

스테레오 영상의 실시간 정합을 위한 보간 신경망 설계

김종만

* 남도대학 전자제어과

A Propagation Programming Neural Network for Real-time matching of Stereo Images

Jong-man Kim

Abstract

Depth error correction effect for maladjusted stereo cameras with calibrated pixel distance parameter is presented. 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 obojects. All of them request much memory space and time. Therefore the most reliable neural-network algorithm is driven for real-time matching of obojects, which is composed of a dynamic programming algorithm based on sequence matching techniques.

1. 서 론

산업현장에서 물류 이송이나 이동중인 자동차 등의 영상 자료에서 거리를 산출하여 실시간으로 전송하는 기술은 산업 경제 뿐 아니라 안전 문제에서도 중요한 분야이다. 전기전자 기술의 급속한 발달에 따라 원격지의 정보를 송신하는 기술 또한 매우 유용하다. 특히 화상탐사 로봇을 이용한 과학 기술 탐사분야나 오지 벽지섬 병원등의 비상 수술을 위한 의료 기술 분야 등에 있어서 원격지의 화상정보를 기술탐사 본부나 중앙병원 본부 등으로 전송하는 기술들을 위하여 스테레오 카메라를 이용한 정확한 거리 계측 방법은 필수적이다. 또한

개인,직장일과 사업일로 인하여 자동차를 이용시, 바쁘고 복잡한 운전 환경하에 졸음이나 극도의 피로 상태 하에서도 불가피하게 운전을 해야하는 경우에 인간의 목숨과 관련된 자동차의 안전 운전과 관련한 연구는 매우 중요시 되는 분야이다. 특히 자동차가 운전자의 도움없이 스스로 자율운행하는 시스템 제어기술은 앞에서 언급된 여러 불가피한 운행상태하의 환경에서 매우 필수적으로 요구되기 때문에, 전력전자 및 계측제어 등 전기전자 기술의 발달과 아울러 많은 연구,개발이 진행되어지고 있다. 갑자기 출현한 도로 장애물에 실시간 인식하는 기술은 주로 스테레오 영상을 이용한 기법으

로 정합기술과 아울러 영상 보간 기술이 중요하게 사용되는 정보추출기법이다.

보간 기법을 이용한 경우로서 저해상도를 가지는 영상시스템에서 고해상도의 원래의 영상 정보로 보간시키는 방법이 수행되었다[1][2]. 또한 지형의 고

도 윤곽 정보에 대한 응용 부분으로써, 국토 건설, 도시 계획시의 채광 및 통신 장애 지역 등에 압축된 등고선 지형도 등을 이용하여 3차원 실재 지형도를 복원시켜 사용하였다[8]. 정보 보간 특성을 갖는 기존모델로 Inverse Distance 기법, TIN(삼각불규칙망) 기법, DT기법 등의 알고리즘을 이용한 기법 등 수학적인 계산방식들은 보편적으로 주위의 정보를 잘 계산해내지만, 굴곡이 심한 부분 등에 부분적인 오차가 자주 발생되며, 특히 계산 시간이 오래 걸리는 단점을 가지고 있어 불편함이 있다. 이와 반면, 병렬처리의 원리를 이용하는 신경회로망은 내부의 연결강도를 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[4-7].

본 연구에서는 자동차 주행중이나 필요한 영상 정보를 원격지 거리 전송시에 거리 산출하는 목적을 위하여 정보를 보간시키는 알고리즘을 제안하였다. 먼저 추출한 영상 정보를 기초로 하여 스테레오 비전 처리시 임의의 위치에 대한 일치성 문제인 동적계획법의 해결 방법을 가능케 해주고, 아울러 얻어진 비전 정보를 통해 특징점 중심의 거리정보로부터 전체 거리정보로 실시간보간 기능을 실시하므로 스테레오 비전의 실시간 동작처리를 가능케 해준다.

2. 일반 스테레오 비전

2.1 스테레오 카메라를 이용한 화상 계측 정보의 전송 절차

스테레오 비전은 인간의 거리 인식 원리와 같이 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 두 대의 스테레오 카메라를 사용하여 화상을 획득한다. 스테레오 방식으로 깊이 정보를 추출하는데 중요한 문제는 3차원 실좌표계의 한점에 대한 왼쪽의 화상과 오른쪽 화상의 같은 위치점을 찾아내는 정합문제이다. 좌우 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 목표로 하는 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 원격지 목적지까지 송신후 다시 재생시키는 기본 절차를 한 예로 하여 스테레오 비전 처리하는 경로에 대한 전체 흐름도를 보였다.

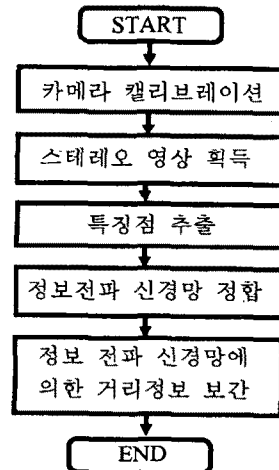


그림 1. 스테레오 비전 정보 전송의 거리산출 기본 절차

2.2 스테레오 카메라의 투사시스템

카메라의 일반적 투사과정은 임의의 3차원 공간상의 점을 평면에 투사시키는 투사 변환법을 이용하는 핀홀 모델을 이용한다. 이 핀홀 모델을 이용한 두대의 카메라를 사용하는 스테레오 카메라의 화상처리 과정은 그림 2와 같이 일정 거리의 3차원 공간상의 점의 거리 정보가 스테레오 카메라를 통하여 투사된 두개의 화상 평면으로 나타나는 원리를 이용한다.

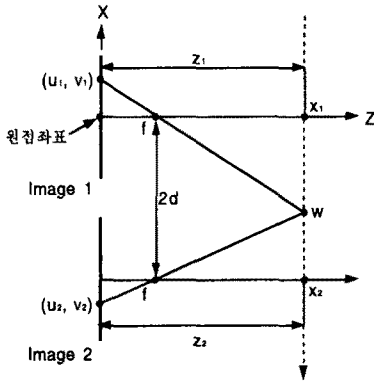


그림 2 스테레오 비전의 기하학적 구조

여기서 f(focus)는 카메라의 초점거리이고, z (distance 혹은 depth)는 물체의 위치에서 화상면까지의 거리를 의미하며, 2d는 epipolar선 즉, 두 카메라의 중심과 중심을 연결한 선분을 의미한다. 물체의 좌표를 $P_1(x_1, y_1, z_1)$, 좌측 화상에 대한 투사되는 좌표를 (u_1, v_1) , 물체의 우측화상에 대한 투사되는 좌표를 (u_2, v_2) 라 하면 앞의 기하학적 구조는 다음과 같은 관계식으로 표현된다.

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

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

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

$$z = f - \frac{2df}{u_2 - u_1} \tag{3}$$

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

3장. 실시간 보간을 위한 정보전파신경망

3.1 보간특성을 갖는 정보전파 신경회로망

본 연구에서 제안한 정보전파 신경회로망(LIPN)은 그림 3과 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 1차원구조이다. 그림에서 I_i 와 P_i 는 각각 i 번째 입력과 출력을 나타내며, N_i 는 i 번째 노드를 나타낸다. 연결강도 c_i 는 i 번째 노드의 입력연결 연결강도이다. 이러한 1차원 LIPN의 i 번째 노드의 출력은 식(4)과 같다.

$$P_i = f [c_i I_i + \sum c_{i,i+1} P_{i+1}]; l = \pm 1 \tag{4}$$

여기서 f는 신경망 노드의 활성화 함수이다.

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

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.

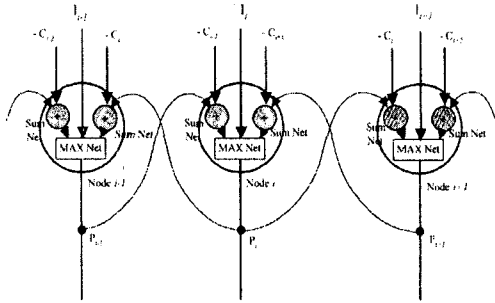


그림 3. 제안한 신경회로망의 1차원 연결과 노드의 구조

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

3.2 정보전파 신경회로망의 연결강도

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.

식 (1)의 f 가 선형 활성화함수일 때, 연결강도 c_i 를

입력측 연결강도 w^i 와 출력의 연결강도 $w^o_{l,m}$ 그리고 출력의 스케일 연결강도 w^s 를 적용하여 식 (5)로 표기할 수 있다.

$$O_i = w^s (w^i I_i + \sum w^o_{i,i+1} O_{i+1}) ; l = \pm 1 \quad (5)$$

만일 w^o 가 w^i 보다 작고 w^s 가 w^i 의 역관계이면, 식 (5)는 식 (6)과 같다.

$$O_i \cong w^s (w^i I_i) = I_i \quad (6)$$

여기서, 노드의 w^o 는 조건 [A]를 만족시키기 위해서는 입력이 있는 노드에 있어서 매우 작아야 한다. 또한 만일 w^o 가 w^i 보다 매우 크면, 식 (5)는 식 (7)로

$$O_i \cong w^s (\sum w^o_{i,i+1} O_{i+1}) ; l = \pm 1 \quad (7)$$

w^s 가 $\frac{1.0}{\sum w^o_{i,i+1}}$ 이면, 입력값이 없는 노드

는 인접노드의 연결강도의 평균값으로 보간시키게 된다. 그러므로 조건 [B]를 만족시키도록 노드의 연결강도 w^o 는 w^i 보다 훨씬 커야 한다.

따라서 프로그램 수행중 보간망은 입력값의 존재 유무에 의존하여 다른 고정된 연결강도의 값을 할당하므로써 새로운 연결강도 값을 생성시킨다.

3.3 제안한 정보전파 신경회로망을

이용한 동적계획법 구현

동적계획법은 주어진 제한 조건하에서 시작점으로부터 목적지까지 이르는 최적의 경로를 구하는 효과적인 방법이다. 그림 4와 같은 2차원 격자형 경로의 경우 시작점

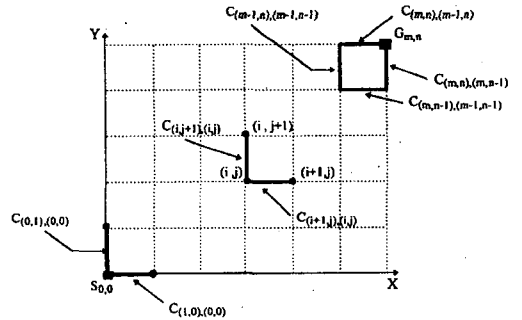


그림 4. 동적계획법을 위한 2차원 격자 경로

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

$$C^*_{0,0} = \text{MAX} \left\{ C_{(1,0),(0,0)} + C^*_{1,0}, C_{(0,1),(0,0)} + C^*_{0,1} \right\} \quad (8)$$

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

$$C^*_{i,j} = \text{MAX} \left\{ C_{(i+1,j),(i,j)} + C^*_{i+1,j}, C_{(i,j+1),(i,j)} + C^*_{i,j+1} \right\} \quad (9)$$

이다.

$C_{(i+1,j),(i,j)}$ 와 $C^*_{i+1,j}$ 사이에 합(+) 유니트를 설치하고 $C_{(i,j+1),(i,j)}$ 와 $C^*_{i,j+1}$ 사이에 합(+) 유니트를 설치한다. 합해진 두 신호는 MAX 유니트에 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 스테레오 영상 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.

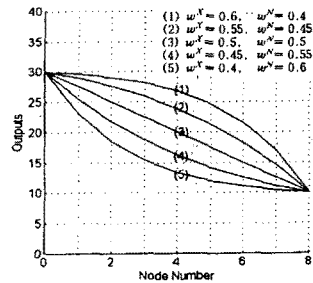


그림 5. 다양한 연결강도를 갖는 LIPN의 보간 곡선

그림 5에서 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 주행시에 얻어지는 스테레오 영상을 그림 6과 같이 실험실 내의 영상을 모델로 대체하여 특성 실험을 수행하였다. 또한 사용된 특징값으로는 에지 성분(윤곽선 정보)을 사용하였으며, 임계값보다 큰 에지 영상의 화소들 간에 상호 상관값을 계산하였다. 에지를 추출하기 위해 sobel operator를

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 7은 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 8은 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.



(a) 좌측 영상 정보 (b)우측 영상 정보

그림 6. 스테레오 비전 실험을 위한 영상 정보

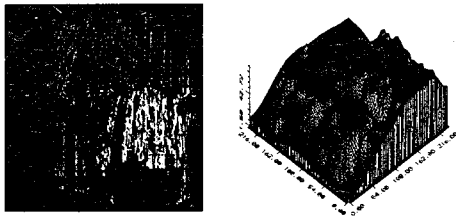


그림 7. 특징점의 정합된 후의 거리 영상 정보
그림 8. 특징점에 대한 LIPN을 이용한 거리정보 보간 결과

5. 결 론

원격지의 영상자료를 송신하거나 도로 장애물을 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 양끝단에서

로 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 입력의 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 주행시의 자동차에 응용할 때, 최적의 안전 운행 시스템을 환경을 만드는데 일익을 담당할 것이다.

5. 참 고 문 헌

- [1] Y. Yakimovsky & R. Cunningham, "A System for Extracting Three-Dimensional Measurements from a Stereo Pair of TV Cameras", Computer Graphics and Image Processing 7, pp.195-210, 1978
- [2] G. Borgefors, "Distance Transformations in Digital Images", Computer Vision, Graphics, and Image Processing, Vol. 34, pp. 344~371, 1986.
- [3] 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.
- [4] P. Puget and T. Skordas, "An optimal solution for mobile camera calibration", Proc. IEEE Trans. Syst. Man Cybern. 19(6), 1426-1445(November/December 1989)
- [5] Z. Zhang and O.D. Faugeras, "Calibration of a mobile robot with application to visual navigation", Proc. IEEE Work. Visual Motion, Irvine, California, pp. 306-313, March(1989)