

이동물체의 정확한 추출을 위한 스테레오 알고리즘

김종만

남도대학 컴퓨터응용전기과

Stereo vision Techniques for Correct extract of Moving object

Jong-Man Kim

Namdo Provincial College

Abstract - The proposed neural network technique is the real time computation method based theory of inter-node diffusion for searching the safety distances from the sudden appearance-objects during the work driving. The main steps of the distance computation using the theory of stereo vision like the eyes of man is following steps. One is the processing for finding the corresponding points of stereo images and the other is the interpolation processing of full image data from nonlinear image data of objects. All of them request much memory space and time. Therefore the most reliable neural-network algorithm is driven for real-time matching of objects, which is composed of a dynamic programming algorithm based on sequence matching techniques in moving objects.

1. 서 론

보간 기법을 이용한 경우로서 저해상도를 가지는 영상 시스템에서 고해상도의 원래의 영상 정보로 보간시키는 방법이 수행되었다[1][2]

병렬처리의 원리를 이용하는 신경회로망은 내부의 연결 강도를 적응적으로 변화시키는 학습성, 병렬성 등의 특성에 의하여 신속하게 계산처리를 하는 보간기능을 가진다. 이러한 특성으로 인하여 기존의 몇몇 연구자들이 신경회로망의 보간 특성을 이용하여 패턴인식에 응용 및 함수의 근사화에 적용하였다. [2-4]

본 연구에서는 이동하는 물체에서 정확한 거리추정을 위하여 스테레오 알고리즘을 제시하였다. 기존의 근사화 특성의 불확실성이 존재하는 부분들의 단점을 보완코자 실시간 전파를 목적으로 정보를 보간시키는 알고리즘을 제안하였다. 현재 실제 지형정보를 기초로 하여 얻어진 임의의 위치에 대한 일치성 문제인 동적계획법의 해결 방법을 가능케 해주고, 얻어진 정보의 표본 지형 정보를 통해 특징점 중심의 거리정보로부터 전체 거리정보로 실시간 보간 기능을 실시하므로 실시간 동작처리를 가능케 해준다.

2. 스테레오 계측 기법

2.1 스테레오 카메라의 화상 계측 정보 시스템

스테레오 비전은 두 대의 카메라에 맺히는 물체 화상의 위치에 대한 이격값을 거리로 계산하는 원리이다. 정확한 화상정보 계측을 위하여 먼저 카메라 캘리브레이션을 수행한다. 3차원 공간과 2차원 카메라 화상 평면과의 대응관계에서 나타나는 카메라 내부의 기하학적, 광학적 특성 인수를 결정하는 내부 파라미터 교정과 기준 좌표계에 대한 카메라 좌표계의 위치와 방향 인수인 외부 파라미터를 결정한다. 카메라 캘리브레이션 과정이 완료되면 두대의 스테레오 카메라를 사용하여 화상을 획득한다. 스테레오 방식으로 깊이 정보를 추출하는데 중요한 문제는 3차원 실좌표계의 한점에 대한 왼쪽의 화상과 오른쪽 화상의 같은 위치점을 찾아내는 정합문제이다. 좌

우 카메라에 얻어진 화상들의 일치점 파악을 위하여 특징점 처리(Feature Extraction) 과정을 수행한다. 이 특징점 정보들에 대하여 신경회로망 등의 화상신호처리 알고리즘을 사용하여 정합과정을 수행한다. 정합이 완료된 2차원 특징점 화상 정보는 작은 용량의 정보로 목표표하는 원격지 등에 전송된후, 원격지의 중앙센터에서 최적의 지능 알고리즘 등을 이용하여 3차원 전체 영역의 거리 정보로 보간시켜 원래의 전체 화상을 재생 복원시켜 사용한다.

그림 1에 스테레오 카메라를 이용하여 화상신호를 습득하여 코딩 처리후 원거리 목적지까지 송신후 다시 재생시키는 기본 절차를 한 예로 하여 스테레오 비전 처리하는 기하학적인 구조를 보였다. 카메라의 일반적 투사 과정은 임의의 3차원 공간상의 점을 평면에 투사시키는 투사 변환법을 이용하는 편환 모델을 이용한다. 이 편환 모델을 이용한 두대의 카메라를 사용하는 스테레오 카메라의 화상처리 과정이 그림과 같이 일정 거리의 3차원 공간상의 점의 거리 정보가 스테레오 카메라를 통하여 투사된 두개의 화상 평면으로 나타나는 원리를 이용한 다.

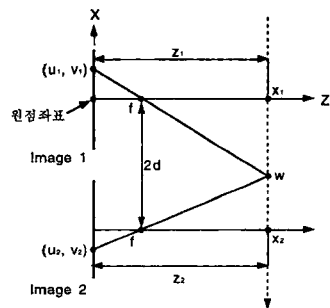


그림 1 스테레오 비전의 기하학적 구조

2.2 거리 추출 관계식.

여기서 f(focus)는 카메라의 초점거리이고, z(distance 혹은 depth)는 물체의 위치에서 화상면 까지의 거리를 의미하며, 2d'는 epipolar선 즉, 두카메라의 중심과 중심을 연결한 선분을 의미한다. 물체의 좌표를 P1(x1, y1, z1), 좌측 화상에 대한 투사되는 좌표를 (u1, v1), 물체의 우측 화상에 대한 투사되는 좌표를 (u2, v2)라 하면 앞의 기하학적 구조는 다음과 같은 관계식으로 표현된다.

$$x_1 = \frac{(f-z_1)u_1}{f} \quad (1)$$

$$x_2 = \frac{(f-z_2)u_2}{f} \quad (2)$$

이 식과 $x_2 = x_1 + 2d$ 및 $z_1 = z_2 = z$ 라는 조건을 이용하면 다음과 같은 관계식을 쉽게 얻을 수 있다.[1]

$$z = f - \frac{2df}{u_2 - u_1} \quad (3)$$

이 식은 초점거리와 두 화상좌표의 x 좌표의 화소 차이 $u_2 - u_1$ 만 알면 거리 z 를 산출해낼 수 있음을 의미한다.

3. 실시간 보간을 위한 정보전파 신경회로망

3.1 보간특성을 갖는 정보전파 신경회로망

본 연구에서 제안한 정보전파 신경회로망(LIPN)은 그림 2와 같이 그 구조가 단층 신경망 노드로 구성되어 있다. 각 노드가 자신의 입력값 뿐 아니라 인접한 노드의 출력값에 점점 수렴하도록 하는 연결구조를 가지며, 임의의 위치 정보를 주변의 상태 공간에 신속하게 전파시켜 자연스런 정보의 보간화가 이루어진다. 그림 3은 제안한 정보전파 신경회로망의 1차원 구조이다.

$$P_i = f [c_i I_i + \sum_l c_{i,l+1} P_{i+1,l}]; \quad l = \pm 1 \quad (4)$$

여기서 f 는 신경망 노드의 활성화 함수이다.

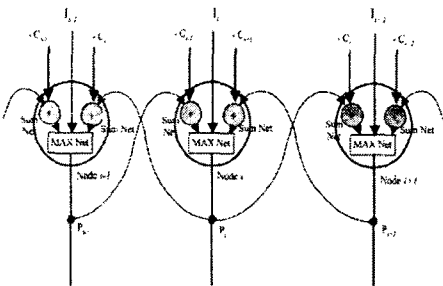


그림 2 제안한 신경회로망의 1차원 연결과 노드의 구조

위 신경회로망 구조는 Sum과 MAX 등의 연산을 수행하는 부 신경망(sub networks)들의 합성구조(compound structure)를 가지며 입력된 cost에 의해 최소의 비용이 소요되는 방향을 찾는 역할을 한다. 각 노드에는 자신의 입력 I가 인가되며 인접노드의 출력으로 부터 노드간의 cost가 감해진 값들도 함께 인가된다. 각 노드의 출력으로는 인가된 값 들 중 최대의 값이 출력된다.

신경회로망의 각 노드는 양자화된 입력공간상의 상태와 일치한다. 노드의 입력 신호와 인접노드의 신호정보는 평균 가중되어 출력을 산출시킨다.

산업현장의 물류시스템 등 작업환경하의 돌발 물체의 영상 데이터를 잡아 신경망 기법을 통해 실시간으로 인식하기 위해서는 비선형 영상정보 부분에 대하여 선형적인 보간기능 수행을 통해 출력을 생성시키며, 이때 연결강도도 새롭게 생성시키는 기능을 가져야 한다. 본 연구에서 제안한 정보전파 신경회로망의 연결강도는 보간을 위해서 다음과 같이 최적으로 정해지는 출력 특성을 갖는다.

[A] 입력이 있는 노드는 입력과 같은 연결강도의 값을 출력시킨다.

[B] 입력이 없는 노드는 인접 노드로부터의 연결강도의 평균된 값을 출력시킨다.

3.2 제안한 정보전파 신경회로망을 이용한 동적 계획법 구현

동적계획법은 주어진 제한 조건하에서 시작점으로부터 목적지까지 이르는 최적의 경로를 구하는 효과적인 방법이다. 그림 3과 같은 2차원 격자형 경로의 경우 시작점

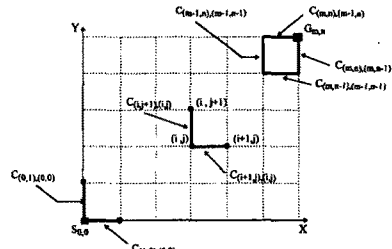


그림 3 동적계획법을 위한 2차원 격자 경로

$S(0,0)$ 에서 목적지 $G(m,n)$ 지점에 이르는 최대 이익(benefit)의 경로는 $(0,0)$ 에서 $(1,0)$ 점을 지나는 경로와 $(0,1)$ 점을 지나는 경로 중 이익(benefit)이 더 큰 경로이다. 만약 $(1,0)$ 과 $(0,1)$ 로 부터 목적지에 이르는 최대 이익 $C_{1,0}^*$ 과 $C_{0,1}^*$ 을 알고 있다면, $(0,0)$ 점으로부터 $(1,0)$ 을 거쳐 목적지에 이르는 이익은 $C_{(1,0),(0,0)} + C_{1,0}^*$ 으로 표기되며, $(0,1)$ 을 거쳐서 목적지에 이르는 이익은 $C_{(0,1),(0,0)} + C_{0,1}^*$ 가 되며, $(0,0)$ 점으로부터 목적지에 이르는 최대 이익 $C_{0,0}^*$ 은

$$C_{0,0}^* = \text{MAX} \{ C_{(1,0)(0,0)} + C_{1,0}^*, C_{(0,1)(0,0)} + C_{0,1}^* \} \quad (5)$$

이다.

스테레오 영상에 동적계획법을 적용시키기 위해서는 좌, 우 영상의 epipolar 선을 각각 X, Y 축으로 하며 각 격자점 위치에 해당 화소들 간의 상호 상관값을 각 노드의 이익값으로 갖는 정합공간을 구성한다. 따라서 그림 2에서 구하고자 하는 (i,j) 번째 노드를 지나는 경로의 최대 이익 $C_{i,j}^*$ 는

$$C_{i,j}^* = \text{MAX} \{ C_{(i+1,j),(i,j)} + C_{i+1,j}^*, C_{(i,j+1),(i,j)} + C_{i,j+1}^* \} \quad (6)$$

이다.

$C_{(i+1,j),(i,j)}$ 와 $C_{i+1,j}^*$ 사이에 합(+) 유닛을 설치하고 $C_{(i,j+1),(i,j)}$ 와 $C_{i,j+1}^*$ 사이에 합(+) 유닛을 설치한다. 합해진 두 신호는 MAX 유닛에 인가시키고 그 중 최대치를 취하여 $C_{i,j}^*$ 로 하는 신경회로망을 구성하면 동적계획법이 구현된다. $C_{(i+1,j),(i,j)}$ 와 $C_{(i,j+1),(i,j)}$ 는 $(i+1,j)$ 방향과 $(i,j+1)$ 방향에 대한 bias로 간주하면 쉽게 구현될 수 있으며, MAX net는 다이오우드 회로를 이용하면 간단히 구성된다.

4장. 실험 및 검토

4.1 실시간 측정을 위한 보간모델의 실험

산업현장의 돌발물체 출현시 얻어진 비선형 영상정보에서 연결강도의 보간 특성을 입증하기 위해 다음 그림 5에서는 신경망 노드의 0번째 노드와 8번째 노드에 각각 30과 10의 값을 인가하여 연결강도를 각기 달리하여 보간시킨 특성 실험 값을 보였다.

그림 4에서 보인 것 처럼 제안한 정보전파 신경회로망은 연결강도 값이 $w_1=0.5$ 와 $w_2=0.5$ (3)의 경우에 선형보간을 수행함을 알 수 있으며, 또한 다른 여러 경우의 연결강도 설정시에 비선형 보간을 수행할 수 있음을 보여주고 있다.

제안한 신경회로망의 보간 특성을 확인하기 위해 수행시에

얻어지는 스테레오 영상을 그림 5와 같이 실험실 내의 영상용 모델로 대체하여 특정 실험을 수행하였다.

5. 결 론

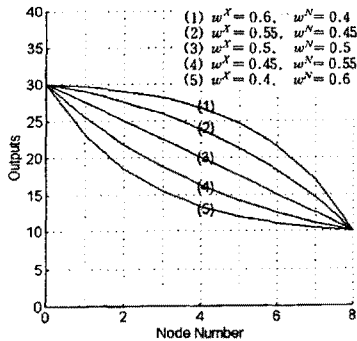


그림 4. 다양한 연결강도를 갖는 LFN의 보간 곡선

사용하였으며, 특징으로 사용된 에지의 임계값은 10.0 이었다. 에지 위치에 대한 정합값은 문헌[3, 4]에서 제안한 상호 상관 계산법을 사용하였으며 사용한 마스크 창은 X,Y 축 및 45도, 135도 선상의 17×17 의 크기였다. 그림 6은 각 특징점들에 대한 거리정보를 제안한 정보전파 신경회로망을 사용하여 동적계획법 특성실험을 수행후 일치점(정합점)을 찾아낸 후 화소 밝기로 표현한 영상정보이다. 여기서 밝은 점은 카메라로부터 가까운 점의 위치를 표현한다. 그림 7은 특징점들만의 영상 정보값을 제안한 정보전파 신경회로망에 의해 전체 영상 정보로 보간시켜 구해진 거리정보값을 삼차원적으로 나타낸 그림이며, Z 값이 큰 값을 갖는 화소의 위치는 카메라와의 거리가 가까움을 의미한다. 이 실험 결과로, 특징점에 해당되는 거리 정보로부터 전체 영상점에 대한 위치점을 점차로 보간시켜 자동차의 고속 주행시에 얻어지는 물체에 대한 완전한 거리정보를 구해낼 수 있음을 입증시켜 준다.

원격지의 영상자료를 송신하거나 이동물체를 실시간으로 인식해낼 수 있도록 신경회로망 내부의 정보전파 기법을 통하여 실시간 보간되는 특성을 갖는 정보전파 신경회로망을 제안하였다.

제안한 신경회로망에 있어서 연결강도의 서로 다른 값들을 할당하여 다양한 형태의 보간 특성을 보임으로 최적의 보간 특성을 갖는 신경구조 방법을 제시하였다.

정보전파신경회로망 하드웨어의 노드 양끝단에 서로 다른 동적입력을 인가하여 보간 실험한 결과, 각 노드들이 두 입력 신호 사이에서 매우 좋은 보간특성을 보여주고 있음을 확인하였다. 신경회로망 노드의 한쪽에 가까울수록 그쪽의 입력파형에 가까워지게 합성 보간하였고, 또 다른 쪽 노드에 가까이 갈수록 그쪽 파형의 모양에 근접하여 합성 보간하여감을 확인하였다. 두 신경회로망 노드 사이의 보간시 시간 지연은 약 $6 \mu s$ 정도였다. 결과적으로 원격지의 영상을 획득하여 처리시키도록 임의의 좌우 영상을 취하여 실시간 보간 처리특성을 갖는 신경망 알고리즘을 보였다. 이러한 실시간 처리 방법으로 얻어진 거리 정보를 바탕으로 하여 원격지의 영상정보를 필요한 장소로 전송할 수 있으며, 또한 이러한 전송 기술을 수행시의 이동 차량에 응용 예와 무인 운전화 시스템 등에 일익을 담당할 것이다.

[참 고 문 헌]

- [1] G. Borgefors, "Distance Transformations in Digital Images", Computer Vision, Graphics, and Image Processing, Vol. 34, pp. 344-371, 1986.
- [2] R.Y.Tsai, "A versatile camera calibration technique for high accuracy 3D machine vision metrology using off the shelf T.V. Cameras and lenses," IEEE Trans. on Robotics and Automation, vol. 3, pp. 323-344, Aug. 1987.
- [3] L. Zhang, B. Curless and S. M. Seitz, "Spacetime stereo: Shape recovery for dynamic scenes," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. II-367-74, 2003.
- [4] J. Pages, J. Salvi, R. Garcia, and C. Matabosch, "Overview of coded light projection techniques for automatic 3D profiling," IEEE International Conference on Robotics and Automation, pp. 133-138, September 2003.



(a) 좌측 영상 정보 (b) 우측 영상 정보
그림 5. 스테레오 비전 실험을 위한 영상 정보

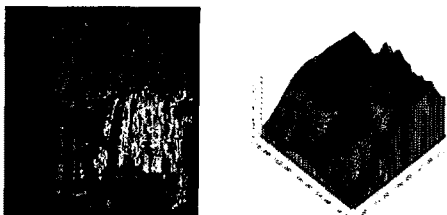


그림 6. 특징점의 정합된 후의 거리 영상 정보
그림 7. 특징점에 대한 LFN을 이용한 거리정보 보간 결과