_i 와 c'_i 은 서로 중심에 대해 대칭 관계에 있는 픽셀을 나타내기 위한 위치 변화량이다. 이는 픽셀 p 에 더해져 비교할 픽셀 좌표를 나타낸다. n 은 비교하는 비트 쌍의 개수이다. ξ 함수는 식(10)에 제시되었다.

Fig. 4은 3×3 윈도우를 이용한 일반화된 센서스 변환 예제를 나타낸다. 중심 픽셀을 제외한 8개의 픽셀을 비교하지만, 실제 비교 횟수는 4번으로 변환 결과는 4비트가 된다.

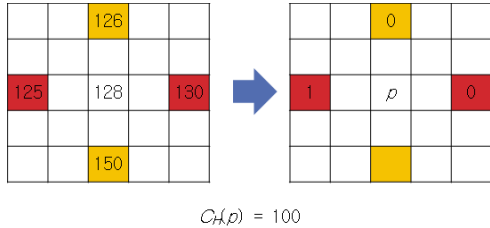
3.3 제안한 센서스 변환

본 연구에서 제안한 알고리즘은 MCT와 GCT를 사용한 하이브리드 센서스 변환(hybrid census transform, HCT)이다. MCT는 잡음이 없을 시, 계산이 간단하고 낮은 스테레오 매칭 어려움을 나타낸다. GCT는 동일한 비트열일 때, MCT에 비해 많은 픽셀을 선택해야 하지만, 잡음이 있는 경우엔 더 낮은 매칭 어려움을 나타낸다. 따라서 이 두 방법의 장점을 동시에 이용한 방법을 제안하였다. 식(12)은 HCT 식이다.

$$C_H(p) = \begin{cases} \otimes_{1 \leq i \leq m} \xi(p, p + c_i) \\ \otimes_{m < i \leq n} \xi(p + c_i, p + c'_i) \end{cases} \quad \text{식(12)}$$

총 n 개의 비트열로 나타낼 때, 1부터 m 까지는 MCT를

사용하여 나타내고, $m+1$ 부터 다음 n 까지는 GCT를 사용하였다. 사용한 c_i, c'_i 은 픽셀 위치 변화량으로 중심 픽셀 p 에 더해져 비교 픽셀의 좌표를 나타낸다.



[Fig. 5] 5×5 hybrid census transform example

Fig. 5는 HCT를 설명하는 예제 그림이다. 총 3개의 비트 중 2개는 MCT 방법을 사용하고, 1개는 GCT 방법을 사용하였다.

4. 실험 결과

4.1 실험 환경

실험에 사용한 파라미터는 $L_{max}=15, \tau=17$ 이다. L_{max} 와 τ 는 각각 보조 영역 계산 시, 각 방향별 최대 검색 길이와 밝기 유사도 경계값이다. 이는 미니 센서스 변환과 variable-cross를 이용한 스테레오 매칭 연구에서 선정된 파라미터이다[7]. 다른 센서스 변환을 하였을 때의 결과를 비교하기 위해 해당 파라미터를 그대로 사용하였다.

실험은 총 세 개로 이루어졌다. 첫 번째는 제한한 하이브리드 센서스 변환의 윈도우별 그리고 잡음 유무에 따른 결과를 비교하였다. 이 실험을 통해서 윈도우 구성에 따른 결과를 확인하였다. 두 번째 실험은 각 센서스 변환의 잡음에 대한 성능을 비교 하였다. 사용한 잡음은 가우시안 잡음으로 평균은 0이고 표준편차는 각각 2.65, 5.12, 7.68이다. 마지막 실험은 센서스 변환의 계산량으로 각 센서스 변환 시 수행된 명령어 횟수를 비교하였다.

스테레오 매칭에 사용한 이미지는 middleburry에서 제공한 Tsukuba, Venus, Teddy, Cone이다[8]. 스테레오 매칭 결과는 비폐색 영역, 모든 영역, 불연속적인 영역(경계면)에서 에러율을 측정하였다. 알고리즘 평가는 각 이미지에서 구한 에러율들을 종합한 전체 평균을 이용하였다.

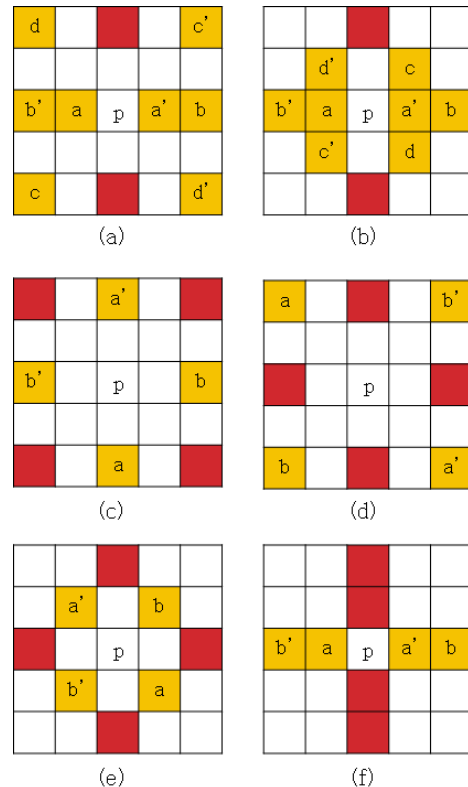
4.2 하이브리드 센서스 변환 윈도우별 비교

사용한 HCT 윈도우는 MCT 비트와 GCT 비트를 사용한 개수에 따라 나뉜다. 이는 변환 결과 6비트 중 MCT 방법

2비트와 GCT 방법 4비트, 그리고 MCT 방법 4비트와 GCT 방법 2비트 조합이다. 그리고 MCT, GCT 픽셀의 위치에 따른 윈도우를 설정하였다. 단, MCT 방법을 이용할 픽셀은 중심에서 2픽셀 떨어진 픽셀로 선정하였다. 이는 중심에 위치한 픽셀보다 1-2 픽셀 떨어진 픽셀이 신뢰도가 더 높기 때문이다[4]. Fig. 6는 조합한 HCT 윈도우들을 나타낸다.

Table 1과 Table 2은 가우시안 잡음의 유무에 따른 HCT 윈도우들을 이용한 스테레오 매칭 에러율을 나타낸다. 표의 non, all, disc 항목은 각각 비폐색, 전체, 불연속 영역을 나타낸다. Table 1을 통해 잡음이 없을 시, MCT 비트를 더 많이 사용한 경우가 낮은 에러율을 보였고, 그 중 실제 MCT 윈도우와 비슷한 형태인 Fig. 6 (f)가 후보들 중 낮은 매칭 에러율을 나타낸다.

Table 2는 평균 0, 표준편차 5.12인 가우시안 잡음이 있을 경우로, Fig. 6 (f)가 높은 에러율을 나타냈다. GCT 픽셀들이 변환 윈도우 내에서 많은 비율이거나 중심에 근접하지 않고 균등히 퍼져있는 경우, 낮은 에러율을 나타냈다.



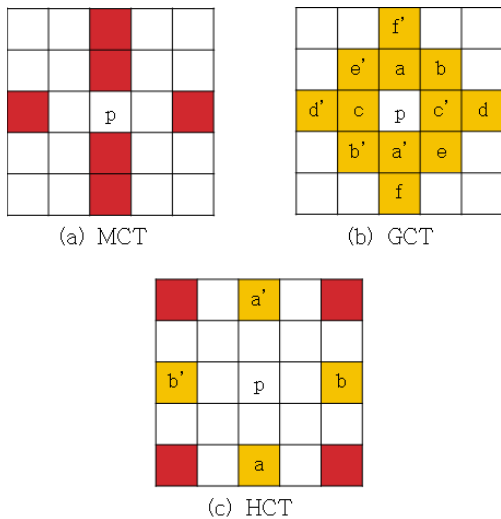
[Fig. 6] Hybrid census transform windows

[Table 1] Error rates of stereo matching result about HCT

HCT	Tsukuba			Venus			Teddy			Cones			avg
	non	all	disc	non	all	disc	non	all	disc	non	all	disc	
(a)	3.45	3.89	14.52	0.35	0.58	4.10	6.70	12.39	17.17	4.87	10.74	12.90	7.64
(b)	4.01	4.47	14.18	0.37	0.55	3.83	6.42	12.20	16.22	4.64	10.29	12.13	7.44
(c)	3.95	4.56	14.40	0.38	0.54	3.74	6.44	11.96	16.38	4.83	10.79	12.76	7.55
(d)	3.65	4.07	14.11	0.33	0.54	4.14	6.70	12.10	16.86	4.63	10.19	12.18	7.46
(e)	3.68	4.15	13.58	0.37	0.55	4.01	6.74	12.20	17.00	4.47	10.08	11.64	7.37
(f)	3.62	3.98	13.08	0.42	0.54	4.05	6.37	11.90	15.98	4.30	9.82	11.05	7.09

[Table 2] Error rates of stereo matching result about HCT with gaussian noise mean=0, $\sigma=5.12$

HCT	Tsukuba			Venus			Teddy			Cones			avg
	non	all	disc	non	all	disc	non	all	disc	non	all	disc	
(a)	16.50	16.91	27.71	18.08	18.82	19.14	20.92	27.29	29.26	8.23	15.59	19.48	19.83
(b)	17.82	18.22	26.81	20.22	20.78	18.07	21.96	27.94	29.54	8.83	15.88	19.88	20.50
(c)	13.78	14.24	22.24	19.54	20.19	23.68	20.24	26.56	29.21	8.23	15.66	18.90	19.37
(d)	18.44	18.74	27.25	20.03	20.68	17.02	23.08	29.06	30.96	9.14	16.22	20.59	20.93
(e)	20.36	20.70	28.64	22.82	23.32	18.77	25.68	31.48	32.84	10.96	17.98	22.46	23.00
(f)	23.80	24.16	33.48	24.92	25.34	23.64	27.76	33.80	35.29	12.59	19.40	25.32	25.79



[Fig. 7] Census transform windows

4.3 각 센서스 변환 비교

각 센서스 변환 윈도우 형태는 Fig. 7과 같다. HCT 윈도우는 Table 2를 통해 낮은 에러율 결과를 나타낸 Fig. 6 (c)로 선정하였다.

Table 3은 MCT, GCT, 그리고 선정된 HCT 윈도우를 이용하여 각각 스테레오 매칭한 결과의 에러율을 나타낸다.

Table 3에서 사용한 잡음은 가우시안 잡음으로 평균 0, 표준편차는 각각 2.65, 5.12, 7.68이다. 기호 σ 는 사용한 표준편차를 나타낸다. 잡음이 없을시, MCT, GCT, HCT 세 방법의 결과는 각각 7.13, 7.34, 7.55로 비슷한 결과를 가졌다. 가우시안 잡음이 추가될 경우, 전체적으로 매칭 결과는 모두 평균 에러가 커졌다. 그중 MCT 방법은 다른 두 방법에 비해 큰 에러율 결과를 나타내었다. 특히 잡음이 없을 경우, 제일 낮은 에러율을 나타내지만 잡음이 추가될수록 에러율이 크게 증가하였다. 이를 통해 MCT의 잡음에 대한 특성을 확인할 수 있다.

GCT와 HCT 방법은 잡음 유무에 상관없이 비슷한 에러율 결과를 나타낸다. 두 방법 모두 잡음이 추가될 경우, MCT보다 낮은 에러율을 가진다. 잡음의 정도가 작은 경우($\sigma=0$, 2.65) GCT 방법이 HCT에 비해 낮은 에러율을 갖지만, 잡음의 정도가 커질수록($\sigma=5.12$, 7.68) HCT 방법이 GCT 방법보다 낮은 결과를 가진다.

가우시안 잡음을 고려하지 않은 각 센서스 변환별 변위맵은 Fig. 8에 나타내었다. 잡음이 없을 시, 세 방법의 변위맵 결과는 Table 3에서 보여주는 바와 같이 큰 차이를 보이지 않았다. Fig. 9는 가우시안 잡음 mean=0, $\sigma=5.12$ 일 때의 변위맵 결과이다. MCT, GCT, HCT 세 방법 모두 매칭 에러가 증가하여 변위맵이 깨지지만, 이중 Fig. 9 (a)는 나머지 결과보다 잘못 매칭된 결과가 더 많고 물체의 경계가 모호해졌다.

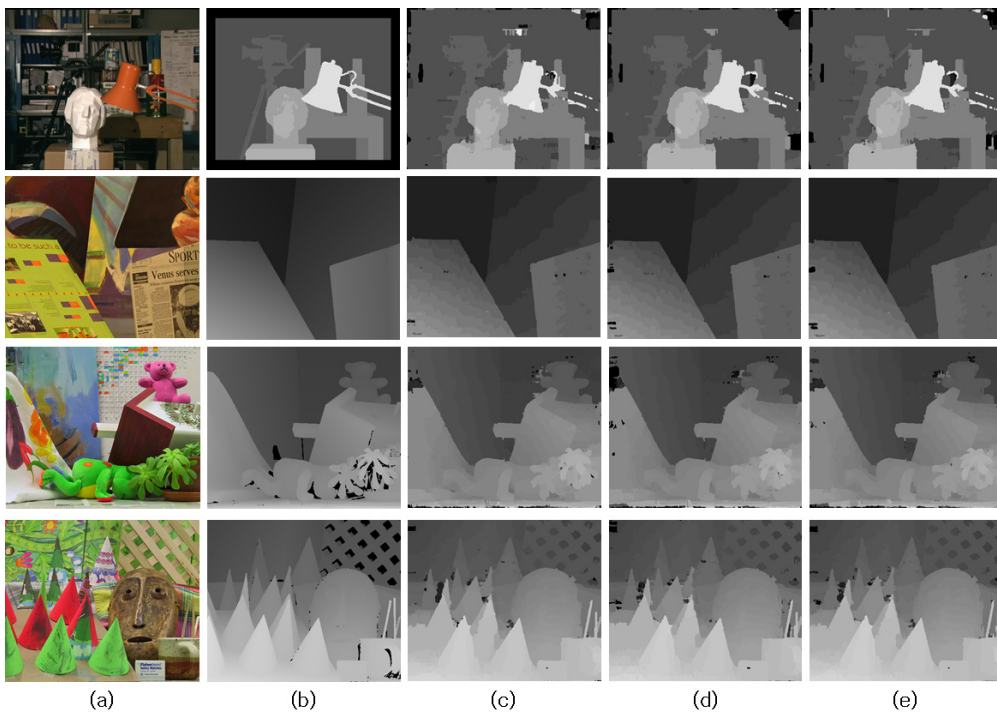
MCT, GCT, HCT 각 계산량은 이미지 크기에 따른 명령어 횟수(move + subtraction + shift)이다. Fig. 7 (a)에서 MCT 변환은 p점과 6개의 점을 하나씩 비교한다. 이때 계산량은 p점을 한번 읽고, 나머지 6개의 점을 차례로 읽으면서 비교하기 때문에 변수 이동 7번, 비교 6번, 계산 결과 이동 6번이다. Fig. 7 (b) GCT 변환은 서로 다른 위치(e.g. a와 a')의 픽셀들을 사용하기 때문에 계산량은 변

수 이동 12번, 비교 6번, 계산 결과 이동 6번이다. Fig. 7 (c) HCT의 계산량은 MCT에 해당하는 변수 이동 5번과 GCT에 해당하는 변수 이동 4번, 비교 6번, 계산 결과 이동 6번이다.

총 계산량은 각각 19(MCT), 24(GCT), 21(HCT)이다. Fig. 10은 이미지 크기별로 각 방법의 계산량을 곱해 그래프로 나타내었다.

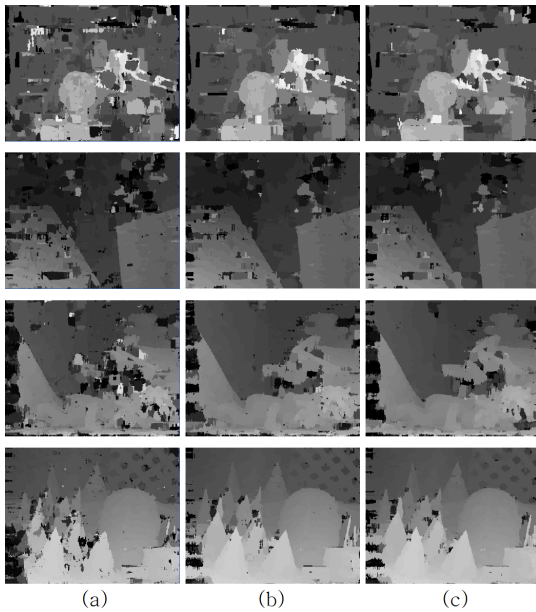
[Table 3] Error rates of stereo matching result

σ	CT	Tsukuba			Venus			Teddy			Cones			avg
		non	all	disc	non	all	disc	non	all	disc	non	all	disc	
0	MCT	4.03	4.40	13.57	0.42	0.54	4.10	6.40	11.60	15.89	4.21	9.60	10.79	7.13
	GCT	3.80	4.22	13.65	0.39	0.54	3.85	6.49	12.08	16.00	4.69	10.26	12.15	7.34
	HCT	3.95	4.56	14.40	0.38	0.54	3.74	6.44	11.96	16.38	4.83	10.79	12.76	7.55
2.65	MCT	12.03	12.34	20.19	8.88	9.24	10.38	13.01	18.37	23.53	5.40	11.34	13.47	13.18
	GCT	7.54	7.93	16.07	6.27	6.65	8.85	11.05	16.56	21.77	5.23	11.25	13.57	11.06
	HCT	7.70	8.16	16.88	6.62	7.03	9.72	11.35	16.93	22.50	5.41	11.78	14.11	11.52
5.12	MCT	26.52	26.76	34.72	27.26	27.65	25.55	31.19	37.08	38.86	16.23	22.94	30.32	28.76
	GCT	16.67	17.18	27.65	20.17	20.73	19.12	20.92	27.13	28.57	8.10	15.30	18.65	20.02
	HCT	13.78	14.24	22.24	19.54	20.19	23.68	20.24	26.56	29.21	8.23	15.66	18.90	19.37
7.68	MCT	39.17	39.44	42.04	43.65	44.14	46.40	56.79	61.00	64.89	40.86	46.03	54.89	48.27
	GCT	25.92	26.27	31.98	33.48	34.11	37.29	40.76	46.36	47.57	22.07	28.97	34.41	34.10
	HCT	22.49	22.97	31.29	33.58	34.25	38.05	40.93	46.35	46.76	21.44	28.30	35.26	33.47

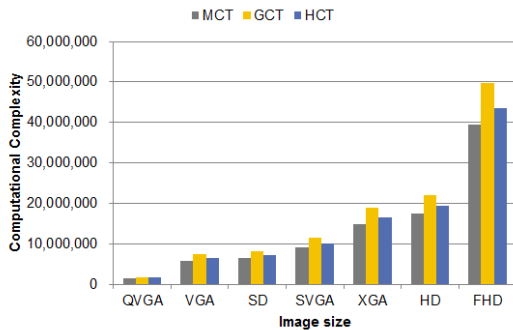


[Fig. 8] Disparity map results.

(a) input images, (b) ground truth, (c) MCT results, (d) GCT results, and (e) HCT results.



[Fig. 9] Disparity map results with mean=0, $\sigma=5.12$.
(a) MCT results, (b) GCT results, and (c) HCT results.



[Fig. 10] Graph of computational complexity

5. 결론

기존 실시간 처리에 유리한 방법으로 많이 사용된 미니 센서스 변환(MCT)은 변환 시 중심 픽셀과 이웃 픽셀 간 밝기 차이를 이용하기 때문에 중심 픽셀에 잡음이 생길 시, 매칭 에러율이 증가한다. 이에 대안으로 일반화된 센서스 변환(GCT)은 MCT와는 달리 주변 대칭 픽셀간의 비교 방식으로 변환을 시도하였다. 이를 통해 잡음이 있을 때 MCT에 비해 낮은 매칭 에러율을 얻었지만, 동일한 비트열을 사용할 시 더 많은 비트를 사용하는 단점이 있다. 본 논문은 미니 센서스 변환의 계산량과 일반화된 센

서스 변환의 잡음 특성을 동시에 이용한 하이브리드 센서스(HCT)를 제안하였다.

제한한 방법의 특성을 알아보기 위해 MCT, GCT, HCT 각 방법에 middlebury[8]에서 제공한 Tsukuba, Venus, Teddy, Cone 이미지를 이용하여 스테레오 매칭 실험을 하였다. 잡음이 없을 시, MCT와 GCT, HCT의 차이는 미미하지만, 잡음이 있는 상황에선 $MCT > GCT \approx HCT$ 로 평균 에러가 증가 하였다. 각 방법의 계산량은 이미지 한 픽셀 당 센서스 변환 시 필요한 명령어 횟수 (move + subtraction + shift)로 측정하였다. MCT, GCT, HCT 변환 윈도우에 필요한 계산량은 각각 19, 24, 21로 이미지 크기가 커질수록 계산량 차이는 증가한다.

실험 결과는 제시한 HCT가 잡음을 고려한 경우에 GCT와 에러율이 비슷하면서 MCT보다 에러에 강인한 특성을 보인다. 또한 계산량을 비교하였을 경우, HCT는 MCT보다 계산량이 많지만 GCT보다 적은 계산량을 가진다.

본 논문은 MCT의 잡음에 대한 특성을 보완하고, GCT의 계산량을 줄이는 방법에 대해서 알아보았다. MCT 방법과 GCT 방법을 동시에 사용한 HCT는 GCT 방법의 특징을 이용하여 MCT의 잡음 특성을 보완하고, MCT 방법을 통해 GCT의 계산량을 줄였다. 이와 같은 특징은 잡음을 고려한 실시간 스테레오 매칭에 유용할 것이다. 추후 향상된 잡음 제거와 실시간 처리를 위해 하드웨어 구현에 필요한 알고리즘 최적화 및 병렬화 연구가 필요할 것이다.

References

- [1] H. Hirschmuller and D. Scharstein, "Evaluation of Stereo Matching Costs on Images with Radiometric Differences," IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(9), pp.1582-1599, 2009. DOI: <http://dx.doi.org/10.1109/TPAMI.2008.221>
- [2] R. Zabih and J. Woodfill, "Non-parametric Local Transforms for Computing Visual Correspondence," European Conference on Computer Vision, pp.151-158, 1994. DOI: <http://dx.doi.org/10.1007/BFb0028345>
- [3] N. Chang, T. Tsai, B. Hsu, Y. Chen, and T. Chang, "Algorithm and Architecture of Disparity Estimation With Mini-Census Adaptive Support Weight," IEEE Transactions on Circuits and Systems for Video Technology, 20(6), pp.792-805, 2010. DOI: <http://dx.doi.org/10.1109/TCSVT.2010.2045814>

[4] W. Fire and J. Archibald, "Improved Census Transforms for Resource-Optimized Stereo Vision," IEEE Transactions on Circuits and Systems for Video Technology, PP(99), pp.1, 2012.
DOI: <http://dx.doi.org/10.1109/TCSVT.2012.2203197>

[5] K. Zhang, J. Lu, and G. Lafuit, "Cross-Based Local Stereo Matching Using Orthogonal Integral Images," IEEE Transactions on Circuits and Systems for Video Technology, 19(7), pp.1073-1079, 2009.
DOI: <http://dx.doi.org/10.1109/TCSVT.2009.2020478>

[6] L. Zhang, K. Chang, T. Chang, G. Lafuit, G. Kuzmanov, and D. Verkest, "Real-time high-definition stereo matching on FPGA," Field Programmable Gate Arrays, pp.55-64, 2011.

[7] Lu Zhang, "Design and Implementation of Real-Time High-Definition Stereo Matching SoC on FPGA," M.Sc. Thesis, Delft University of Technology, 2010

[8] D. Scharstein and R. Szeliski,
<http://vision.middlebury.edu/stereo/>

강 성 진(Seong-Jin Kang)

[정회원]



- 1998년 8월 : 연세대학교 전자공학과 공학박사
- 1998년 12월 ~ 2002년 1월 : ETRI 무선방송기술연구소 선임연구원
- 2000년 2월 ~ 2002년 8월 : (주) 이노텔리텍 기술이사
- 2002년 9월 ~ 2007년 2월 : KETI 통신네트워크연구센터 책임연구원
- 2007년 3월 ~ 현재 : 한국기술교육대학교 전기전자통신공학부 부교수

<관심분야>

WPAN, WLAN, 임베디드 시스템

정 성 환(Seong-Hwan Jeong)

[정회원]



- 2011년 2월 : 한국기술교육대학교 전기전자통신공학부 (학사)
- 2011년 3월 ~ 2013년 2월 : 한국기술교육대학교 전기전자통신공학과 (석사)
- 2013년 3월 ~ 현재 : 엑시콘 선행기술팀 연구원

<관심분야>

스테레오 매칭, 3D, 임베디드 시스템