

論文95-32B-1-5

정합 일치성을 이용한 반복 이완 스테레오 정합

(A Stereo Matching by the Iterative Relaxation Using the Consensus of Matching Possibility)

李 旺 國^{*}, 金 容 奭^{**}, 都 敬 勳^{**}, 河 永 浩^{**}
 (Wang-Gug Lee, Yong-Suk Kim, Kyeong-Hoon Do, and Yeong-Ho Ha)

요 약

스테레오 시각은 서로 다른 위치에서 획득한 두 영상으로부터 3차원 거리 정보를 추출하는데 유용하다. 본 논문에서는 스테레오 정합을 위한 대응점 추출을 위해 intra-주사선(scanline) 제한조건으로부터 대응점 탐색 영역을 줄이고, 이완 과정을 위해 inter-주사선 제한조건과 인접하는 화소와의 변이의 연속성을 이용하였다. 3차원 배열 구조를 가지는 노드들은 좌·우영상의 변이축 상에 위치하며, 각 노드들은 두 영상에 위치하는 화소들 간의 정합가능도를 가진다. 좌·우영상의 각 변이축의 교차점에 위치하는 노드의 정합가능도(matching possibility)가 최대값일 때는 정합이 이루어지며, 최대값이 아니면 주위 노드의 정합가능도로부터 새로운 값으로 갱신되는 이완 과정을 가진다. 양안의 경쟁을 통한 이완 과정이 무작위로 선택된 노드에 대하여 반복적으로 수행되며, 정합 가능성을 두 영상의 변이축을 따라 동시에 고려함으로써 정합의 신뢰성이 높고 물체의 경계선 부분에서 일어나는 뭉롱화 현상을 제거한다. 랜덤 도트 영상과 항공 촬영 스테레오 영상에 대한 실험에서 좋은 정합 결과를 얻을 수 있었다.

Abstract

Stereo vision is useful to obtain three dimensional depth information from two images taken from different view points. In this paper, we reduce searching area for correspondence by using the intra-scanline constraint, and utilize the inter-scanline constraint and the property of disparity continuity among the neighboring pixels for relaxation. Nodes with 3-D structure are located on the axes of two views, and have matching possibility of correspondent pixels of two images. A matching is accepted if a node at the intersection of the disparity axes has the greatest matching possibility. Otherwise, the matching possibility of the node is updated by relaxation with the cooperation of neighboring nodes. Further relaxation with competition of two views is applied to a matching possibility of randomly selected node. The consensus of two views increases the confidence of matching, and removes a blurring phenomenon on the discontinuity of object. This approach has been tested with various types of image such as random dot stereogram and aerial image, and the experimental results show good matching performance.

* 正會員, 國防科學研究所
 (Agency for Defence Development)

** 正會員, 慶北大學校 電子工學科

(Department of Electronic Engineering, Kyungpook National University)

接受日字: 1994年 5月 16日

I. 서 론

컴퓨터 비전의 궁극적 목적은 컴퓨터에 인간의 시각 기능을 부여하는 것이다. 시각 정보는 인간의 시각 활동에 아주 중요한 위치를 차지하므로 컴퓨터 비전 분야에 수많은 연구가 진행되어 왔으며, 서로 다른 위치에서 획득한 2개의 2차원 영상으로부터 3차원 정보를 추출하는 스테레오 시각이 그 한 분야를 이루어 왔다.^[1]

스테레오 시각은 영상 획득, 카메라 모델링, 대응점 추출, 3차원 깊이 정보 추출 등 일련의 과정을 거치며, 좌·우 두 영상으로부터 대응점을 추출하는 정합 과정이 가장 어렵고 대부분의 연산 시간이 소요된다. 스테레오 정합 과정은 수많은 연산이 필요하여 실제적인 응용이 힘들었으나, 프로세서의 속도 및 병렬 처리 방법의 발전으로 더욱 복잡하고 강력한 스테레오 모델의 개발이 가능하게 되었다. 정합을 위한 대응점의 추출은 적절한 제약 조건들을 사용하여 가능한 해의 수를 제한한 뒤 최적화 알고리즘을 적용한다. 최적화 방법은 주어진 제한 조건과 잘 일치할 때 비용 함수가 최소가 되도록 함수를 정의하고 이 함수의 최소값을 찾음으로써 문제를 해결한다. 이와 같은 최적화 방법에 기초를 둔 정칙화(regularization) 방법은 스테레오 정합뿐만 아니라 Surface reconstruction, optical flow, Edge detection 등에서도 유사한 형태의 에너지 식을 사용한 방법이 시도되었다^[2,3]. 확률적 방법과 MRF(Markovian Random Field) 모형에 기초를 두고 영상 처리에 적용한 방법은 최적화 문제를 simulated annealing과 같은 통계적으로 동작하는 일반적인 방법으로 해결하였으나 에너지가 증가하는 방향으로도 상태가 변할 수 있기 때문에 수렴 속도가 느린 단점을 가진다^[4,5]. 좌·우영상의 대응점을 결정하는 스테레오 정합 과정에서 대응점의 후보 수를 줄이고 인접하는 화소의 변이의 연속성을 고려한 Ohta 등^[6]은 intra- 및 inter-주사선 제한 조건을 사용하여 회로를 구성한 뒤 Dynamic programming 기법을 도입해 주어진 비용 함수가 최소가 되는 최적 경로를 2차원 또는 3차원 탐색 방법으로 찾아 대응점을 결정하였다. 이와 같은 방법은 병렬 처리에 제한점이 많으며, 빠른 시간에 많은 양의 연산을 필요로 하는 스테레오 정합은 기존의 Von-Neumann 방식의 컴퓨터로는 실시간 처리가 어려워, 최근에는 병렬 회로망을 이용한 모델에 대한 연구가 많이 진행되어 왔다.

스테레오 정합 문제와 모션(motion)문제를 해결하기 위해 Hopfield 회로를 이용한 Zhou^[7] 등은 제한 조건에 기초하여 에너지 함수를 정의하고 에너지 함수

로부터 연결선의 강도를 구한 후 반복 계산 과정을 거쳐 에너지 함수의 값을 최소화하는 방법을 제안하였다. Hopfield 회로망을 이용한 방법은 에너지가 감소하는 방향으로만 상태를 변화시켜 나가기 때문에 통계적으로 동작하는 simulated annealing 방법에 비해 빠른 수렴 속도를 가지나 국부 극소(local minimum)에 빠질 우려가 있으며 중복 정합의 가능성이 존재한다.

양안의 정합 정보를 고려하는 회로망 구조로는 Marr^[8,9] 등이 제안한 인간의 시각 체계를 바탕으로 세 가지의 제한 조건을 만족하는 협동 정합 알고리즘(cooperative matching algorithm)이 있으며, Cochran^[10] 등은 영역 기반 정합으로 양안에서의 정합 상관도를 고려하면서 얻어진 변이 정보에 영상의 특징을 이용하여 불연속점 부근에서의 변이값을 개선하였다. Khotanzad^[11] 등은 Marr와 Poggio가 제안한 대응 화소의 유일성 및 변이의 연속성 제한 조건을 역전파(Back-propagation) 학습 방법에 적용하여 비반복적 학습 방법으로 랜덤 도트 영상에 적용하여 실시간 처리가 가능하였으나 2비트 이상으로 양자화 되는 실제 영상에의 적용이 어렵다.

본 연구에서는 intra-주사선 제한 조건으로부터 대응점의 탐색 범위를 줄이고 inter-주사선 제한 조건 및 서로 인접하는 화소의 변이의 연속성을 이용하며, 두 영상의 변이축 상에서 양방향 정합 가능성을 동시에 고려한 반복적 이완 정합 방법을 제안한다. 제안하는 이완화 방법은 각 영상의 변이축을 따라 가지는 정합가능도를 두 영상의 변이축이 교차하는 지점에 위치하는 노드와 비교하여 정합의 신뢰도가 떨어지는 경우는 정합의 가능성을 억제한다. 정합의 신뢰도가 높은 경우에는 변이의 연속성을 유지하면서 그 주위 노드로 정합의 가능성을 전파하게 되며 주위 값들에 따라 자신의 정합 가능성을 반복적으로 개선한다. 두 영상의 변이축 상에 위치하는 노드들의 정합 일치성에 의한 대응점 결정 방법은 공간적으로 좁은 범위에서 협동하는 과정과 공간적으로 떨어진 범위에서 경쟁하는 과정으로 이루어져 있으며 이 두 가지 반복적 협동-경쟁에 의한 이완 과정에 의해 목적하는 대응점을 찾게 된다. 이와 같이 이완화 과정에서 양안의 정합 가능성을 동시에 고려하여 물체의 경계선 부근에서 일어나는 몽롱화(blurring) 현상이 제거되고 정합의 신뢰성이 높다. 또한, 양안에서 얻은 각각의 변이 정보로부터 불연속성을 나타내는 물체의 경계부에서의 중복 정합을 피하고 물체의 인식 과정에 유용한 정보를 얻을 수 있다.

II. 변이 정보 추출을 위한 반복 이완 정합 알고리즘

1. 스테레오 정합을 위한 카메라 모델링

스테레오 정합은 좌측 영상의 한 화소와 대응되는 화소를 우측 영상에서 탐색하는 문제로 정의할 수 있으며, 영상 획득시 카메라의 설치 방법에 따라 대응점 후보의 수를 감소시키고 정합 전략을 간략화 할 수 있다. 그림 1에서와 같이 극상선(epipolar line)이 수평 주사선과 평행하게 두 카메라를 설치함으로써 대응점 탐색의 범위를 영상 전체에서 수평 주사선 내로 제한할 수 있다. 즉, 좌영상의 한 화소와 대응되는 우영상에서의 대응점은 좌영상의 화소가 위치하는 주사선과 동일한 우영상의 주사선 상에 존재하게 되고, 영상에 존재하는 물체의 거리를 선형 지식 또는 응용 목적에 따라 제한할 수 있으면 수평 주사선 내에서의 탐색 범위도 줄일 수 있게 되어 대응점 탐색의 영역이 대폭 줄어들게 된다.

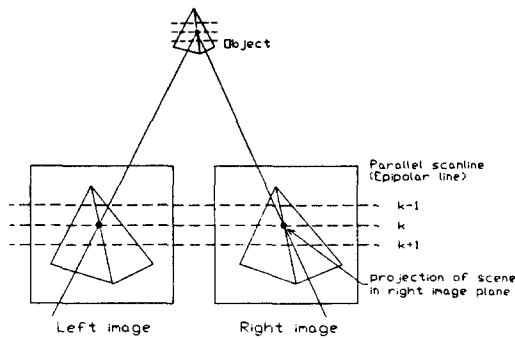


그림 1. 스테레오 정합을 위한 제한 조건
Fig. 1. Constraints for stereo matching.

대응점의 결정은 영상에 존재하는 잡음 등으로 인해 수평 주사선 탐색만으로는 많은 정합 오류를 가져오게 된다. 이와 같은 정합의 오류를 제거하기 위해 그림 1과 같이 좌·우영상의 k번째 주사선 내에서 대응되는 대응쌍이 존재할 경우, 그 인접하는 k-1 혹은 k+1번째 주사선 상의 동일한 변이값 위치에서 서로 대응되는 대응쌍이 존재할 가능성이 크다는 변이의 연속성을 이용한다. 즉, 인접하는 대응 후보쌍의 정합 가능성을 고려해 반복적인 이완 방법으로 자신의 정합가능도를 갱신함으로써 잡음 등으로 인한 오정합을 제거한다.

이상의 제한 조건을 이용하여 Henderson^[12] 등은 영상의 모든 주사선에 대하여 순차적으로 정합 과정을 진행하며, 한 주사선에서의 정합 결과를 다음 주사선에서의 정합 과정에서 참조하였다. 이러한 방법은 한 주사선에서 발생한 정합 오류가 전체에 전파될 가능성이

있다. Baker^[13] 등은 영상의 모든 주사선에 대하여 독립적으로 정합을 실시한 후 정합 결과를 개선해 나가는 과정에서 주변에서의 결과와 협동하는 과정을 적용하였으나 inter-주사선 제한 조건을 탐색 과정에서 직접 적용하지 않아 최적 탐색 방법으로 보기 어렵다. Ohta^[6] 등은 두 가지 제한 조건을 대응점 결정의 탐색 과정에서 동시에 고려하였으나 Dynamic programming 방법을 적용하여 얻은 변이 정보의 수평 주사선 방향으로 많은 오정합이 발생한다.

2. 제안한 스테레오 정합 모델

이와 같은 정합 오류를 배제하고 변이의 연속성을 유지하면서 신뢰성 있는 정합을 위해, 위에서 언급된 두 가지의 제한 조건과 두 영상의 변이축 상에 위치하는 노드들의 양방향 정합 일치성을 이용한 정합 회로망 모델을 제안한다. 첫 번째 제한 조건으로부터 정합 회로망 모델의 기하학적 좌표계는 그림 2(a)와 같다. 그림 2(a)에서 K-축과 I-축이 이루는 면이 좌영상이고 K-축과 J-축이 이루는 면이 우영상이며, 각 영상은

$$g_L(k, i) = f_L(k, i) + n_L(k, i)$$

$$g_R(k, j) = f_R(k, j) + n_R(k, j) \quad (1)$$

과 같이 주어진다. 여기서, f_L 과 f_R 은 3차원 공간상에 위치하는 한 점이 잡음이나 카메라의 왜곡이 없이 좌·우영상에 각각 투영되었을 때의 명암도이며, k 는 좌·우영상의 행, i 및 j 는 각 영상의 열의 위치를 나타내며 $1 \leq k \leq \text{NROW}$, $1 \leq i, j \leq \text{NCOL}$ 의 값을 가진다. 그리고, NROW 및 NCOL은 각 영상의 행 및 열의 크기이며, n_L 과 n_R 은 영상에 존재하는 잡음이나 카메라의 왜곡을 나타낸다.

좌표계의 각 (k, i, j) 좌표에 하나의 노드가 위치하며, I-축 및 J-축에 45° 방향이고 원점을 지나는 면에 위치하는 노드들의 변이값 d 는 0으로 주어진다. 변이가 0인 면과 평행하며 거리가 먼 곳의 면에 위치할수록 노드의 변이값은 증가하게 된다. 이때, d 는 $j-i$ 의 값으로 주어지며 최대 변이값이 d_{max} 로 제한되는 경우 $|d| \leq d_{\text{max}}$ 이다. 최대 변이가 d_{max} 로 제한되는 경우, 그림 2(b)에서와 같이 두 영상의 k번째 행에서 좌영상의 i번째 화소와 우영상의 $(i+d)$ 번째 화소와의 정합가능도를 가지는 노드 $N_{k, i, d}^L$ 은 I-축의 i번째 변이축을 따라 위치하게 되며, 우영상의 j번째 화소와 대응되는 좌영상의 화소들 간의 정합가능도를 가지는 노드들은 J-축의 j번째 변이축을 따라 위치한다.

이때, J-축의 j번째 변이축 상의 변이 d 에 위치하는 노드는 I-축의 $(j-d)$ 번째 변이축 상의 변이 d 에 위치

하는 노드와 동일하며

$$N'_{k,i,d} = N'_{k,i,d,d} \quad (2)$$

와 같은 등식을 만족한다. 식 (2)에서와 같이 3차원 형태의 배열을 이루고 있는 각 노드는 좌측 및 우측 영상의 변이축에서의 좌표값으로 표현이 가능하며, 본 논문에서는 설명의 편의를 위해서 좌영상에서의 변이축에 대한 좌표값으로 표현하기로 한다.

같이 정합 창틀내에 위치하는 좌·우영상의 평균과 분산으로 정규화한 NCC(normalized cross-correlation)값을 구한 후, 이를 0과 1사이의 정규화된 값으로 주어진다.

$$P'_{k,i,d}(0) = \frac{\sum_l (g_l(k,i) - \mu_l)(g_R(k,i+d) - \mu_R)}{\{ \sum_l (g_l(k,i) - \mu_l)^2 \sum_l (g_R(k,i+d) - \mu_R)^2 \}^{1/2}} \quad (3)$$

여기서, A는 좌영상의 화소 $g_l(k, i)$ 와 우영상의 화소 $g_R(k, i+d)$ 를 중심으로 하는 상관관계 창틀의 영역에 포함되는 모든 화소에 대한 연산을 의미한다. μ_l 및 μ_R 은 좌·우영상의 상관관계 창틀내 화소들의 평균값이며, 0은 초기값임을 나타낸다. 또한, 우영상에서의 변이축을 따라 위치하는 각 노드에서의 정합가능도 $P'_{k,i,d}(0)$ 는 식 (3)과 동일한 방법으로 주어진다.

3. 반복 이완 정합 알고리즘

1) 정합 시스템의 동작

그림 3은 제안한 스테레오 정합 시스템의 전체적인 동작 과정을 보여준다. 먼저 입력된 좌·우영상에 가우시안 필터(Gaussian filter)를 적용하여 영상에 존재하는 잡음을 제거한 후 식(3)에 의해 3차원 배열 형태를 이루고 있는 모든 노드들에 대한 초기 정합가능도를 구한다.

이완화 과정은 무작위로 선택된 노드에 대해 순차적으로 수행되며, 전체 노드의 갯수만큼 노드를 무작위로 선택하여 각 노드에서의 정합가능도를 갱신하였을 때 1회의 반복 과정을 마치게 된다. 무작위로 선택된 노드 $N'_{k,i,d}$ 에서 교차되는 좌영상 및 우영상의 변이축을 따라 $-d_{max} \sim d_{max}$ 내에 위치하는 노드들의 정합가능도를 노드 $N'_{k,i,d}$ 의 값과 비교하여 정합가능도의 갱신 여부를 결정한 후 인접하는 노드의 정합가능도를 고려한 이완 과정을 거친다. 반복 과정은 미리 설정된 반복 횟수에 도달하거나 정합가능도가 수렴할 때까지 수행된다. 이와 같은 이완 과정을 반복함으로써 초기 정합가능도가 가지는 오류를 제거하고, 두 영상의 변이축 상에 위치하는 노드들에 대한 정합의 일치성을 동시에 고려하므로 정합 결과의 신뢰성을 향상시킨다. 여기에서, 정합가능도를 갱신할 때 고려하는 주변의 노드들은 Marr^[8] 등이 제안한 여기 강도(excitatory connection)로 연결되는 노드들과 동일하며, 각 변이축 상에 놓인 노드들은 억제 강도(inhibitory connection)로 연결되는 노드들과 동일한 형태를 가진다.

Marr 등은 노드의 초기 활성화 상태를 좌·우영상의 대응되는 각 화소의 배타적 논리값(XOR)으로 결정하였고, 각 노드에 대한 갱신은 여기 강도 및 억제 강도

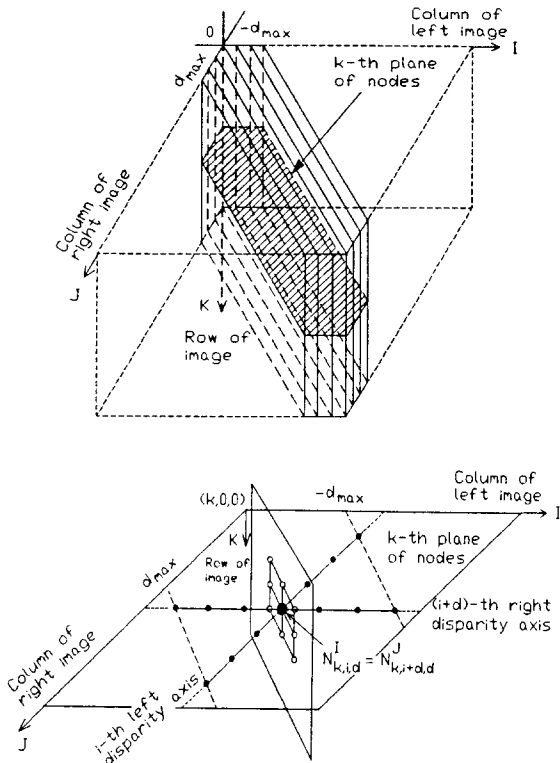


그림 2. 정합 회로망 모델

- (a) 회로망 모델의 좌표계
- (b) 노드의 배열 형태

Fig. 2. The network model for stereo matching.

- (a) The coordinates of the model.
- (b) The arrangement of nodes.

노드 $N'_{k,i,d}$ 의 정합가능도는 인접 화소간의 변이의 연속성으로부터 갱신될 수 있으며, 식 (1)에 존재하는 잡음으로 인한 오정합의 결과는 주위의 정합가능도를 고려한 반복적인 이완에 의해 억제되게 된다. 좌영상의 한 화소와 대응되는 우영상의 화소들 간의 상호 상관도 값을 가지는 노드들은 좌영상에서의 변이축을 따라 위치하며, 각 노드에서의 초기 정합가능도는 식 (3)과

로 연결된 노드들로부터 입력을 받아들여 출력 함수의 임계값에 따라 여기 상태일 경우에는 1의 값을 가지며 억제 상태일 경우에는 0의 값으로 하였다. 이와 같은 노드의 갱신은 유일성 및 연속성 제한 조건을 만족시키면서 진행되고, 간단한 반복 과정을 통해 정합이 이루어진다.

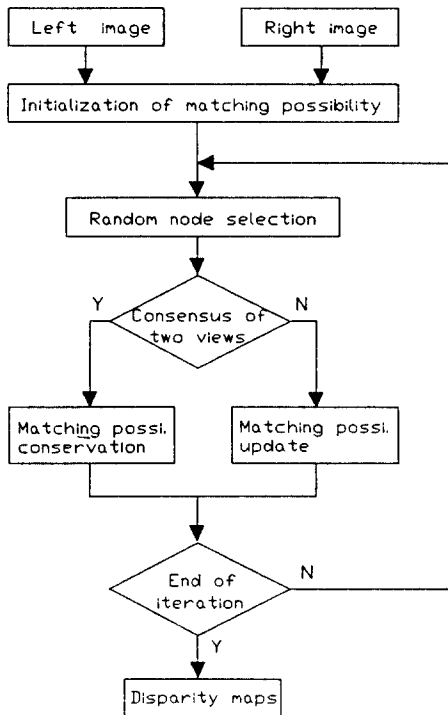


그림 3. 제안한 반복 이완 알고리즘
Fig. 3. The flow-chart of proposed iterative relaxation algorithm.

신경 회로망의 일반적인 문제 해결 방법은 회로망의 동작이 에너지 함수 값을 줄이는 방향으로 뉴런의 상태를 변화시키는 것이다. 인간의 시각 체계에 근거를 둔 Marr 등의 방법은 제한 조건이 주어진 비용 함수에 최적화 알고리즘을 적용한 신경 회로망의 해결 방법을 사용하였다. 따라서, 해결하고자 하는 문제의 에너지 분포 상태에는 여러 개의 국부 극소점(local minima)이 존재하기 때문에 최소점에 도달하기가 어려워지며, 초기 에너지의 위치에 따라 문제의 해결 결과가 달라지는 신경 회로망의 일반적인 문제점을 가진다.

제안한 방법에서는 유사도의 척도로 좌·우영상간의 정합가능도를 사용하므로 배타적 논리값에 의해 결정

되는 방법보다 국부 극소점이 적게 존재하여 에너지의 분포 상태가 단순해지고, 양방향의 정합 결과가 일치할 경우 에너지의 최소점에 도달해 있을 가능성이 크게 되어 수렴 속도가 빠르다. 또한, 정합 결과가 일치하지 않을 경우 인접 노드의 정합 결과와 협동하여 국부 극소점으로부터 빠져 나와 최소점에 도달하게 된다.

2) 정합가능도의 갱신

본 논문에서는 이완화 과정에서 변이의 연속성을 유지하기 위해 주변 노드들의 정합가능도를 고려하여 정합가능도를 갱신한다. 이때, 반복 이완 과정에서 갱신되는 정합가능도의 값이 0과 1사이의 값을 가져 정규화(normalization)과정이 필요하지 않도록 하는 것이 바람직하다.

노드 $N'_{k,i,d}$ 의 정합가능도는 그림 2(a)의 I-축 및 J-축의 대각선 방향이며 변이가 d인 면에서 $N'_{k,i,d}$ 를 중심으로 하여 3x3 또는 5x5창 내에 존재하면서 $N'_{k,i,d}$ 의 인접 주사선에 위치하는 노드들과 $N'_{k,i,d}$ 와 동일 주사선에 위치하는 노드들의 정합가능도로부터 갱신된다. 노드 $N'_{k,i,d}$ 의 정합가능도가 노드 $N'_{k,i,d}$ 를 교차하여 지나는 좌측 및 우측 영상의 변이축에 위치하는 다른 모든 노드들의 정합가능도보다 크면 정합의 신뢰성이 크므로 식(4)에서와 같이 현재의 값을 유지하고, 노드 $N'_{k,i,d}$ 의 정합가능도가 최대값이 되지 않으면 좌영상 및 우영상에서의 정합 결과가 서로 상이하여 정합의 신뢰성이 모호하게 되므로 주변 노드의 값을 고려하여 새로운 값으로 갱신하여 준다.(SRR 알고리즘, Self-Reliant Relaxation algorithm)

$$P'_{k,i,d}(t+1) = \begin{cases} P'_{k,i,d}(t), \\ \text{if } P'_{k,i,d}(t) = \max\{P'_{l,m,n}(t), (l, m, n) \in \mathcal{E}(k, i, d)\} \\ f(x+\Delta x), \\ \text{otherwise} \end{cases} \quad (4)$$

여기서, t는 갱신 시간 변수로 $t = 0, 1, 2, \dots, T$ 이며, T는 최대 반복 회수이다. 그리고, \mathcal{E} 는 노드 $N'_{k,i,d}$ 를 교차하여 지나는 좌·우영상의 두 변이축에서 최대 변이 a_{max} 내에 위치하는 노드들의 집합이다.

정합가능도를 갱신하기 위한 함수로는 주변 노드들의 정합가능도가 $N'_{k,i,d}$ 노드의 값보다 크면 주변 노드와의 변이의 연속성을 유지하기 위해 갱신되는 정합가능도를 증가시켜 주고, 0과 1사이의 값을 가지며, 반복 이완 과정을 거쳐 이 두 값으로 수렴하도록

$$f(x) = \frac{1}{1 + \exp[-\lambda(x-0.5)]} \cdot \lambda = 4 \quad (5)$$

와 같이 sigmoid 함수를 사용한다. 여기서 식 (4) 및 식 (5)에서 주어지는 x는 정합가능도 $P'_{k,i,d}(t)$ 의

sigmoid 역변환 값이다. 식(5)의 λ 값에 따라 정합가능도의 갱신 분포가 결정되고, 너무 큰 값을 가지게 되면 정합가능도가 이치화 되므로 적당한 값을 선택하여야 한다. 본 논문에서는 정합가능도가 0.5일 때 sigmoid 함수의 기울기가 1이 되도록 λ 의 값을 결정하였다. 식 (4) 및 식 (5)로부터 x 는

$$x = 0.5 - 0.25 \ln \left[\frac{1 - P'_{k,i,d}(t)}{P'_{k,i,d}(t)} \right] \quad (6)$$

과 같이 주어지며, 정합가능도의 갱신을 위한 sigmoid 역변환 값의 변화분은 식(7) 및 식(8)에서와 같이 노드 $N'_{k,i,d}$ 의 정합가능도와 인접하는 노드들의 정합가능도 평균값과의 차이로 둔다.

$$\Delta x = \bar{P} - P'_{k,i,d}(t) \quad (7)$$

$$\bar{P} = \frac{1}{M} \sum_m \sum_n P'_{k,i,d}(t), N_{k,i,d} \in \Theta N_{k,i,d} \quad (8)$$

여기서, Θ 는 노드 $N'_{k,i,d}$ 의 인접 노드들의 집합이며, M 은 인접 노드의 갯수이다. 식(7)의 sigmoid 역변환 값의 변화분에 의해 시간 t 에서의 정합가능도가 0.5 부근에서는 갱신되는 변화량이 커지게 되고 0 또는 1 부근에서는 거의 수렴 상태이므로 변화량이 미소하게 된다. 또한, 노드 $N'_{k,i,d}$ 에서의 정합가능도가 주위의 값과 비교하여 많은 차이가 나는 경우 갱신되는 변화량이 크게 되어 수렴의 속도를 증가시킨다.

3) 협동-경쟁 이완(Cooperative-Competitive Relaxation, CCR) 알고리즘

부작위로 선택된 노드 $N'_{k,i,d}$ 의 정합가능도와 인접 노드들의 정합가능도를 평균하여 $N'_{k,i,d}$ 를 교차하여 지나는 좌·우영상의 두 변이축 상에 위치하는 노드들과 비교하는 SRR 알고리즘은 두 변이축의 정합 결과가 일치하더라도 영상에 존재하는 잡음 등으로 인해 잘못된 정합이 이루어질 수 있으며, 최종적으로 얻어지는 변이 영상에서 고립점(isolated point)으로 존재하게 된다. 이와 같은 고립점들은 양방향 정합의 일치성을 확인하는 과정에서 탐색 영역을 확장하여 새로운 변이축에서의 정합가능도를 고려하는 협동-경쟁 이완의 방법을 사용하여 제거한다.

노드 $N'_{k,i,d}$ 의 정합가능도를 노드 $N'_{k,i,d}$ 를 교차하여 지나는 두 변이축 상의 노드들의 정합가능도와 비교할 때, 각 노드에서의 정합가능도는 인접 노드와의 평균값을 사용한다. 이는, 정합 여부를 결정할 때 인접 노드의 정합 가능성으로부터 지원을 받는 협동의 형태로 해석할 수 있으며, 잡음의 영향으로 정합가능도가 크게 되어 오정합을 일으키는 노드의 정합 가능성을 배제하

게 된다.

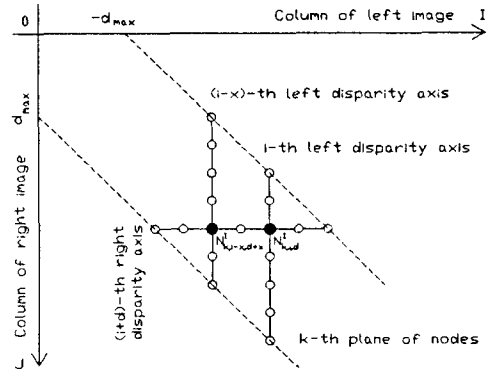


그림 4. 두 영상의 변이축 탐색 영역
Fig. 4. The searching space on disparity axes of two views.

이와 같이 인접 노드의 평균값을 고려한 노드 $N'_{k,i,d}$ 의 정합가능도가 $N'_{k,i,d}$ 를 교차하여 지나는 두 변이축 상에 위치하는 노드들 중에서 최대인 경우, 그림 4에서와 같이 두 번째로 큰 정합가능도를 가진 노드 $N'_{k,i,x,d+x}$ 가 속하는 좌영상의 $(i-x)$ 번째 변이축을 따라 위치하는 노드들의 정합가능도를 비교한다. 좌영상의 $(i-x)$ 번째 변이축을 따라 위치하는 노드들 중에서 노드 $N'_{k,i,x,d+x}$ 의 정합가능도가 최대인 경우에는 노드 $N'_{k,i,x,d+x}$ 에서의 정합 가능성이 크게 되어 노드 $N'_{k,i,d}$ 에서의 정합 신뢰성이 확실하지 않으므로 현재 선택된 노드 $N'_{k,i,d}$ 의 정합가능도를 식 (4)에 의하여 갱신한다. 이는 좌영상의 i 번째 화소와 $(i-x)$ 번째 화소가 우영상의 $(i+d)$ 번째 화소와의 정합 가능성을 서로 경쟁하는 것으로 해석할 수 있다. 반복 과정을 거치면서 우영상의 $(i+d)$ 번째 화소와의 정합가능도를 가진 노드 $N'_{k,i,x,d+x} (|d+x| \leq d_{max})$ 들은 우영상의 $(i+d)$ 번째 화소와 정합을 위해서 서로 경쟁을 하게 되고, 정합 신뢰성이 확실하지 않은 경우 자신의 정합가능도를 갱신한다. 노드 $N'_{k,i,x,d+x}$ 의 정합가능도가 최대값이 아닌 경우에는 노드 $N'_{k,i,d}$ 의 정합 가능성이 상대적으로 더 커지므로 현재 선택된 노드 $N'_{k,i,d}$ 의 정합가능도를 갱신하지 않는다.

현재 선택된 노드 $N'_{k,i,d}$ 의 정합가능도 $P'_{k,i,d}$ 가 노드 $N'_{k,i,d}$ 를 교차하여 지나는 두 변이축 상에 위치하는 노드들 중에서 최대가 되지 않을 경우에는 현재 선

택된 노드 $N_{k,i,d}$ 에 대한 정합의 신뢰성이 낮으므로 식 (4)에 의해 갱신한다.

이상에서와 같이 새로운 변이축을 고려하여 탐색 영역을 확장한 CCR 알고리즘은 인접 노드의 정합가능도의 평균 정합가능도를 구하는 협동 과정에서 잡음으로 인한 오정합의 가능성을 억제하며, 정합의 신뢰성이 확실하지 않는 양방향 정합 일치성을 가지는 노드의 정합가능도를 갱신하여 고립점으로 존재하는 오류를 제거한다. 제안하는 양안의 협동-경쟁에 의한 대응점 결정 방법은 공간적으로 좁은 범위에서 협동하는 과정과 한 변이축을 따라 공간적으로 떨어져 위치하는 각 노드들이 경쟁하는 과정으로 이루어져 있으며, 이 두 가지의 반복적 협동-경쟁에 의한 이완 과정으로 목적하는 대응점을 찾게 되며 잡음에 의한 오정합 결과를 줄여 주는 장점이 있다.

4) 변이값 결정

반복 이완 과정에 의해 미리 설정된 최대 반복 횟수에 이르면 이완 과정의 진행은 끝나게 되고, 좌·우영상의 각 열의 위치에서 정합 변이축을 따라 최대 정합가능도를 가지는 노드가 위치하는 d 를 변이값으로 결정하며

$$d_L(k,i) = \underset{d}{\operatorname{argmax}} [P_{k,i,d}^L(T+1), (k,i,d) \in E_1]$$

$$d_R(k,j) = \underset{d}{\operatorname{argmax}} [P_{k,j,d}^R(T+1), (k,j,d) \in E_2] \quad (9)$$

와 같이 주어진다. 여기에서, E_1 는 좌영상의 i 번째 정합 변이축에 위치하는 노드들의 위치이며, E_2 는 우영상의 j 번째 정합 변이축에 위치하는 노드들의 위치이다.

식(9)로부터 얻은 양안의 두 변이 영상은 양안의 정합 일치성을 고려한 정합가능도의 전파효과로 물체와 배경 부분에서는 신뢰성 높은 정합 결과를 보이나, 변이의 연속성 제한 조건에 의해 불연속점 부근에서 정합가능도의 전파 효과가 대립되는 현상에 의해 양안의 정합 결과가 달라지는 오류를 발생시키게 된다. 이는 화소가 물체의 경계 부분에서 가까이 위치할수록 좌·우영상의 상관관계가 적어질 가능성이 커지게 되며, 멀리 떨어질 경우에는 정합가능도가 높은 신뢰성 있는 정합 결과를 가지기 때문이다. 이러한 연속성 제한 조건에 의한 불연속점 부근에서의 심각한 오류의 발생에 대해서는 Grimson^[14]에 의해 확률적인 접근 방법으로 잘 해석되어 있다. 물체의 경계선 부근에서의 오류를 해결하기 위하여 Mayhew와 Frisby^[15]는 경계선의 연속성 제한 조건을 사용하였으며, 기존의 폐색 영

역(occlusion region) 탐지 방법^[16-18]으로는 물체 경계 부분의 연결성, 명도의 윤곽, texture의 갑작스런 변화 등 입력 영상의 정보를 분석하는 방법이 있다. 제안한 방법은 고수준(high-level) 처리 과정을 거치지 않고 반복 이완 과정을 통해 경계선으로부터 멀리 떨어진 화소의 신뢰성 높은 정합 결과가 불연속점에 위치하는 화소로 전파되어 경계 영역에서의 오류를 줄이게 된다.

양안에서 얻은 두 변이 영상의 결과를 서로 비교하여 변이 결과가 일치하지 않는 경우 물체의 불연속점 및 폐색 영역으로 결정하여 차후 물체의 인식 단계에서 유용한 정보로 사용한다. 왼쪽 변이 영상에 있어서 물체의 경계로 인한 불연속점이나 폐색 영역은

$$d_L(k,i) \neq d_R(k,i+d_L(k,i)) \quad (10)$$

에 의해 결정된다. 여기에서, 좌·우영상의 정합 결과로부터 얻은 물체의 경계선으로 인한 불연속점들은 양안의 정합가능도를 동시에 고려한 정합가능도의 전파 효과로 경계선을 따라 가늘게 나타난다. 오른쪽 변이 영상에 있어서 물체의 경계로 인한 불연속점이나 폐색 영역은 식 (10)과 마찬가지로

$$d_R(k,j) \neq d_L(k,j+d_R(k,j)) \quad (11)$$

과 같이 주어진다.

III. 실험 및 고찰

본 논문에서는 10%의 잡음이 포함된 50% 랜덤 도트 스테레오 영상과 항공 영상인 펜타곤(Pentagon) 영상에 대하여 제안한 알고리즘의 타당성과 성능을 실험하였다. 잡음이 섞인 랜덤 도트 영상에 대해 Marr^[8] 등의 반복 협동 정합 알고리즘과 제안한 정합 알고리즘을 적용하여 변이 정보 추출 결과를 비교하였다.

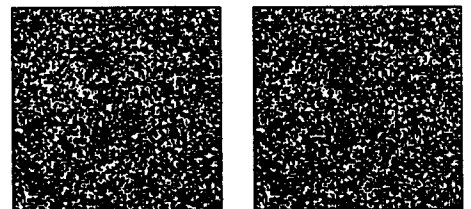


그림 5. 50% 랜덤 도트에 10%의 잡음이 섞인 스테레오 영상

Fig. 5. A 50% random dot stereogram with 10% noise.

제안한 방법에 있어서 주어진 두 영상의 정합가능도를 구하기 위한 NCC의 창틀 크기를 5x5로 하여 유사

도를 구하였으며, 최대 변이 d_{max} 를 3으로 하여 결과적으로 얻어지는 변이는 -3에서 3사이의 값을 가지게 된다. 그림 5는 이치화 되어 있는 128x128 크기의 랜덤 도트 스테레오 영상으로 가장자리의 변이는 0이며 중앙은 3의 변이값을 가지는 4개의 계층을 가진다.

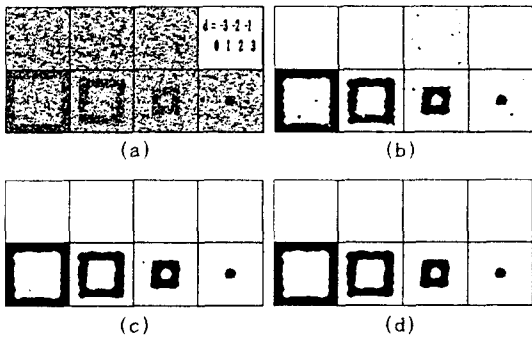


그림 6. 반복 횟수에 따른 각 변이층의 활성화 상태
(a) 초기 상태 (b) 5회 반복
(c) 10회 반복 (d) 20회 반복

Fig. 6. The activation state of each disparity layers with iterations.
(a) Initial state.
(b) After 5 iterations.
(c) After 10 iterations.
(d) After 20 iterations.

그림 6은 Marr 등이 제안한 알고리즘을 적용하여 각 변이층 뉴런들의 활성화 상태를 매 반복 횟수에 따라 나타낸 결과이다.

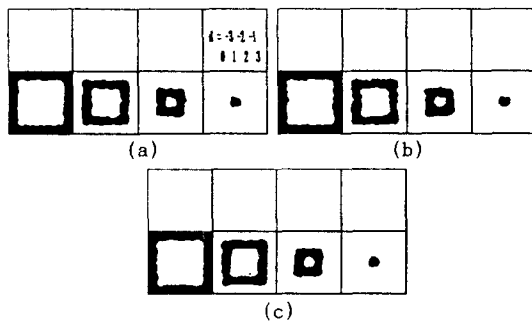


그림 7. 랜덤 수 발생기의 seed값에 따른 각 변이층의 활성화 상태
(a) 랜덤 seed = 1 (b) 랜덤 seed = 3
(c) 랜덤 seed = 5

Fig. 7. The activation state of each disparity layers with different values of random seed.
(a) random seed = 1.
(b) random seed = 3.
(c) random seed = 5.

그림 6의 변이 영상에서 볼 수 있듯이 각 변이층에 대한 뉴런의 활성화 상태가 반복 횟수가 증가하면서 평활성과 유일성의 제한 조건에 따라 점차적으로 수렴하게 되나, 변이가 변하는 불연속점 부근에서 20회의 반복 과정을 거쳐도 한 화소에 대응되는 변이가 하나 이상인 중복 정합이 일어난다. 그림 7은 반복 이완시 무작위로 선택되는 노드의 순서를 결정하기 위한 랜덤 수 발생기의 seed 값에 따라 랜덤 도트 영상을 20회 반복 이완한 정합 결과들로 불연속점 부근에서 변이의 연속성 제한 조건에 의해 주변 화소의 영향을 받아 몽롱화 현상을 동반하며, 랜덤 수 발생기의 seed 값을 변경하였을 때 회로망의 동작이 초기 에너지 상태가 다른 지점에서 시작하게 되어 활성화되는 뉴런의 상태가 서로 달라지는 시스템의 불안정성을 확인하였다.

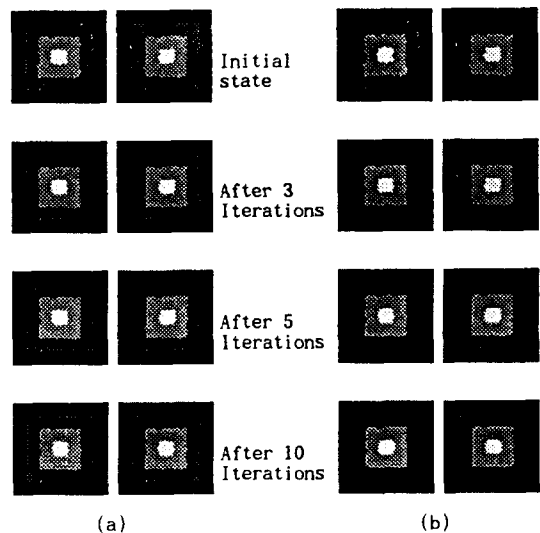


그림 8. 제안한 알고리즘을 적용한 좌·우 변이 영상

(a) SRR 알고리즘 (b) CCR 알고리즘

Fig. 8. The left and right disparity map of the proposed algorithms.

(a) Srr algorithm. (b) CCR algorithm.

그림 8은 제안한 SRR 및 CCR 알고리즘을 적용하여 매 반복 횟수에 따른 정합 결과를 디더링(dithering)하여 나타낸 영상이다. 제안한 알고리즘은 두 영상의 변이축이 교차하는 노드에서의 상호 상관도가 최대값을 가지는 경우 그 값을 보존하며, 잡음 등에 의해 최대 상관도 값을 가지는 노드는 주위 노드로부터 억제력을 받아 반복 이완을 진행하는 과정에서 오정합의 가능성이 배제된다. 결과 변이 영상에서 볼 수 있듯이 5회 이내의 반복으로 수렴된 결과를 얻을 수 있으며, 변이가 변하는 불연속점 부근에서의 몽롱화 현상이 제거되었다.

SRR 알고리즘에서는 잡음이나 물체의 불연속점으로 인해 잘못된 정합 결과가 이완 과정을 반복하여도 계속 고립점 형태로 존재하는 경우가 발생하며, 양안의 협동-경쟁에 의한 CCR 알고리즘은 오정합이 일어난 노드들의 정합가능도의 억제 효과로 반복 이완을 진행 하면서 고립점 형태의 오정합을 제거하였다. SRR 및 CCR 알고리즘에 대해 랜덤 수 발생기의 seed 값을 달리하여 얻은 변이 영상에서 불연속점에서의 수렴된 변이 형태가 그림 8과 같이 거의 동일하여 시스템이 안정되게 동작함을 확인하였다. Marr 등의 알고리즘과 성능을 객관적으로 평가하기 위해 좌영상에 대한 실제 변이 $d_{k,i}$ 와 알고리즘을 적용 후 얻은 결과 변이 영상의 차이값을 식(12)와 같이 정의할 때, 반복 횟수에 따른 변이의 평균 오차와 표준 편차는 표 I과 같이 주어진다.

$$\Delta d = d_{k,i} - d_L(k, i) \tag{12}$$

그림 9는 양안에서 얻은 변이 영상으로부터 식(10) 및 식(11)에 의해 얻은 양안에서의 불연속점 및 폐색 영역으로, 반복 이완 과정을 통해 불연속점으로부터 떨어진 신뢰성 높은 정합 결과가 양안의 정합가능도를 동시에 고려하면서 불연속점으로 전파되기 때문에 경계선을 따라 가늘게 나타난다.

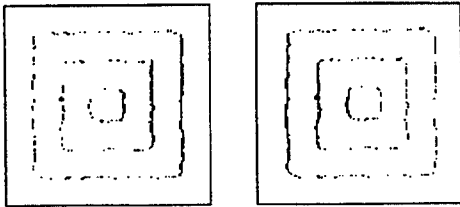


그림 9. 양안의 폐색 영역
Fig. 9. The occlusion region of two views.

그림 10은 실외에서 촬영된 512x512 크기의 스테레오 영상으로부터 영상에 존재하는 잡음을 억제하고 표본화의 오차를 줄이기 위하여 가우시안 저역 통과 필터를 적용한 후 256x256 영상으로 표본화한 영상으로, 최대 변이 d_{max} 를 5로 제한하였다.

그림 11은 제안한 알고리즘으로 10회의 반복 처리를 통해 얻은 양안에서의 변이 영상이며, 영상 우측의 위쪽에 위치하는 다리뿐만 아니라 건물이 존재하는 영역의 세밀한 부분에서도 정확한 변이 정보를 찾고 있다. 특히, 건물 좌측의 웅달진 영역에 존재하는 폐색 영역에서 두 변이 영상의 정합 결과가 상이하게 나타나 양안의 변이 영상으로부터 중복 정합을 제거하고

신뢰성 높은 폐색 영역의 결정이 가능하다.

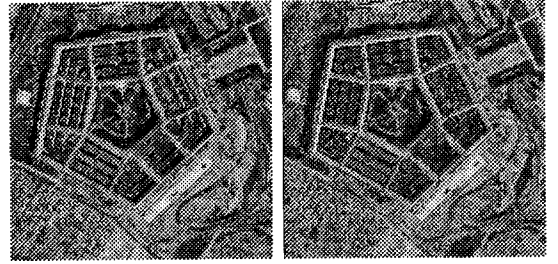
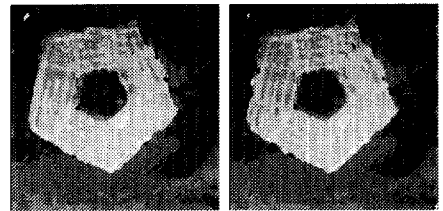
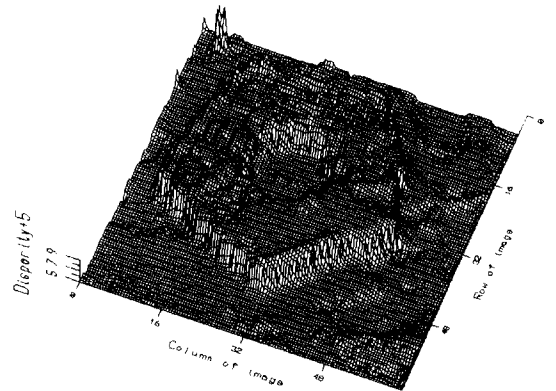


그림 10. 펜타곤 스테레오 영상
Fig. 10. "Pentagon" stereo image.



(a)



(b)

그림 11. 펜타곤 스테레오 영상에 대한 양안의 변이 영상

- (a) 디더링된 좌·우 변이영상
- (b) 좌측 변이 영상의 3차원 plot

Fig. 11. The disparity computation results for the "Pentagon" stereo image.

- (a) The dithered left and right disparity images.
- (b) The 3-D plot of left disparity image.

IV. 결 론

본 연구에서는 intra-주사선 제한 조건으로부터 대

응점의 탐색 범위를 줄이고, inter-주사선 제한 조건 및 인접하는 화소 사이의 변이의 연속성을 이용하며, 좌·우영상의 각 변이축 상에 위치하는 노드들에 대한 정합가능도의 일치성을 고려한 반복적 이완 정합 방법 제안한다. 제안하는 두 변이축 상에서 정합가능도의 일치성을 고려한 알고리즘은 두 영상의 변이축이 교차하는 노드에서의 상호 상관도가 최대값을 가지는 경우 그 값을 보존하며, 잡음 등에 의해 최대 상관도 값을 가지는 노드는 주위 노드로부터 억제를 받아 반복 이완을 진행하는 과정에서 오정합의 가능성이 배제되어 신뢰성 높은 정합 결과를 가져온다.

표 1. 랜덤 seed값에 따른 각 알고리즘의 변이 오차 분포

Table 1. Disparity error distribution of each algorithm for various random seed value.

seed No. / average & standard deviation	1		3		5		
	$\overline{\Delta d}$	$\sigma_{\Delta d}$	$\overline{\Delta d}$	$\sigma_{\Delta d}$	$\overline{\Delta d}$	$\sigma_{\Delta d}$	
Marr	0	0.791	1.999	0.791	1.999	0.791	1.999
	5	0.118	0.501	0.111	0.482	0.114	0.474
	10	0.102	0.365	0.099	0.367	0.101	0.357
	20	0.100	0.354	0.097	0.364	0.103	0.354
Proposed	0	0.105	0.393	0.105	0.393	0.105	0.393
	5	0.092	0.321	0.091	0.320	0.092	0.319
	10	0.092	0.318	0.092	0.318	0.093	0.320
	20	0.093	0.314	0.093	0.314	0.092	0.314

뿐만 아니라, 변이의 연속성 제한 조건으로 인한 물체의 경계 부분인 불연속점 부근에서 심각하게 발생하는 몽롱화 현상이 억제되는 장점이 있다. 반복 이완을 마친 후 각 변이축을 따라 최대 정합가능도를 가지는 노드의 위치를 변이값으로 선택하기 때문에 폐색 영역에서도 다른 영상의 한 화소와 대응되어 오정합의 결과를 가져오나 양안에서 동시에 얻은 변이 영상으로부터 폐색 영역의 결정이 가능하므로 후처리 과정인 물체 인식 과정에 유용한 정보로 사용이 가능하다. 제안된 방법은 병렬 처리가 가능한 구조로 많은 정보량을 빠른 시간에 처리할 수가 있으나 영상의 변이값이 큰 경우에는 노드의 수와 탐색 영역이 대폭 증가하게 되어 시스템의 성능이 떨어진다.

이러한 단점을 보완하기 위해서는 최근에 많은 연구

가 진행되고 있는 계층 구조^[10,19,20]를 이용하여 상위 계층의 초기 변이 정보를 하위층에 전파하여 탐색 영역 및 오정합의 가능성을 줄이는 연구가 계속 진행되어야 하겠다.

참 고 문 헌

- [1] Yoshiaki Shirai, Three-Dimensional Computer Vision, Springer-Verlag, pp.1, 1987
- [2] M. Bertero, T.A. Poggio, and V. Torre, "Ill-posed problems in early Vision", IEEE proc., vol.76, no.8, pp.869-889, August 1988.
- [3] T. Poggio and V. Torre, "Ill-posed problems and regularized solution to edge detection", AI memo 773, AI Lab. MIT, April 1984.
- [4] J. Zhang and J. W. Modextino, "A Markov Random Field Model-Based Approach to Image Interpretation," SPIE Visual Communication and Image Processing IV, Vol.1199, pp.328-339, 1989.
- [5] S.Geman and D. Geman, "Stochastic relaxation, Gibbs Distribution, and the Bayesian Restoration of images," IEEE Trans. on PAMI, vol.6, pp.721-741, 1984.
- [6] Y. Ohta and T. Kanade, "Stereo by Intra and Inter-Scanline Search Using Dynamic Programming," IEEE Trans. on PAMI, vol.7, no.2, pp.139-154, 1985.
- [7] Y. H. Zhou and R. Cellappa, "Stereo matching using a neural network," in Proc. Intl. Conf. on ASSP, New York, 1988.
- [8] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," Science, vol. 194, pp.283-287, 1976.
- [9] D. Marr and T. Poggio, "A Computational theory of human stereo vision," Proc. Royal Society of London, B.204, pp.301-328, 1979.
- [10] S. D. Cochran and G. Medioni, "3D su-

- rface description from binocular stereo." IEEE Trans. on PAMI, vol.14, no. 10, pp. 981-994, Oct. 1992.
- [11] A. Khotanzad, A. Bokil, and Y. W. Lee, "Stereoopsis by Constraint Learning Feed-Foward Neural Networks," IEEE Trans. on Neural Networks, vol. 4, no.2, pp.332-342, Mar. 1993.
- [12] R. L. Henderson, W. J. Miller, and C. B. Grosch, "Automatic stereo reconstruction of man-made targets," SPIE, vol.186, pp. 240-248, 1979.
- [13] H. H. Baker, "Depth from edge and intensity based stereo," Stanford Artif. Intell. Lab., Stanford Univ., Stanford, CA, Tech. Rep. AIM-347, 1982.
- [14] W. E. L. Grimson, "Computational experiments with a feature-based stereo algorithm," IEEE Trans. on PAMI, vol.7, no.1, pp.17-34, Jan. 1985.
- [15] J. E. W. Mayhew and J. P. Frisby, "Psychophysical and computational studies towards a theory of human stereoopsis," Artif. Intell., vol.17, pp. 349-385, 1981.
- [16] Koenderink J.J and Van Doom D., "The shape of smooth objects and the way contours end," Perception 11, pp. 129-137, 1982.
- [17] Horn B. K. P., "Understanding Image Intensities," Artif. Intell. 8, pp.201-231, 1977.
- [18] Riley M.D., "The representation of image texture," MIT AI-Lab., AI-TR-649, Sep. 1981.
- [19] D. Terzopoulous, "Multilevel computational processes for visual surface reconstruction," Comput. Vision Graphics Image Processing, vol.24, pp.52-63, 19 83.
- [20] J. Weng, N. Ahuja and T. S. Huang, "Matching two perspective views," IEEE Trans. on PAMI, vol.14, no.10, pp. 981- 994, Oct. 1992.

 저 자 소 개

李 旺 國(正會員) 第 30卷 B 編 第 6號 參照
 현재 국방과학연구소 선임연구원

金 容 奭(正會員) 第 31卷 B 編 第 8號 參照
 현재 경북 대학교 전자공학과 박사과정

都 敬 勳(正會員) 第 30卷 B 編 第 6號 參照
 현재 경북 대학교 전자공학과 박사과정 수료

河 永 浩(正會員) 第 31卷 B 編 第10號 參照
 현재 경북 대학교 전자공학과 부교수