

웨이브릿 변환을 이용한 계층적 스테레오 정합

(A Hierarchical Stereo Matching Algorithm Using Wavelet Representation)

金 容 奭*, 李 駿 宰*, 河 永 浩*

(Young-Suk Kim, Jun-Jae Lee and Yeong-Ho Ha)

要 約

본 논문에서는 웨이브릿 변환된 영상에서 국소 적응 가중치와 가변 창틀을 이용하여 변이(disparity)를 구하는 계층적 정합 방법이 제안되어 진다. 웨이브릿 변환에 의해 얻어진 계층적 구조는 기존의 Gaussian이나 Laplacian pyramid가 가지는 정보의 손실을 해결할 수 있다. 웨이브릿 변환된 영상은 입력 영상이 뭉롱화된 영상, 수평 경계 성분, 수직 경계 성분, 대각 경계 성분으로 분할된다. 각 특징에 대한 가중치는 웨이브릿 부대역간의 유사도로 부터 결정되어 지고, 영역 기반(area-based)과 특징 기반(feature-based) 정합 방법을 적응적으로 고려할 수 있게 해준다. 웨이브릿 변환에 의해 얻어진 특징값들은 정보의 중첩이나 손실없이 전영상에서 얻어질 수 있다. 한편, 변이값의 정확하고 안정된 추정을 위해 필요한 가변 창틀은 낮은 해상도에서 추정된 변이값과 웨이브릿 변환된 영상의 LL(low-low) 부대역으로 부터 적응적으로 구해진다. 그리고 변이값의 경계 부분에서의 뭉롱화 현상없이 오정합을 줄일 수 있는 새로운 이완화 방법을 제안한다. 다양한 영상에 대한 실험결과와 스테레오 정합에 많은 어려움을 가진 영상에 대해서도 제안한 방법이 우수한 결과를 가져옴을 보여준다.

Abstract

In this paper, a hierarchical stereo matching algorithm to obtain the disparity in wavelet transformed domain by using locally adaptive window and weights is proposed. The pyramidal structure obtained by wavelet transform is used to solve the loss of information which the conventional Gaussian or Laplacian pyramid have. The wavelet transformed images are decomposed into the blurred image, the horizontal edges, the vertical edges, and the diagonal edges. The similarity between each wavelet channel of left and right image determines the relative importance of each primitive, and make the algorithm perform the area-based and feature-based matching adaptively. The wavelet transform can extract the features that have the dense resolution as well as can avoid the duplication or loss of information. Meanwhile, the variable window that needs to obtain precise and stable estimation of correspondense is decided adaptively from the disparities estimated in coarse resolution and LL(low-low) channel of wavelet transformed stereo image. Also, a new relaxation algorithm that can reduce the false match without the blurring of the disparity edge is proposed. The experimental results for various images show that the proposed algorithm has good performance even if the images used in experiments have the unfavorable conditions.

*正會員, 慶北大學校 電子工學科
(Dept. of Elec. Eng., Kyungpook Nat'l Univ.)

接受日字 : 1993年 10月 13日

I. 서론

컴퓨터 시각 분야에서는 거리를 측정하기 위해 여러가지 방법들이 이용되어 왔으며, 이 중에서 좌 우 영상간의 변이(disparity)를 추정하여 거리 정보를 추출하는 방법이 매우 유용한 것으로 알려져 있다.^[1] 이러한 스테레오 정합 과정에서 일어날 수 있는 어려움은 특징 추출의 과정, 추출된 특징들을 동시에 고려하기 위해 결합하는 과정, 큰 변이값을 가진 영상을 다루는 것, 가장 적당한 창틀 크기와 형태의 결정, 제한 조건들이 성립하지 않는 변이값의 경계 부분의 처리등이다. 본 논문에서는 이러한 문제점들을 해결할 수 있는 스테레오 정합 방법을 제안한다.

기존의 스테레오 정합 방법들은 정합에 사용되는 특징값에 따라 몇 가지로 구분될 수 있다. 필터링된 영상의 영점 교차점^[5], 연결성을 가진 경계선^[6], 분할(segmentation)된 경계선^[7], 모서리 성분^[8] 등을 특징점으로 가지는 특징 기반(feature-based) 정합 방법을 이용하여 추출된 변이값은 특징점에서는 매우 정확한 값을 가진다. 그러나 경계선의 세선화(thinning)와 연결(linking) 과정과 같은 복잡한 처리를 거쳐야 하고, 전 영상에서 변이 성분을 얻기 위해서는 내삽(interpolation) 과정을 필요로 한다. 그리고 영상내의 특징점들의 분포 상태가 너무 조밀하거나 특징점의 수가 너무 적은 경우 정확한 변이값을 구해내기가 어렵다. 반면에, 영상의 밝기 정보를 그대로 정합에 이용하는 영역 기반 정합 방법은 특징점의 추출이나 내삽 과정없이도 전영상에서 변이값을 얻을 수 있으나^[9-11], 잡음에 매우 약하므로 잡음을 제거하기 위한 전처리 과정을 필요로 하고, 다루어야 할 변이값이 클 경우 많은 잘못된 정합이 일어난다. 그러므로 이러한 영역 기반 정합 방법의 단점을 보완하기 위해 여러장의 영상을 정합에 이용,^[12] 계층적인 방법의 도입^[13], 다양한 제한 조건을 만족시키는 신경회로망의 사용^[14] 등을 시도해 왔다. 이외에 다른 스테레오 정합 방법에는 분할된 영역(segmented region)을 단위로 하여 정합을 행하는 방법^[15], 보다 안정된 특징이 될 수 있는 두 영상간의 위상차를 이용하는 방법^[16], 영상의 형태론적 특징을 이용하는 방법^[17], 신호의 전후 관계를 고려해서 정합을 이루는 방법^[18], Gabor 필터링된 결과를 이용하는 방법^[19], 신경 회로망을 이용해서 정합 양식을 학습시키는 방법^[20] 등이 있다.

계층적 구조에서, 영상에 대한 정보들은 해상도가 낮은 계층에서 해상도가 높은 계층으로 점진적으로 해석되어 진다. 큰 변이값을 다루어야 하는 영상에서 coarse to fine strategy라고 불리는 다해상도 해석

개념이 스테레오 정합에 많이 이용되어 왔다.^{[13] [21]} 이러한 coarse to fine 방법은 우선 전체적인 특징을 고려하고, 점차적으로 해상도를 높여 정확한 변이값을 추정한다. 계층적 구조를 형성하기 위해 이용되는 filter는 잡음의 감소 효과 뿐만 아니라 계층 형성 과정에서 특징값에 대한 정보를 보존하는 특징을 가지고 있어야 하지만 기존의 Gaussian이나 Laplacian filter를 이용해서 형성된 계층 구조에서는 정보의 손실과 함께 계층간의 상관 관계(correlation)가 존재하게 된다. 더욱이 Laplacian pyramid는 공간적 방향 정보를 가지지 않으므로 방향 선택성이 없다. 그러므로 손실된 정보로부터 추출된 특징을 정합에 이용하는 낮은 해상도 계층에서의 정합 결과는 부정확하게 된다. 반면, 웨이브릿 변환에 의해 얻어진 계층적 구조에서는 정보의 손실이 없으며, 계층 형성 과정에서 동시에 특징값의 추출이 이루어진다. 또한, 최근에 연구 되고 있는 영역 기반과 특징 기반을 동시에 고려하는 방법들에서 가장 큰 어려움인 각 특징들을 결합하는 문제를 해결할 수 있다.

한편, 정확한 변이값의 추정을 위해 적당한 창틀의 크기와 형태를 결정하는 것은 매우 중요하다. Okutomi와 Kanade^[23]는 입력 신호와 추정된 변이값의 변화 정도에 따라 창틀의 크기를 결정하였고 이러한 개념을 2차원으로 확장시켰다. 제안한 스테레오 정합 방법에서는 비반복적인 방법으로 웨이브릿 변환된 영상의 LL 부대역과 추정된 변이값으로부터 창틀의 크기와 형태를 결정한다.

계층적 방법에 기초한 앞서의 방법과 함께 보다 정확한 변이값의 추정을 위해 주변 변이값에 따라 정합 확률을 바꾸어 줄 수 있는 새로운 이완화 방법을 제안한다. 웨이브릿 변환된 영상에서 국소 적응 가중치와 가변 창틀을 이용하여 얻어진 각 부대역간의 유사도로 부터 얻어진 초기의 확률값은 주변의 값들과의 차이에 따라 반복적으로 개선시킴으로써 변이값의 경계 부분에서의 몽롱화 현상없이 오정합을 감소시킨다.

제안한 정합 방법을 RDS(random dot stereogram), 합성 영상, 실내 영상, 실외 영상에 적용한 결과, 제안한 정합 방법이 특징점의 분포 상태가 너무 조밀하거나 특징점의 수가 절대적으로 부족한 경우, 혹은 심한 잡음이나 반복되는 형태를 가진 영상에 대해서도 매우 효과적임을 보여준다.

II. 스테레오 정합을 위해 사용된 웨이브릿 변환

1. 웨이브릿 변환

WT(wavelet transform)은 STFT(short time

Fourier transform)이나 Gabor 변환을 대신해서 non-stationary 신호 해석에 이용되고 있으며^[24], 이산 신호에 적용되는 DWT(discrete wavelet transform)은 다해상도 해석^[25, 27], 영상 압축^[28, 29], 정제선 검출^[30, 31] 등에 사용되어 진다. 웨이브릿 변환은 웨이브릿이라 불리는 기본함수들의 집합으로 분할하는 과정이다. 웨이브릿들은 식(1)에서 처럼 하나의 기본되는 웨이브릿(prototype wavelet) $h(x)$ 의 확대(dilation), 이동(shift), 축소(contraction)에 의해 형성되어 지고 이러한 웨이브릿들로부터 식(2)와 같은 CWT(continuous wavelet transform)이 유도되어 진다.^[32]

$$h_{a,b}(x) = \frac{1}{\sqrt{|a|}} h\left(\frac{x-b}{a}\right) \quad a, b \in R, \quad a \neq 0 \quad (1)$$

$$F(a,b) = \langle h_{a,b}, f \rangle = \int f(x) h_{a,b}(x) dx \quad (2)$$

$$\text{where } \int \frac{|H(\omega)|^2}{|\omega|} d\omega < \infty$$

이때 $h_{a,b}(x)$ 는 $h(x)$ 가 a 만큼 scaling 되고 b 만큼 이동된 함수이고, $F(a,b)$ 는 웨이브릿 $h(x)$ 로 부터 변환된 함수이며, $H(\omega)$ 는 $h(x)$ 의 Fourier 변환된 함수이다. 웨이브릿들은 constant-Q 개념에 따라 형성되므로 고주파에서는 좁은 창틀, 저주파에서는 넓은 창틀이 적용된다. 2의 해상도에서 $f(x)$ 의 이산 추정 신호(discrete approximation signal) $A_{2^j}^d f$ 와 이산 세부 신호(discrete detail signal) $D_{2^j}^d f$ 는 scaling 함수 $\varphi(x)$ 와 웨이브릿 $\psi(x)$ 로 표현되어 질 수 있다. 2차원 신호에서의 DWT는 식(3)-(6)과 같다.^[26]

$$A_{2^j}^d f = ((f(x,y) * \varphi_{2^j}(-x) \varphi_{2^j}(-y))(2^j n, 2^j m))_{(n,m) \in Z^2} \quad (3)$$

$$D_{2^j}^1 f = ((f(x,y) * \varphi_{2^j}(-x) \psi_{2^j}(-y))(2^j n, 2^j m))_{(n,m) \in Z^2} \quad (4)$$

$$D_{2^j}^2 f = ((f(x,y) * \psi_{2^j}(-x) \varphi_{2^j}(-y))(2^j n, 2^j m))_{(n,m) \in Z^2} \quad (5)$$

$$D_{2^j}^3 f = ((f(x,y) * \psi_{2^j}(-x) \psi_{2^j}(-y))(2^j n, 2^j m))_{(n,m) \in Z^2} \quad (6)$$

이때, $\varphi_{2^j} f$ 와 $\psi_{2^j} f$ 는 각각 2^j 의 해상도를 가지는 scaling 함수와 웨이브릿이다.

2. 스테레오 영상에서 shiftability를 유지하기

위한 변형된 웨이브릿 변환

스테레오 정합에서 좌 우 영상사이에 변이가 존재할 때 filtering된 결과가 단지 위치의 변화만으로 나타날 때 shiftability가 유지된다고 말할 수 있다^[33].

주파수 영역에서는, sampling된 신호의 shiftability가 유지되기 위해 aliasing이 일어나지 않아야 한다. 그러므로 Nyquist 주파수 이하의 비율로 sampling된 신호는 위치와 해상도에 따라 불안정한 특성을 가지게 되는데, 유한한 갯수의 tap을 가진 H 와 G filter를 통과한 신호가 decimation과정을 거치게 되면 aliasing이 일어나게 되고, 이로 인해 웨이브릿 변환의 불안정성이 발생한다. 이러한 aliasing을 제거하기 위해서는 H 와 G filter를 보다 이상적인 형태로 설계하여야 하지만 제한된 filter tap의 갯수로 불가능하다. 그러므로 본 논문에서는 decimation 되기전의 aliasing 효과가 적은 신호를 정합에 이용했으며, 비교적 shiftability의 유지가 쉬운 Mallat의 웨이브릿^[26]을 사용했다. 그림 1은 스테레오 영상에서 shiftability를 유지하기 위한 변형된 웨이브릿 변환의 과정을 보여준다.

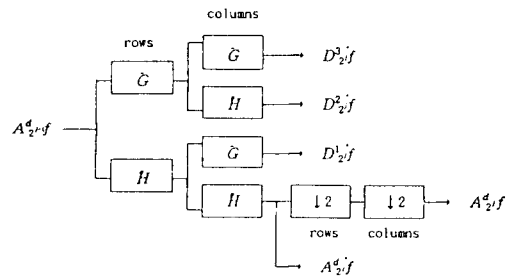


그림 1. 변형된 웨이브릿 변환 과정

Fig. 1. Modified wavelet transform preserving shiftability.

III. 제안한 계층적 스테레오 정합방법

1. 계층적 정합구조

Mallat가 제안한 웨이브릿^[26]을 사용하는 변형된 웨이브릿 변환을 이용하여 그림2와 같은 계층적 구조를 형성한다. $LL(A_{2^j}^d)$ 부대역은 몽롱화된 영상이고, $LH(D_{2^j}^1)$, $HL(D_{2^j}^2)$, $HH(D_{2^j}^3)$ 부대역은 각각 수평 경계 성분, 수직 경계 성분, 대각 경계 성분을 나타낸다. 현재의 계층에서의 초기치는 낮은 해상도를 가지는 계층에서 얻어진 변이 정보를 2배로 하여 결정되어지고, 가장 낮은 해상도를 가지는 계층의 초기치는 0으로 설정하였다. 그러므로 탐색 영역이 큰 영상에 대해서도 각 계층에서의 탐색 영역은 크게 줄어든다. 예를 들어 계층의 수가 l 인 경우, 각 계층에서의 탐색 영역을 $-s \sim s$ 으로 설정하면, 전체 영상에서는 2

$(2^j - 1)s + 1$ 의 탐색 영역을 가지는 것과 같다. 이러한 탐색 영역의 감소는 영상에서의 정합 오류나 중복 정합을 크게 줄일 수 있으며, 계산 시간의 감소 효과를 가져온다. 실제로 위의 예에서, $N \times N$ 크기의 영상에서 탐색해야할 영역은 모두 $(2(2^j - 1)s + 1)N^2$ 이 된다. 그러나, 계층적 구조를 이용할 경우 그 양은

$$C = (2s + 1)N^2 \times \left(1 + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^{2^j - 1}}\right) \quad (7)$$

$$= (2s + 1)N^2 \left(\frac{1 - \frac{1}{2^{2^j}}}{1 - \frac{1}{2^2}}\right) = \left(\frac{4}{3} \cdot \frac{2^{2^j} - 1}{2^{2^j}}\right) (2s + 1)N^2$$

과 같이 되고, j 이 4이고 s 가 1인 경우 탐색을 위한 계산량은 전체 영역을 탐색할 때의 10.37% 밖에는 되지 않는다. 본 논문에서와 같이 웨이브릿 부대역을 모두 고려하더라도 계산량은 계층적인 구조를 사용하지 않았을 때의 반 이하로 줄어든다.

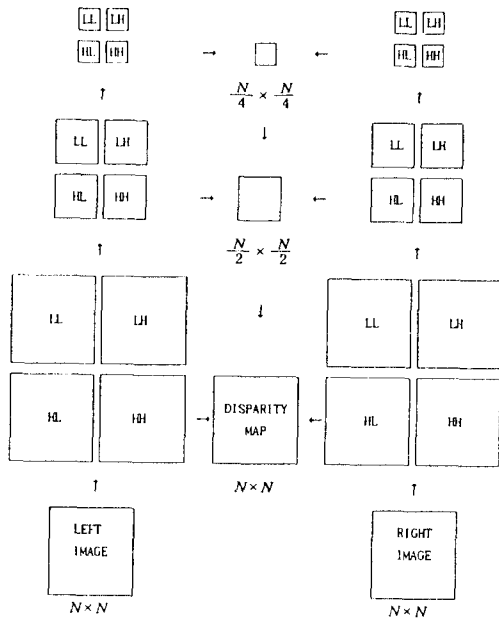


그림 2. 제안한 계층적 스테레오 정합의 구조
Fig. 2. The structure of the proposed hierarchical stereo matching.

2. 국소 적응 가중치와 가변 창틀을 이용한 웨이브릿 변환된 영상에서의 정합

1) 정합을 위한 유사 함수의 선택

정합에 사용되는 유사 함수는 식 (8)과 같은 p-Holder norm으로 정의할 수 있다.

$$E_p = \sqrt[p]{\sum_{j=1}^N \sum_{i=1}^N |x_{ij} - y_{ij}|^p} \quad p \geq 1 \quad (8)$$

p 가 1 인 경우,

$$E_1 = \sum_{j=1}^N \sum_{i=1}^N |x_{ij} - y_{ij}| \quad (9)$$

이 되고, E_1 을 최소화하게 되면 Dirac peaks를 만들어 내는 경향이 있다. 반면 p 인 경우에는

$$E_\infty = \lim_{p \rightarrow \infty} \sqrt[p]{\sum_{j=1}^N \sum_{i=1}^N |x_{ij} - y_{ij}|^p} \quad (10)$$

$$= \sup |x_{ij} - y_{ij}|$$

가 되어, E_∞ 를 최소화 하는것은 각 지점에서 두 영상 사이의 차이를 균일하게 만들어 주는 효과가 있다.

[34] 만약 SAD(sum of absolute difference)를 유사 함수로 선택하게 되면 SSD(sum of squared difference)값이 같더라도 오차가 한 화소에 모두 몰려 있는 경우(전체 창틀 내부의 화소들중 일부가 잡음에 의해 정합 오차를 크게 발생시킨 경우)가 최소값으로 선택된다. 이러한 것은 백색 잡음에 대해서는 유사 함수로서 SAD가 SSD보다 더 유리하게 작용할 수 있음을 보여준다.

제한한 정합 방법에서는, 식(11)과 같은 WSMAD(weighted sum of mean absolute difference)가 각 계층에서의 변이값을 결정하기 위해 사용되어진다.

$$E_i = \frac{1}{m \times n} \sum_{j=1}^m \sum_{k=1}^n [w_{ij}^{A^d} |A_i^d \cdot (i, j) - A_i^d \cdot (i, j + d_j)| + w_{ij}^{D^1} |D_i^1 \cdot (i, j) - D_i^1 \cdot (i, j + d_j)| + w_{ij}^{D^2} |D_i^2 \cdot (i, j) - D_i^2 \cdot (i, j + d_j)| + w_{ij}^{D^3} |D_i^3 \cdot (i, j) - D_i^3 \cdot (i, j + d_j)|] \quad (11)$$

이때 $w_{ij}^{A^d}, w_{ij}^{D^1}, w_{ij}^{D^2}, w_{ij}^{D^3}$ 는 (i, j) 위치에서의 각 웨이브릿 부대역에 대한 가중치, d_j 는 (i, j) 위치에서 변이값이고, $n \times m$ 은 창틀의 크기이다. 이때 각 화소에 대한 가중치는 WSMAD의 에너지가 최소로 되는 값으로 결정되어진다.

2) 국소 적응 가중치

앞에서 살펴본 바와 같이 각 웨이브릿 부대역들은 2차원 filter bank를 형성하게 되고, 영상은 서로 다른 성질을 가지는 여러 값들로 분해된다. WSMAD에서 각 부대역에 적용되는 가중치는 식(12)와 같이 가정합이 일어난 위치에서의 NCC(normalized cross-correlation)를 이용해서 초기화 되어진다.

$$w_{ij}^a = \frac{M[q(i,j)q(i,j+\hat{d}_{ij}^a)] - M[q(i,j)]M[q(i,j+\hat{d}_{ij}^a)]}{\sigma[q(i,j)]\sigma[q(i,j+\hat{d}_{ij}^a)]} \quad (12)$$

이때 $M[\cdot]$ 는 모든 WSMAD값의 평균치이고, $\sigma[\cdot]$ 는 표준편차이다. 그리고 $q(i,j)$ 는 각 웨이브릿 부대역의 값이다. 만약 high-pass 부대역의 가중치가 low-pass 부대역에 대한 가중치보다 더 높다면 특징 기반 정합이 더 큰 효과를 가지게 되고, 반대로 low-pass 부대역의 가중치가 더 크다면 영역 기반 정합이 변이값의 결정에 주로 영향을 미친다. 그러므로 각 부대역에서 얻어진 가중치들은 영역 기반 (A^{σ_2})과 특징 기반 (D^{σ_2} , D^{σ_3} , D^{σ_4}) 정합을 동시에 고려할 수 있게 해준다. 가정합이 행해지는 초기화 과정에서는 \hat{d}_{ij}^a 가 각 부대역에 대해 서로 독립적으로 결정된다. 일단 초기 가중치가 결정되면 그 값들은 각 위치에서 하나의 변이값 d_{ij} 를 추정하기 위해 WSMAD를 구하는데 이용되어진다. 가중치들은 반복 수행 과정에서 d_{ij} 가 바뀌어 감에 따라 점차 개선되어진다. 이 값들은 실험과정에서 15회의 반복 처리만으로도 안정된 값으로 수렴함을 볼 수 있었다.

3) 국소 적응 창틀

스테레오 정합에서 변이값의 정확하고 안정된 추정을 위해 적당한 창틀 크기의 결정은 매우 중요하다. 창틀의 크기는 창틀내의 신호와 변이값의 변화 정도에 따라 결정되어야 한다. 그러나 신호의 변화는 찾아내기가 쉬운 반면 변이값의 변화는 고려하기가 매우 어렵다. 실제적으로 변이값의 경계 부분에서와 같이 창틀내에서의 변이값이 서로 다를 경우에는 정합시 많은 문제를 일으킨다.

본 논문에서는 변이값의 정합 오류를 줄이기 위한 새로운 적응 창틀 결정 방법을 제안한다. 보다 낮은 해상도에서 얻어진 변이값의 추정치와 웨이브릿 변환된 영상의 LL 대역의 값을 이용하여 각 화소에서 가장 적당한 창틀의 형태를 결정하게 된다. 우선, 추정된 변이값들로 부터 변이값의 변화가 없는 부분과 있는 부분으로 구분한다. 즉, 주변 5×5 창틀내에서 추정된 변이값이 모두 동일하면 창틀내의 변이값들이 모두 같게 되는 조건을 만족하는 최대 크기의 창틀을 그 위치에서의 창틀로 선택한다. 반면, 5×5 창틀내에 다른 변이값이 존재하면 그림 3과 같은 9가지의 창틀중 LL대역의 MAD(mean absolute difference)가 최소가 되는 형태의 창틀을 선택한 후, 창틀의 형태는 유지하면서 창틀의 크기만을 가변시켜 그 위치에서의 적응 창틀을 결정한다. 제안한 방법은 설정된 창틀내의 변이값이 서로 달라짐으로써 일어나는 잘못된 정합을 줄일 수 있으므로 보다 정확하고

안정된 변이값의 추정이 가능하다.

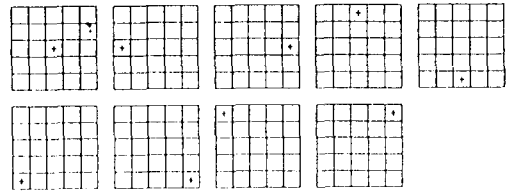


그림 3. 제안한 가변 창틀의 초기 창틀 형태(5×5). 이때 "+"는 정합 확률이 결정되어질 중심 화소의 위치이다

Fig. 3. Proposed initial window type(5×5). "+" is the center pixel to determine the matching probability.

3. 이완화 방법

Marr와 Poggio,^[5] Barnard와 Thompson,^[35] Kim과 Aggarwal^[36], 그리고 Nasrabadi와 Choo^[8]는 스테레오 정합 과정에서 이완화를 사용했다. Marr와 Poggio는 간단한 반복 과정으로 주변의 정보를 이용했고, Barnard와 Thompson은 Moravec operator에 의해 얻어진 특징점의 유사도로 부터 계산된 확률값을 반복적으로 개선해 나갔다. Kim과 Aggarwal은 변이값, 영상의 형태, 정합 확률의 연속성을 결합시킨 이완화 방법을 제안했으며, Nasrabadi와 Choo는 이완화 과정에서 Hopfield 신경망을 사용하였다. 본 논문에서는 이완화 과정에서 주변의 변이값을 고려하기 위한 이완화 방법을 제안한다. 제안하는 방법은 변이값의 동공화 현상을 방지할 수 있으며 반복 과정에서의 정규화(normalization)를 요구하지 않는다.

1) 정합 확률의 초기화

WSMAD값을 0과 1사이의 값으로 변환하기 위해 식(13)과 같은 sigmoid함수를 사용하였다.

$$p = \frac{1}{1 + \exp[\lambda(E'_i / M[E'_i] - 1)]} \quad (13)$$

$$M[E'_i] = \frac{1}{2^i \times 2^j \times D} \sum E_i \quad (14)$$

이때 $M[E'_i]$ 는 전체 WSMAD값의 평균이고 D 는 각 계층에서의 탐색 영역의 수이다. 식(13)에서 λ 값에 따라 sigmoid함수의 기울기가 결정되고 결과적으로 λ 가 커짐에 따라 변환된 확률값은 보다 넓은 분포를 가지게 된다. 그러나, λ 가 너무 클 경우에는 확률값이 $M[E'_i]$ 를 임계치로 이치화되므로 적당한

λ 를 결정하는 것이 중요하게 된다(그림 4). 그러므로 본 논문에서는 이때 E'_i 이 Gaussian분포를 가진다고 가정하고 E'_i 의 누적 분포 함수(cdf)의 값이 0.9가 될 때 λ 의 값이 0.1의 확률을 가지도록 λ 를 결정하였다. λ 는 식(15)과 같이 E'_i 의 평균과 분산만으로 쉽게 구해질 수 있다.

$$\lambda = \ln 9 \times \frac{M[E'_i]}{1.282 \times \sigma[E'_i]} \quad (15)$$

이와 같이 구해진 확률값들은 $N \times N \times D$ 개의 node로 구성된 이완화를 위한 회로망의 초기치가 된다(그림 5). 각 node간의 연결 상태는 그림 6과 같다.

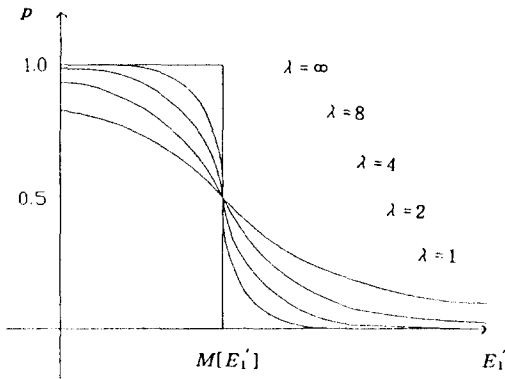


그림 4. 식(13)의 sigmoid 함수
Fig. 4. The sigmoid function as shown in (13).

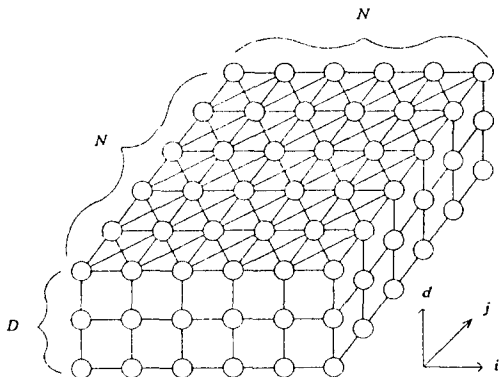


그림 5. 제안한 이완화를 위한 회로망
Fig. 5. The network for the proposed relaxation algorithm.

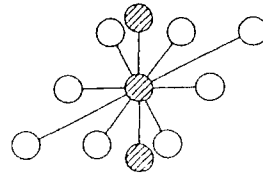


그림 6. 각 node간의 연결 상태
Fig. 6. The connection of each node.

2) 각 node에 대한 상태값의 변화
각 위치에서의 변이값의 결정이 모호한 경우에만 주변값을 고려하여 상태값을 바꾸어 준다. 위치 (i, j) 에서 변이 d 를 가지는 node의 새로운 값 $p_{i,j,d}^{t+1}$ 는 식(16)에 따라 결정된다.

$$p_{i,j,d}^{t+1} = \begin{cases} f(u + \Delta u) & \text{if } \max(p_{i,j}) - \text{sub max}(p_{i,j}) < T_p \\ p_{i,j,d}^t & \text{otherwise} \end{cases} \quad (16)$$

이때, T_p 는 상태값을 변화시키기 위한 확률값의 임계치이다.

$$f(u) = \frac{1}{1 + \exp[-\lambda(u - 0.5)]}, \quad \lambda = 4 \quad (17)$$

식(16)에서 u 는 식(17)로부터 다음과 같이 결정되어 진다.

$$u = 0.5 - 0.25 \ln \frac{1 - p_{i,j,d}^t}{p_{i,j,d}^t} \quad (18)$$

중심 화소에서의 확률값과 주변의 확률의 평균값과의 차이는

$$\Delta u = \bar{p} - p_{i,j,d}^t \quad (19)$$

이고, 이웃 node들의 확률값의 평균치 \bar{p} 는

$$\bar{p} = \frac{1}{K} \sum_r \sum_s \sum_k p_{r,s,k} \quad \text{if } k = d, \quad (20)$$

$(r, s) \neq (i, j), \text{ and } (r, s) \in N(i, j)$

이다.
이때 K 는 동일한 변이값을 가지는 주변 node의 갯수이고 $N(i, j)$ 는 (i, j) 위치 주변의 node들이다. 이와 같은 방법으로 결정된 새로운 값 $p_{i,j,d}^{t+1}$ 는 그림 7과 같이 Δu 에 의해 결정되어 진다. 또한, 식(17)의 sigmoid 함수에서 0.5 주위에서의 기울기가 1이 되도록 하기 위해서 λ 는

$$\frac{df(u)}{du} = \lambda \cdot f(u)(1 - f(u)) \quad (21)$$

$$\left. \frac{df(u)}{du} \right|_{u=0.5} = \frac{\lambda}{4} = 1 \quad (22)$$

를 만족해야 하고 그값은 4가 된다. 만약 변이값의 결정이 모호한 node가 존재하지 않거나 설정된 최대 반복 수행 한계에 도달하게 되면 반복 수행은 종결되어 지고, 그 위치에서의 확률값이 최대가 되는 d 를 그때의 변이값으로 선택하였다.

$$d_{i,j} = \text{arg max}(p_{i,j,d}) \quad (23)$$

이러한 과정은 그림 5에서 세로로 연결된 node들간에 MAXNET^[37,38]을 형성하여 구현할 수 있다. 본 논문에서는 주변 node로 8-neighbor를 고려하였고, 각 계층에서 -1과 1사이를 탐색 영역으로 하였으며, T_p 는 0.6으로 두고 실험하였다. 실제로 실험에서는 10 epochs이내에 전체 neuron이 안정된 상태에 도달함을 확인할 수 있었다. 제안한 이완화 과정에서는 동일한 변이값을 가지는 node들 간에만 강화가 일어나므로 변이값이 크게 변하는 부분에서 일어나기 쉬운 변이값의 붕괴 현상을 방지할 수 있다. 또한 각 node들의 값이 비교적 빨리 안정화 상태로 들어가고, 모호한 node들에서 발생할 수 있는 오정합을 많이 줄여준다.

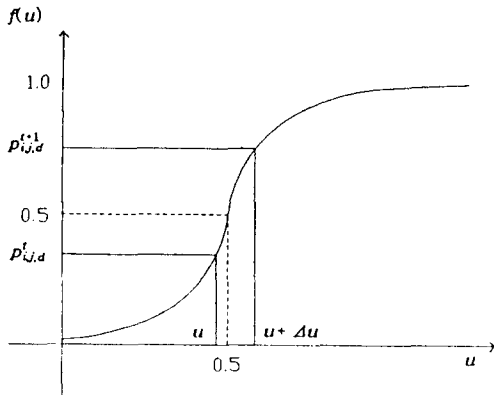


그림 7. node의 새로운 값
Fig. 7. The new value of nodes.

IV. 실험 및 고찰

본 논문에서는 제안한 정합 방법의 타당성을 확인

하기 위해 RDS 영상, 합성 영상, 실내 영상, 실외 영상등에 대해 다양하게 실험을 해 보았다. 또한, 제안한 방법들을 여러 가지로 실험함으로써 제안한 방법들이 가지는 효과를 검토하였다. 실험은 크게 3가지로 구분할 수 있는데

- 1) WT된 영상에서 가중치 없이 계층적 구조만을 사용하여 정합을 시도하는 방법 (SMW-FW)
- 2) 가중치를 가지고 반복적인 방법으로 계층적 정합을 사용한 방법 (SMW-VW)
- 3) 국소 적응 창틀과 가중치를 모두 가지고 이완화 과정을 거치는 방법 (SMW-RN) 이다.

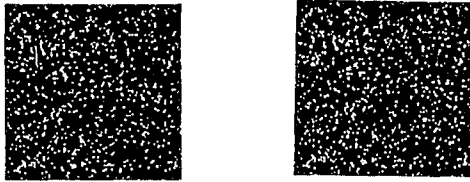
1. RDS와 합성영상

실험을 위해서 그림 8에서 그림 12까지의 잡음이 섞이지 않은 5% RDS, 30% RDS 영상에 대해 우영상에 10%의 잡음이 섞인 영상, 50% RDS 영상에 20%의 잡음이 섞인 영상, 밝기값이 0 255 사이에서 uniform 분포를 가지는 random gray dot 영상에 표준편차가 50인 Gaussian 분포를 가지는 백색 잡음이 실린 영상, 서로 다른 간격을 가진 줄무늬가 합성된 영상에 20%의 잡음이 섞인 영상을 사용했다. 위의 모든 영상은 128×128 크기이고 4층의 wedding cake 형태의 변이값을 가지고 있다. 이때 각 방법의 성능을 객관적으로 평가하기 위한 값으로는 변이값 d_{ij} 와 이상적인 변이값 \bar{d}_{ij} 사이의 SAE (sum of absolute error)와 MSE(mean squared error)를 사용했다.

$$SAE: s_1 = \sum_i \sum_j |d_{ij} - \bar{d}_{ij}| \quad (24)$$

$$MSE: s_2 = \frac{1}{N^2} \sum_i \sum_j (d_{ij} - \bar{d}_{ij})^2 \quad (25)$$

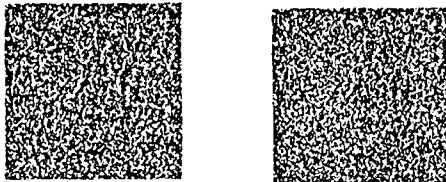
이때 모든 방법의 최대 탐색 영역은 0~15 화소로 하였다. 제안한 계층적 방법에서는 4개의 계층을 사용하였고, 각 계층당 탐색 영역을 0~1 화소로 하였다. 가변 창틀을 사용하는 SMW-RN 방법에서는 최대 창틀의 크기를 7×7로 하였다. 뿐만 아니라, 후처리 과정없이 전 영상의 변이값을 얻을 수 있는 NCC (normalized cross-correlation)을 유사 함수로 사용한 영역 기반 정합 방법의 결과를 제안한 방법들과 비교하였다. 기존의 방법과 제안한 3가지 방법을 각 영상에 적용했을때의 실험 결과가 그림 13, 그림 14, 표.1 에 주어져 있다. 실험 결과를 보면, 제안한 방법이 정합에 불리한 조건들을 가진 영상들에 대해서도 매우 뛰어난 결과를 가져옴을 알 수 있다. 즉,



(a) (b)

그림 8. 5% RDS (a) 좌영상 (b) 우영상

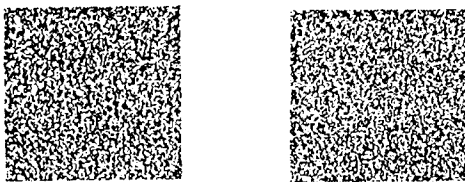
Fig. 8. 5% RDS. (a) Left image. (b) Right image.



(a) (b)

그림 9. 우영상에 10%의 잡음이 실린 30% RDS (a) 좌영상 (b) 우영상

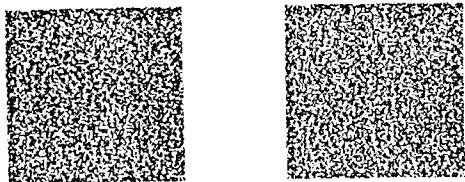
Fig. 9. 30% RDS. 10% dots of right image are decorrelated. (a) Left image. (b) Right image.



(a) (b)

그림 10. 우영상에 20%의 잡음이 실린 50% RDS (a) 좌영상 (b) 우영상

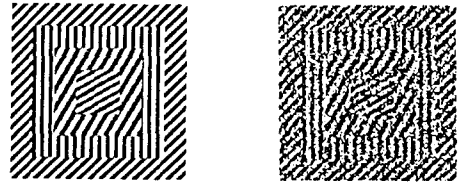
Fig. 10. 50% RDS. 20% dots of right image are decorrelated. (a) Left image. (b) Right image.



(a) (b)

그림 11. 표준편차가 50인 white Gaussian 잡음이 실린 gray dot 영상이 실린 gray dot 영상 (a) 좌영상 (b) 우영상

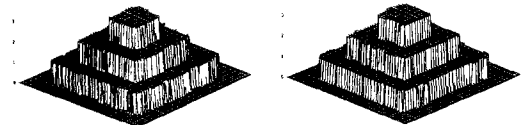
Fig. 11. The gray dot image with white Gaussian noise in which the standard deviation is 50. (a) Left image. (b) Right image.



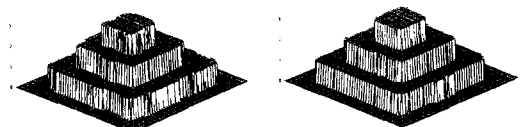
(a) (b)

그림 12. 우영상에 20%의 잡음이 실린 "Stripes" 영상 (a) 좌영상 (b) 우영상

Fig. 12. "Stripes" image. 20% dots of right image are decorrelated. (a) Left image. (b) Right image.



(a) (b)



(c) (d)

그림 13. RDS 영상들과 합성 영상에 대해 제안한 방법(SMW-RN)을 사용해서 얻어진 결과 변이값 (a) 5% RDS(그림 8) (b) 30% RDS(그림 9) (c) 50% RDS(그림 10) (d) Gray dot 영상(그림 11)

Fig. 13. The disparity map for RDSs and synthetic images. The results is obtained using the proposed algorithm (SMW-RN). (a) 5% RDS as shown in Fig. 8. (b) 30% RDS as shown in Fig. 9. (c) 50% RDS as shown in Fig. 10. (d) Gray dot image as shown in Fig. 11.

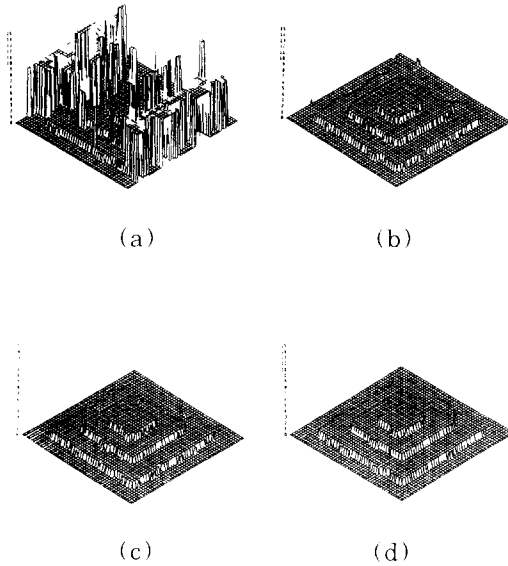


그림 14. "Stripes" 영상(그림 12)에 대해 다양한 실험 방법으로 얻어진 결과 변이값 (a) NCC(기존의 영역 기반 정합 방법으로 얻어진 가장 좋은 결과) (b) SMW-FW 방법으로 얻어진 가장 좋은 결과(창틀의 크기가 11×11 일때) (c) SMW-VW 방법으로 얻어진 가장 좋은 결과(창틀의 크기가 9×9 일때) (d) SMW-RN(가변 창틀을 사용하였고, 최대 창틀의 크기는 7×7 로 두었다)

Fig. 14. The experimental results for "Stripes" image as shown in Fig. 12. They are obtained by the various methods. Each result is the best one obtained for each method. (a) NCC (the best result obtained by the classical area-based method). (b) SMW-FW. The window size is 11×11 in which the result have best performance. (c) SMW-VW. The window size is 9×9 . (d) SMW-RN. The variable window is used and the maximum window size is 7×7 .

제안한 방법은 특징점의 수가 부적당하거나, 심한 잡음이 실린 경우, 또는 동일한 형태가 반복적으로 나타남으로 해서 오정합의 가능성이 큰 경우에도 적용될 수 있다. 특히 이완화 과정을 포함한 방법은 후처

리 과정없이도 정확한 정합 결과를 얻을 수 있었다.

표 1. RDS 영상들과 합성 영상에 대해 제안한 방법(SMW-RN)으로 얻어진 결과

Table 1. The result of the proposed stereo matching algorithm (SMW-RN) for RDSs and synthetic images.

IMAGES	NOISE	MSE	SAE
5% RDS	none	0.045	730
30% RDS	10% decorrelated	0.031	511
50% RDS	20% decorrelated	0.039	643
Gray Dot	Gaussian ($\sigma = 50$)	0.026	421
Stripes	20% decorrelated	0.055	907

2. 실영상

그림 15과 그림 16에 보여지는 실내 영상("bear")과 실외 영상("pentagon")이 실험에 사용되었다. "bear"영상과 "pentagon"영상에 대한 탐색영역은 각각 0 15와 -31 31로 설정하였다. 이완화 과정을 포함한 실험 결과(SMW-RN)는 그림 17과 그림 18에 나타내었다. 실험 결과는 제안한 방법이 변이값의 몽롱화 현상없이 실영상에 잘 적용됨을 알 수 있었다. 실험 결과에 따르면 특징점의 수가 많거나 적은 영역 모두에서 정합이 효과적으로 이루어 졌으며, 변이값의 경계 부분에서 심각하게 나타나는 오정합 현상이 발생하지 않았다.



(a)

(b)

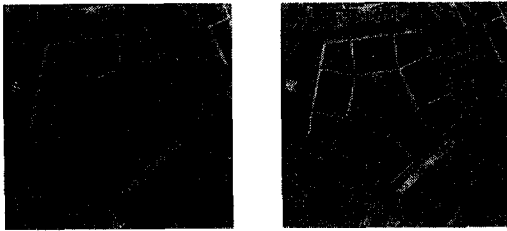
그림 15. "Bear" 영상 (200×200)

(a) 좌영상 (b) 우영상

Fig. 15. "Bear" image (200×200)

(a) Left image. (b) Right image.

V. 결론



(a) (b)

그림 16. "Pentagon" 영상 (512×512)

(a) 좌영상 (b) 우영상

Fig. 16. "Pentagon" image (512×512)

(a) Left image. (b) Right image.

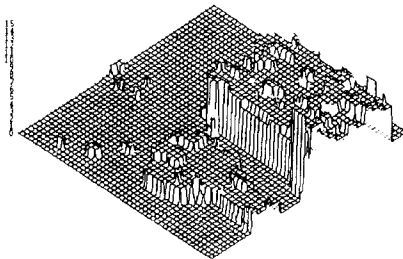


그림 17. "Bear" 영상에 대해 제안한 방법(SMW-RN)을 사용한 실험 결과

Fig. 17. The experimental result for "Bear" image. The proposed algorithm (SMW-RN) is used to obtain the result.

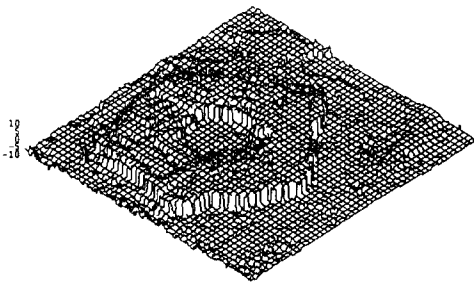


그림 18. "Pentagon" 영상에 대해 제안한 방법(SMW-RN)을 사용한 실험 결과

Fig. 18. The experimental result for "Pentagon" image. The proposed algorithm (SMW-RN) is used to obtain the result.

본 논문에서는 영역 기반과 특징 기반 정합 방법이 적응적으로 고려될 수 있는 계층적 스테레오 정합 방법이 제안되었다. 제안한 방법은 전처리나 후처리 과정없이 얻어진 각 웨이브릿 부대역에 대해 국소 적응 가중치와 각 화소에 대한 가변 창틀을 사용하였으며, 계층적 구조를 형성하기 위해 non-stationary 신호의 다해상도 해석에 유용한 웨이브릿 변환을 사용하였다. DWT는 특징값을 보존하면서 잡음을 제거할 수 있고, 변환된 값이 각각 뭉뚱화된 영상, 수평 경계선, 수직 경계선, 대각 경계선의 특징을 나타낸다. 각 부대역에 대한 가중치는 좌 우 영상의 각 부대역의 유사도로부터 결정되어지므로 가중치에 따라 각 특징값들의 상대적인 중요도를 적응적으로 고려할 수 있다. 가변 창틀의 크기와 형태는 웨이브릿 변환된 영상의 LL 부대역과 해상도가 낮은 계층에서 추정된 변이값으로부터 결정되어 지고, 각 웨이브릿 부대역에 대해 국소 적응 가중치와 가변 창틀을 적용해서 얻어진 정합 확률에 대해 주변의 변이값을 고려하기 위한 새로운 이완화 과정이 제안되어 진다. 제안한 가변 창틀을 사용할 경우 변이값의 경계 부분에서의 오정합을 감소시킬 수 있었으며, 이완화를 거친 변이값들은 뭉뚱화 현상없이 오정합이 감소되었다. RDS, 합성 영상, 실내 영상, 실외 영상등의 다양한 영상에 대한 실험 결과에 따르면, 심한 잡음이나 반복적으로 나타나는 동일한 형태를 가진 영상이나, 혹은 특징점의 수가 너무 적거나 많은 경우에도 제안한 스테레오 정합 방법이 잘 적용됨을 알 수 있었다. 한편, 제안한 이완화 과정에 대한 수학적 분석과 적응 창틀의 설정 과정이 전체 알고리즘에 미치는 영향에 대한 연구가 계속 수행되어 지고 있다.

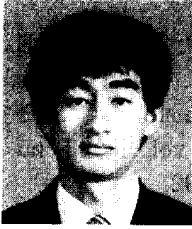
參考文獻

- [1] K. Sugihara, "Three principle in stereo vision," *Robotics*, vol. 1, no. 4, pp. 391-400, 1986.
- [2] B. K. P. Horn, *Robot Vision*, McGraw-Hill, 1987.
- [3] R. M. Haralick and L. G. Shapiro, *Computer And Robot Vision*, Addison-Wesley Publishing Co., 1993.
- [4] U. R. Dhond and J. K. Aggarwal, "Structure from stereo : a review," *IEEE Trans. Syst. Man Cybern.*, vol.

- 19, no. 6, pp. 1489-1510, Nov./Dec. 1989.
- [5] D. Marr and T. Poggio, "A computational theory of human stereo vision." *Proc. Roy. Soc. London*, B204, pp. 301-308, 1979.
- [6] S. B. Pollard, J. E. W. Mayhew, and J. P. Frisby, "PMF : A stereo correspondence algorithm using a disparity gradient limit." *Perception*, vol. 14, pp. 449-470, 1985.
- [7] G. Medioni and R. Nevatia, "Segment-based stereo matching." *Compt. Vision Graphics Image Processing*, vol. 31, pp. 2-18, 1985.
- [8] N. M. Nasrabadi and C. Y. Choo, "Hopfield network for stereo vision correspondence." *IEEE trans. Neural Networks*, vol. 3, no. 1, Jan. 1992.
- [9] H. P. Moravec, "Towards automatic visual obstacle avoidance." *Proc. 5th Int. Joint Conf. Artificial Intell.*, 1977
- [10] L. Matthies, "Stereo vision for planary rovers : stochastic modeling to near real-time implementation." *Int. J. Computer Vision*, vol. 8, no. 1, pp. 71-91, 1992.
- [11] M. D. Levine, D. A. O'Handley and G. M. Yagi, "Computer determination of depth maps." *Comput. Graphics, Image Processing*, vol. 2, pp. 131-150, 1973.
- [12] M. Okutomi and T. Kanade, "A multiple-baseline stereo." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-15, no. 4, pp. 353-363, Apr. 1993.
- [13] D. Terzopoulos, "Multilevel computational processes for visual surface reconstruction." *Compt. Vision Graphics Image Processing*, vol. 24, pp. 52-96, 1983.
- [14] J. J. Lee, S. J. Cho, and Y. H. Ha, "Neural network modeling of new energy function for stereo matching." *SPIE. Intel. Rob. and Comp. Vis. X*, vol. 1608, Nov. 1991.
- [15] S. B. Marapane and M. M. Trivedi, "Region-based stereo analysis for robotic applications." *IEEE Trans. Syst. Man Cybern.*, vol. 19, no. 6, pp. 1447-1464, Nov./Dec. 1989.
- [16] D. J. Fleet, A. D. Jepson, and M. R. M. Jehkin, "Phase-based disparity measurement." *CVGIP : Image understanding*, vol. 53, no. 2, pp. 198-210, Mar. 1991.
- [17] M. M. Fleck, "A topological stereo matcher." *Intern. J. Comp. Vision*, vol. 6, no. 3, pp. 197-226, 1991.
- [18] Y. C. Cheng and S. Y. In, "Waveform correlation by tree matching." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-7, no. 3, pp. 299-305, Apr. 1985.
- [19] T. D. Sanger, "Stereo disparity computation using Gabor filters". *Biol. Cybern.* 59, pp. 405-418, 1988.
- [20] A. Knotanzad, A. Bokil, and Y. W. Lee, "Stereopsis by constraint learning feed-forward neural networks." *IEEE Trans. Neural Networks*, vol. 4, no. 2, Mar. 1993.
- [21] J. Weng, N. Ahuja and T. S. Huang, "Matching two perspective views." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-14, no. 8, pp. 806-825, Aug. 1992.
- [22] S. D. Cochran and G. Medioni, "3D surface description from binocular stereo." *IEEE Trans. Pattern Anal. Machine intell.*, vol. PAMI-14, no. 10, pp. 981-994, Oct. 1992.
- [23] M. Okutomi and T. Kanade, "A locally adaptive window for signal matching." *Intern. J. Comp. Vision*, vol. 7, no. 2, pp. 143-162, 1992.
- [24] O. Rioul and M. Vetterli, "Wavelets and signal processing." *IEEE SP magazine*, pp. 14-38, Oct. 1991.
- [25] A. Cohen, I. Daubechies, and J. C. Feauveau, "Biorthogonal bases of

- compactly supported wavelets." *Comm. Pure Appl. Math.*, vol. 45, pp. 485-560, 1992.
- [26] S. G. Mallat, "A theory for multiresolution signal decomposition : the wavelet representation." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-11, no. 7, pp. 674-693, July 1989.
- [27] R. Wilson, A. D. Calway, and E. R. S. Pearson, "A generalized wavelet transform for Fourier analysis : the multiresolution Fourier transform and its application to image and audio signal analysis." *IEEE Trans. Inform. Theory*, vol. IT-38, no. 2, pp. 674-690, Mar. 1992.
- [28] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform." *IEEE Trans. Image Processing*, vol. 1, no. 2, pp. 205-220, Apr. 1992.
- [29] R. A. DeVore, B. Jawerth, and B. J. Lucier, "Image compression through wavelet transform coding." *IEEE Trans. Inform. Theory*, vol. IT-38, no. 2, pp. 719-746, Mar. 1992.
- [30] S. G. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets." *IEEE Trans. Inform. Theory*, vol. IT-38, no. 2, pp. 617-643, Mar. 1992.
- [31] S. G. Mallat and S. Zhong, "Characterization of signals from multiscale edge s." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-14, no. 7, pp. 710-732, July 1992.
- [32] I. Daubechies, "Orthonormal bases of compactly supported wavelets." *Commun. Pure Appl. Math.*, vol. 41, pp. 909-996, Nov. 1988.
- [33] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms." *IEEE Trans. Inform. Theory*, vol. IT-38, no. 2, pp. 587-607, Mar. 1992.
- [34] M. Mougeot, R. Azencott, and B. Angeniol, "Image compression with back propagation : improvement of the visual restoration using different cost functions." *Neural Networks*, vol. 4, pp. 467-476, 1991.
- [35] S. T. Barnard and W. B. Thomson, "Disparity analysis of images." *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, no. 4, pp. 333-340, July 1980.
- [36] Y. C. Kim and J. K. Aggarwal, "Positioning 3-D objects using stereo images." *IEEE J. Robotics and Automation*, vol. RA-3, no. 4, pp. 361-373, Aug. 1987.
- [37] J. A. Freeman and D. M. Skapura, *Neural Networks : Algorithms, Applications, and Programming Techniques*. Addison-Wesley Publishing Co., 1991.
- [38] J. M. Zurada, *Introduction to Artificial Neural Systems*. West Publishing Co., 1992.

— 著 者 紹 介 —



金 容 奭(正會員)

1968年 6月 20日生. 1991年 2月
 경북대학교 전자공학과 졸업 공학
 사 학위 취득. 1994年 2月 경북대
 학교 대학원 전자공학과 공학 석
 사학위 취득. 1994年 3月 ~ 현재
 경북대학교 대학원 전자공학과 박
 사과정. 주관심 분야는 영상 처리, 컴퓨터 비전, 신
 호처리, 신경 회로망 등임.

李 駿 宰(正會員) 第 31卷 B編 第 5號 參照

현재 경북대학교 전자공학과 박사
 과정 수료

•

河 永 浩(正會員) 第 31卷 B編 第 5號 參照

현재 경북대학교 전자공학과 부교수