

이동성(shiftability)을 이용한 윈도우 웨이블릿 기반 스테레오 정합

신재민, 이호근, 이호영, 하영호
경북대학교 전자전기공학부

Windowed Wavelet-Based Stereo Matching Using Shiftability

Jae-Min Shin, Ho-Keun Lee, Ho-Young Lee, and Yeong-Ho Ha
School of Electrical Engineering and Computer Science, Kyungpook National University
e-mail: yha@ee.knu.ac.kr

요약

본 논문에서는 스테레오 정합을 위한 특징으로 웨이블릿의 이동성(shiftability)을 이용한 윈도우 웨이블릿 기반 스테레오 정합방법을 제안하였다. 전 영상에 대한 웨이블릿 분해는 웨이블릿의 이동성 유지가 이루어지지 않아서 정합 정확도가 떨어진다. 그래서 웨이블릿의 이동성을 신뢰성 있는 정합정보로 만들기 위해 윈도우로 전체 파형의 일부를 표본화하고 웨이블릿 분해를 수행하여 기준신호와 이동된 신호의 부대역 정보 사이의 상관도(cross-correlation)를 정합정보로 이용하였다. 대역별 상관도는 얻어진 4개의 부대역의 대역별 가중치가 고려되어 계산된다. 제안한 방법은 주파수 대역별 계층적인 정합과 양방향 정합과정을 통해 영상의 경계부분, 동일한 형태의 반복, 잡음(white noise)등이 포함된 영상에서의 오정합을 줄일 수 있었으며 특징정보가 부족한 부분에서의 정합도 개선할 수 있었다.

I. 서론

인간 시각이 감지하는 영상정보는 깊이정보가 추가된 3차원 정보이다. 이와 같은 영상정보를 얻기 위해서 2차원 스테레오 영상으로부터 깊이정보를 추출하는 방법이 연구되어 왔다[1].

정확한 스테레오 정합을 위해서는 영상에서 신뢰성 있는 특징정보의 추출이 중요하다. 일반적인 영상의 특징정보는 필터링(filtering)된 영상의 영점 교차점, 연결성을 가진 경계선, 분할된 경계선, 모서리 성분 등을 사용한다. 이러한 특징정보들은 영상에서 모든 부분에 대해서 추출이 불가능한 문제점을 가지고 있다. 그러므로 전 영상의 스테레오 정합을 위해서는 내삽

(interpolation) 과정을 필요로 한다. 반면에 영상의 밝기 정보를 사용하는 영역기반 정합방법(area-based matching method)은 특징기반 정합방법(feature-based matching method)과 달리 전 영상에 대한 깊이정보를 얻을 수 있으나 잡음에 매우 민감하여 정합이 부정확하고 변이정보가 큰 경우에 오정합의 문제를 일으킨다. 그래서 두 가지 정합방법의 장점들을 결합하여 정확한 깊이정보를 추출하는 방법이 연구되어 왔다.

이러한 방법들 중 웨이블릿 변환(wavelet transform)을 이용한 방법은 영역기반의 정합을 하면서 얻어진 경계선에 대한 정보를 활용함으로써 기존의 방법보다 나은 정확한 정합을 할 수 있게 한다[2-3]. 또한 계층적인 구조로의 접근이 쉬워서 잡음에 대한 면역성이 클 뿐만 아니라 각 계층구조에서의 정보가 다음 계층으로 유지되는 특성을 가지고 있어 부정확한 정합을 줄여준다. 그러나, 이러한 전 영상의 웨이블릿 변환을 이용한 스테레오 정합은 신호의 웨이블릿 분해 결과가 이동성을 유지하지 못하기 때문에 오정합의 문제를 발생시킨다[5].

그래서 본 논문에서는 정확한 변위맵을 구하기 위해서 웨이블릿의 이동성(shift-variant wavelet)에 의한 오정합의 문제를 해결하였으며 기존 웨이블릿 변환의 간축(decimation)과정과 상승적분(convolution)에 의해 생기는 부정확성을 제거하였다.

II. 웨이블릿의 이동성(shiftability)

2.1 웨이블릿(wavelets)

신호와 영상처리에 쓰이는 변환은 완전 선형 변환(complete linear transform)의 성질을 가져야한다. 이들 중 푸리에 변환(Fourier transform)이나 웨이블릿 변환은 완전 선형 변환의 특성을 가지고 있어 어떤 신호를

특수한 기저함수(basis function)의 가중합으로 분해한다. 완전 선형 변환은 다음식으로 나타낸다.

$$x(m) = \sum_i y_i f_i(m) \quad (1)$$

여기서 $f_i(m)$ 은 기저함수이고, y_i 는 변환계수이다. 변환계수는 다음식으로 구할 수 있다.

$$y_i = \int h_i(m)x(m)dm \quad (2)$$

여기서 $h_i(m)$ 는 투영함수(projection function)이다.

푸리에 변환은 주어진 신호나 함수 $x(m)$ 을 분석하고 그것의 특성을 파악하기 위해서 정현파(sine wave)나 여현파(cosine wave)와 같은 기저함수를 사용한다. 이 방법은 정상 신호(stationary signal)인 경우에서만 효과적인 신호분석이 가능하다. 그러나 자연계에 존재하는 신호의 대부분은 시간의 변화에 따라 확률적인 특성이 변하는 약정상 신호(non-stationary signal)이므로, 이 신호를 다루는 방법 중 하나가 특수한 신호의 집합을 정하고 그 집합에 적합한 기저신호(basis signal)를 이용하여 주어진 신호를 재구성하는 웨이블릿 변환을 사용하는 것이다. DWT(discrete wavelet transform)를 사용한 웨이블릿 분해(wavelet decomposition)는 다음식으로 나타낼 수 있다.

$$x_K = \sum_j c_{k_0,j} \phi_{0,j} + \sum_{k=k_0}^{K-1} \sum_j d_{k,j} \psi_{k,j} \quad (3)$$

여기서 $\phi_{0,j}$ 와 $\psi_{k,j}$ 는 웨이블릿 함수이고, c_{k_0} 와 $d_{k,j}$ 는 해상도수준 k 에서 x 의 웨이블릿 계수이다.

2.2 이동성(shiftability)

웨이블릿 변환에서 잉여(redundancy) 정보를 제거하는 기저함수의 직교성(orthogonality)과 분해 결과의 이동성 유지는 중요한 특성이다. 그런데 두 가지 중 이동성 유지는 스테레오 정합에서 좌·우 영상의 웨이블릿 분해 결과가 신호의 위치 변화만으로 나타나 정확한 정합을 이루게 하는 요소인데 웨이블릿 변환은 완전한 이동성 유지를 할 수가 없다[3-4].

이동성의 영향을 나타내는 그림 1은 기존의 웨이블릿을 이용한 정합방법에서 오정합을 일으키는 주요 요인이다. 그래서 이동성을 유지할 수 있는 방법이 연구되었는데 steerable pyramid는 선형 다중해상도(linear multi-scale)와 다차원 영상 분해(multi-orientation image decomposition)의 특성을 가지면서도, 주파수 영역에서 극 분리형 분해(polar-separable decomposition)를 수행하므로 해상도와 차원의 독립성이 유지된다. 이

런 특성을 기반으로 쉽게 다차원 확장이 가능한 분리형 정규 직교 웨이블릿(separable orthonormal wavelets)은 이동성이나 회전에 불변하기 때문에 신호처리에 유용하다. 그러나 필터의 설계가 용이하지 못하여 완전한 이동성의 유지는 할 수 없다[4]. 그러나 제한한 알고리즘에서는 이동성을 유지할 수 없는 특성을 이용해서 기준 입력신호의 웨이블릿 분해 결과와 이동된 입력신호의 분해 결과와의 대역별 상관도가 낮아진다는 것을 정합 정보로 이용하였다.

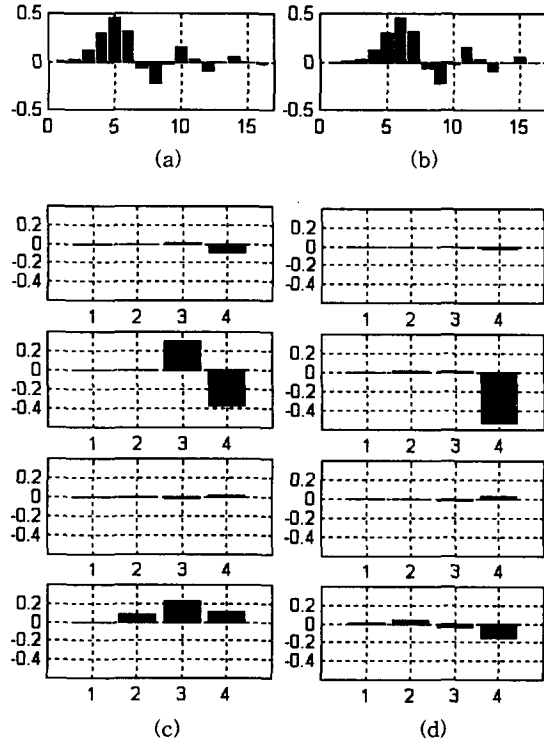


그림 1. 신호의 웨이블릿 분해시 이동성의 영향 (a) 입력신호 (b) 우측으로 1화소 이동된 입력신호 (c) 입력신호의 웨이블릿 분해 (d) 이동된 신호의 웨이블릿 분해

III. 윈도우 웨이블릿 기반 스테레오 정합

영상 전체에서 경계선이나 모서리 정보는 많은 특징 정보를 제공하지 못한다. 그래서 웨이블릿의 이동성이 유지되지 못하는 점에 착안해 좌·우 영상의 신호를 표본화 한 각각의 윈도우를 웨이블릿 변환하고, 얻어진 부대역들의 대역별 유사도로 정합을 수행한다.

전체의 과정은 정확한 탐색구간에 의한 정합의 신뢰도를 높이기 위해 계층적인 구조로 수행된다. 본 논문에서는 영상의 간측과정에 의한 3단계 계층구조를 사용하였다. 그리고 계층별 최종적인 변위맵 구성을 위해

계층별 양방향 정합(bi-directional matching)으로 오정합을 일으키는 반복 형태, 잡음 등의 영향을 한번 더 제거하여 정합의 정확도를 높일 수 있었다. 그림 2는 제안한 알고리즘의 흐름도이다.

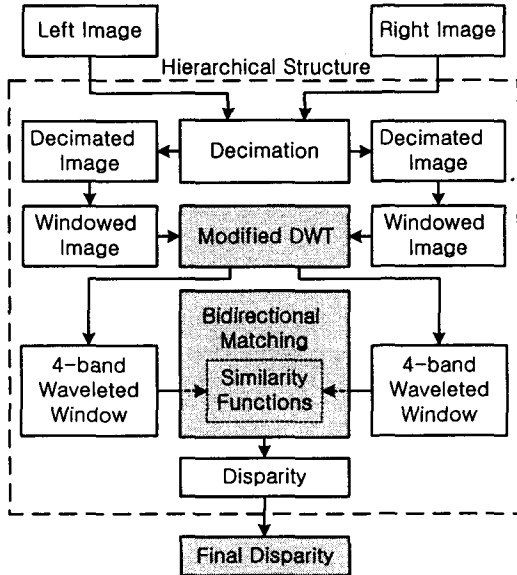


그림 2. 윈도우 웨이블릿 기반 정합과정

3.1 윈도우 웨이블릿 분해

이동성을 특징정보로 사용하기 위해서 극상선 제약(epipolar constraint)을 가정하고 좌·우 영상에서 입력 신호의 압축·팽창의 영향은 신호의 이동에 비해 미약하다고 가정하였다.

제안한 방법에서 윈도우의 웨이블릿 분해는 정확한 주파수 대역별 유사도(similarity)를 얻기 위해서 윈도우만의 정보로 상승적분을 수행하지 않았다. 그래서 동일한 크기의 윈도우 외부 영역을 추가로 이용하여 경계 부분의 오차를 줄인다. 그림 3은 입력영상에서 좌우의 확장된 영역이 상승적분을 위한 외곽정보로 사용되는 것을 보여준다.

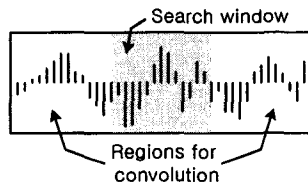


그림 3. 상승적분을 위한 외곽정보

그림 4는 간축과정을 보여주는데, 여기서 입력신호의 반이 제거되므로 얻어진 부대역 정보는 입력신호 중 어

떤 두 개의 다른 주파수 성분이 혼합되어 대역별 정보의 손실이 발생한다. 그래서 제안한 방법에서는 입력신호의 부대역 정보 손실을 줄여서 정확한 정합을 하도록 간축과정을 제거하였다.

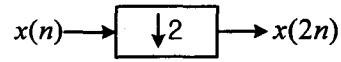


그림 4. 시간영역에서의 간축과정

그래서 간축과정을 제거하여 부대역 정보의 손실을 줄이기 위해 수정된 웨이블릿 분해의 최종결과는 그림 5와 같이 대역별 정합정보를 얻는다.

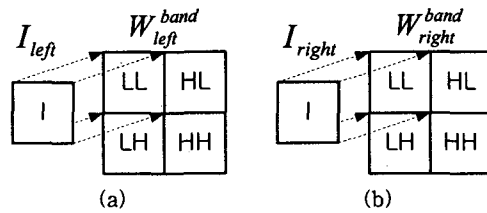


그림 5. 윈도우의 수정된 웨이블릿 분해 (a) 좌측 윈도우 영상 (b) 우측 윈도우 영상

3.2 대역별 유사도

제안한 방법의 유사도 추정에는 대역별 상관도가 사용되었다. SSD(sum of squared difference)나 SAD(sum of absolute difference)는 백색잡음(white noise)이나 조명변화와 부분폐색(partial occlusion)에 의한 잡음에 민감하기 때문에 비교하는 부대역에 나타난 신호의 유무와 크기에 대한 상황을 나타내지 못하는 문제점이 있다. 그러나 상관도 함수는 부대역 안에서 나타난 신호의 크기와 유무에 더 민감한 특성이 있어서 정합 정확도에 결정에 유리하다. 따라서 정확한 유사도 추정을 위해서 NCC(normalized cross-correlation coefficient)를 사용하였다.

정합과정에서 분해된 4개 부대역의 정보를 가지고 주파수 대역별 상관도가 얻어진다. 그러나 얻어진 4개의 부대역에서 정합을 위한 신호의 주파수 대역별 power spectrum은 모두 동일하지 않다. 즉, 정합을 위한 유사도는 4개의 부대역에서 나타난 각각 평균 power에 비례하는 상관도들의 가중합으로 구성된다. 다음식은 정합을 위한 유사도와 가중치를 위한 각 대역별 평균 power를 나타낸다.

$$score_{band}(W_{left}^{band}, W_{right}^{band}) = \frac{P_{band}^{avg}}{\sum_{i=LL}^{HH} P_i^{avg}} \times NCC(W_{left}^{band}, W_{right}^{band}) \quad (4)$$

$$P_{band}^{avg} = \frac{P_{band}^{left} + P_{band}^{right}}{2} \quad (5)$$

여기서 P_{band}^{avg} 는 좌·우 영상 각각의 대역별 평균 파워이고 $NCC(W_{left}^{band}, W_{right}^{band})$ 는 좌·우 영상의 대역별 웨이블릿 계수 사이의 상관도를 나타낸다.

IV. 실험 및 결과

제안한 알고리즘의 성능을 분석하기 위해 128×128 크기의 50% WGN(white Gaussian noise)가 첨가된 'RDS'(random dot stereogram)와 20%의 WGN이 첨가된 'stripe' 영상을 사용하였다.

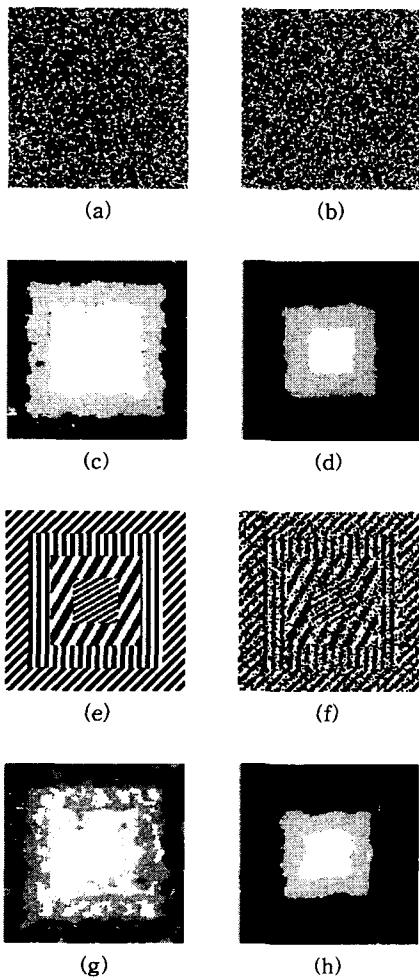


그림 6. 제안한 방법과 기존의 방법에 의한 변위맵 (a) 좌측 RDS (b) 우측 RDS (c) 기존의 방법 (d) 제안한 방법 (e) 좌측 'stripe' 영상 (f) 우측 'stripe' 영상 (g) 기존의 방법 (h) 제안한 방법

그림 6에서 보는 것과 같이 전 영상을 웨이블릿 변환해서 수행되는 기존의 웨이블릿 기반의 정합방법은 웨이블릿의 이동성 유지가 되지 않는 문제점과 첨가된 WGN에 의해 경계부분의 오정합이 많았으나 제안한 방법에서는 경계부분의 오정합이 줄어들었음을 볼 수 있다. 또 'stripe'와 같이 반복된 형태의 무늬와 WGN이 포함된 영상에서도 기존의 방법보다 비교적 정확한 정합이 가능하였다.

V. 결론

본 논문에는 웨이블릿의 이동성을 이용한 윈도우 웨이블릿 정합 방법이 제안되었다. 기존의 웨이블릿 기반의 스테레오 정합은 전 영상을 웨이블릿 분해한 주파수 대역별 정보를 BMA(block matching algorithm)로 정합을 수행했기 때문에 웨이블릿 이동성 유지가 이루어지지 않아서 오정합이 발생하였다. 그래서 제안한 방법은 웨이블릿의 이동성이 유지되지 못하는 특성을 이용하여서 정확한 정합을 위한 특징정보로 사용하였다. 그래서 첨가된 WGN에 의한 주파수 성분이 분해된 부대역 중 높은 주파수 대역에 분포하게 되어 최종 유사도를 위한 가중치에 의해 WGN의 영향을 제거하여 가장 유사한 분포를 찾게 되므로 정확한 정합을 유지할 수 있었다. 향후 보다 더 정확한 정합을 위해 신호의 이동만이 아닌 시야각에 의해 발생하는 신호의 압축과 팽창을 고려한 연구가 계속 수행되어야 할 것이다.

참고문헌

- [1] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," *Sci.*, vol. 194, pp. 283-287, Oct. 15, 1976.
- [2] Fangmin Shi, Neil Rothwell Hughes, and Geoff Roberts, "Wavelet transforms for stereo vision," *3rd Workshop on European Scientific and Industrial Collaboration (WESIC 2001) Enschede, The Netherlands*, pp 267-276, 2001.
- [3] Young-Suk Kim, Jun-Jae Lee, and Yeong-Ho Ha, "Stereo Matching Algorithm Based on Modified Wavelet Decomposition Process," *IEEE Pattern Recognition*, vol. 30, no. 6, pp. 929-952, 1997.
- [4] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: a flexible architecture for multi-scale derivative computation," *IEEE Second Int'l Conf. on Image Processing*, Washington DC, Oct. 1995.
- [5] Eero P. Simoncelli, William T. Freeman, Edward H. Adelson and David J. Heeger, "Shiftable multi-scale transforms," *IEEE Trans. on Information Theory*, vol. 38, no. 2, pp. 587-607, 1992.