

Stereo vision 및 응용

崔祐榮*, 朴來弘**
 明知大學校 電子工學科*
 西江大學校 電子工學科**

I. 서론

컴퓨터비전 (computer vision)의 궁극적 목표는 인간시각 시스템의 기능을 컴퓨터로 구현하는데 있으며 이를 구현하기 위해서는 입력된 영상정보로부터 3차원 물체를 찾아내거나 인식하는 것이 가장 기본적인 과제라 할 수 있다.

3차원 물체 인식에 있어서 2차원 영상으로부터 3차원 정보를 추출하는 과정은 매우 중요한 단계로서 한 장의 영상으로부터 깊이 정보를 추출하는 단안시법과 두 장 이상의 영상으로부터 깊이 정보를 추출하는 복안시법이 있는데 보통 후자를 스테레오 비전 (stereo vision)이라고 한다. 스테레오 비전은 인간의 시각구조와 유사하게 자연 조명상태하에서 물체들이 놓여있는 상태에서 영상을 취득하여, 취득된 좌우 영상으로부터 일치점의 변위를 구하여 간단한 삼각 측량 (triangulation)을 이용하여 깊이 정보를 검출하는 방법으로 특별한 조명이나 주변조건을 제한하지 않는 수동적 방법이다.^[1]

일반적으로 스테레오 비전 시스템은 카메라의 공간적 모형설정, 스테레오 영상정합 (stereo image matching), 깊이 정보의 보간 (depth interpolation) 등 일련의 과정들로 구성된다. 이 중에서 가장 많은 오차를 유발하는 과정은 스테레오 영상정합 과정으로 좌우 두 영상에서 얻은 두점이 3차원 공간상에 존재하는 같은 점임을 밝혀내는 일치성의 문제 (correspondence problem)를 해결해야 한다.

본 고에서는 스테레오 비전 시스템의 각 처리 과정에 대한 소개와 스테레오 비전의 응용에 대해 설명하고자 한다.

II. 스테레오 비전 시스템

1. 카메라의 공간적 모형설정

스테레오 정합을 하기 위한 영상을 취득하는 카메라의 기하학적 구조를 잘 이용하면 스테레오 정합과정을 보다 쉽게 할 수 있다. 공액쌍을 찾아내기 위한 일치성의 문제는 3차원의 물체 영상을 위치가 다른 두 지점에서 취득하기 때문에 각 영상에서 물체의 위치가 틀러지기 때문에 발생하며, 위치 변화의 양은 물체를 보는 각도에 따라 달라지게 된다. 특히 문제를 더 어렵게 하는 이유는 국부적인 점이나 면들이 다른 쪽의 영상에서 하나 이상의 정합 후보를 가지게 되므로 전체적인 일치성을 만족하도록 국부적인 특징들을 서로서로 연계하여 후보를 고려해야 하기 때문이다. 일치성 문제를 해결할 때 정합할 점들에 대한 탐색 영역은 영상을 취득하는 카메라들 사이의 기하학적인 구조를 이용하면 상당히 줄어들 수 있다. 일반적으로 두 카메라의 촛점과 영상의 한 점을 지나는 평면을 epipolar 평면이라고 하고 이 평면과 카메라의 영상면이 만나서 이루는 선을 epipolar line이라고 한다. 한 점에 대한 정합 후보는 이 점에 대한 다른 쪽 영상에서의 epipolar line 위에만 존재하게 되므로 이 선상의 점들에 대해서만 정합을 시도하면 되는데 이를 epipolar 구속 조건이라고 한다. 특히 두 영상의 수평선이 일치된 상에 놓여 있고 수직선이 평행한 경우 epipolar line은 각 영상의 수평선이 된다. 따라서 두 영상의 같은 라인상에 존재하는 특징들간의 정합만을 고려하여 탐색 영역을 줄일 수 있다.

두 대 카메라의 촛점거리 f 가 서로 같고 좌측, 우측 카메라의 각 영상 좌표계를 X_L, Y_L, Z_L 과 $X_R, Y_R,$

Z_R 이라고 할 때, 두 카메라의 광축을 각각 Z_L , Z_R 로 하여 서로 평행하게 되도록 한다. 좌측 렌즈의 중심과 우측 렌즈의 중심사이의 거리를 b 라고 하고 이 두 중심위치를 연결하는 선분을 베이스 라인 (baseline) 이라 한다. 삼차원 공간 좌표계 X - Y - Z 에 대해 좌측, 우측 영상 좌표계의 원점을 각각 $(0,0,f)$ 와 $(b,0,f)$ 에 놓고 좌측 영상의 Z_L 축과 공간 좌표계의 Z 축을 서로 일치시킨다. Epipolar 구속조건을 가정하면 $Y_L = Y_R$ 이고 그림 1과 같은 영상 취득 구조가 된다.

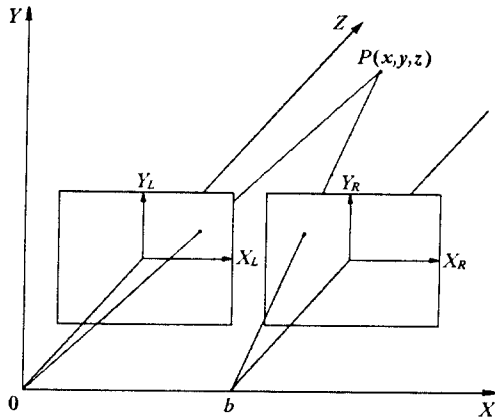


그림 1. 스테레오 영상을 취득하기 위한 카메라의 공간적 모형

만일 3차원 공간상의 한 점 $P(x, y, z)$ 가 좌, 우 영상면의 (x_L, y_L) 과 (x_R, y_R) 에 투영되었다고 가정하면 이들 사이의 관계는 삼각형의 비례관계를 이용하여 다음과 같은 식으로 나타낼 수 있다.

$$x_L = \frac{fx}{z}, \quad (1)$$

$$x_R = \frac{f(x-b)}{z}. \quad (2)$$

위의 두 식으로부터 z 를 구하면

$$z = \frac{bf}{x_L - x_R} = \frac{bf}{d} \quad (3)$$

이 된다. 이 때 $d (= x_L - x_R)$ 를 변위 (disparity)라고 하며 물리적 의미로는 인간시각에서 두 눈의 시차와 같은 의미를 갖는다. 식 (3)에서 b 와 f 는 카메라와 영상취득 구조에 의해 결정되는 값이므로 두 영상의 변위를 구하면 3차원 깊이 정보인 z 값을 구할 수

있다. 또한 이 식으로부터 알 수 있는 것은 깊이 z 와 변위 d 는 서로 반비례 관계가 있으므로 카메라와 가까이 위치한 물체일수록 보다 정확한 깊이 정보를 구할 수 있다. 즉, 멀리 위치한 물체 경우 변위 d 의 값이 작게 되므로 변위의 작은 오차에도 깊이 정보는 크게 틀려질 수 있으나 가까이 위치한 물체 경우에는 변위값이 크게 되므로 변위의 작은 오차에도 깊이 정보의 변화량은 상대적으로 크지 않으므로 보다 정확한 깊이정보를 구할 수 있다. 그리고 변위는 b 에 비례하므로 변위를 구할 때 생기는 오차값이 일정하게 고정되면 구하는 깊이 정보는 b 값이 크면 클수록 보다 정확한 값을 구할 수 있게 된다. 그러나 두 카메라 사이의 거리가 멀어지면 멀어질수록 두 카메라에 투영되는 공동 가시 영역 (common view area)이 줄어들게 되어 카메라에 가까이 위치한 물체들은 공동 가시 영역에서 벗어나게 되어 정합할 수 없게 되는 문제점이 있다. 또한 변위는 초점거리에 비례하고 초점거리에 따라 투영된 영상이 확대 또는 축소되므로 사용하는 카메라의 초점거리가 결정되면 깊이 정보를 구하고자 하는 물체들이 공동 가시 영역에 모두 놓일 수 있도록 두 카메라 사이의 거리 b 를 결정하여야 한다.

2. 스테레오 영상정합

스테레오 정합 방법은 크게 영역기반 정합 방법과 특징기반 정합 방법으로 나눌 수 있다. 영역기반 정합 방법에서는 이상적으로 개개의 화소에 대해서 다른 쪽 영상에서 상응하는 화소를 정합하는 것이 바람직하지만 화소 단위 정합에서 발생하는 모호성을 해결하는 것이 어려우므로 화소들이 이루는 의미있는 영역들간의 상관관계를 이용하여 정합하는 방식을 사용한다. 영역을 가장 쉽게 구할 수 있는 방법은 2차원 창 (2-D window)을 사용하는 것이며 이들 두 영상내의 국부영역 즉, 창의 밝기 패턴 (intensity pattern)의 유사성을 상호 상관값 (cross correlation) 측정 방법을 이용하여 결정하게 된다. 영역기반 정합 방법은 비교적 단순한 방법이지만 다음과 같은 제한조건을 가진다.

- 상호 상관값을 측정할 창내에 텍스처와 같은 특징이 있어야 하고 반복되는 무늬 (pattern)가 없어야 한다.
- 표면에서의 불연속성에 민감하다.
- 화소의 밝기, 대비, 깊이 영역의 급격한 변화,

조명 등에 민감하다.

이러한 조건이 위배되는 경우 대응점의 정합을 얻을 수 없으며 정합의 정확도는 창의 크기에 따라 좌우되며 창의 모든 점들에 대한 상관값을 계산해야 하므로 일반적으로 많은 계산 시간을 필요로 한다. 그러나 정합이 가능한 경우 특징기반 정합 방법처럼 특징점이 아닌 부분을 보간해야하는 단점이 없으므로 정밀한 복원이 가능한 장점이 있다. 영역기반 정합 방법은 항공영상의 분석과 같이 영상이 연속적이고 부드럽게 변하는 분야에서 효과를 거둘 수 있다.

영역기반 정합을 이용한 방법들에 대하여 살펴 보면 Moravec^[2]은 깊이 정보를 구하여 자동 주행장치에 이용하였다. Mori 등^[3]은 영역기반 정합방법에서 영상 취득구조 등에 의한 왜곡현상 (영상 밝기 또는 대비 등의 변화)의 문제점을 개선하고 반복적 예측과 교정을 하는 계층적 정합 방법을 제안했으며 항공영상에 적용하였다. Quam^[4]은 각 계층별로 영상의 해상도를 높이면서 각 계층에서 구한 변위정보를 다음 계층에서 이용하여 지형왜곡의 보상, 정합, 보간 등에 이용하는 자동 스테레오 정합 방법을 제안하여 영역기반 정합의 성능을 향상시키고자 하였다. Ikeuchi^[5]는 photometric 스테레오 시스템에서 구한 두 쌍의 면방향 지도 (surface orientation map)에서 방향성분이 같은 영역들을 분할하여 이들 영역간의 정합을 하여 깊이 정보를 구하는 방법을 제안하였다. 그는 분할된 영역의 중심에서 깊이 정보를 구하고 영역의 나머지 점들의 깊이 정보는 photometric 스테레오로부터 구한 방향정보를 적분하여 구하므로써 연속된 깊이 정보를 구하였다.

특징기반 정합 방법은 영상의 모든 화소에 대해 정합을 하지 않고 영상의 밝기 정보로부터 추출된 심볼 특징 (symbolic feature)를 이용한다. 이러한 심볼 특징에는 에지, 끝점, 모서리, 선소 및 특별한 연산자에 의해 선택된 점 등이 있으며 이러한 특징들은 영상 자체의 밝기보다는 밝기의 변화에 관련됨으로써 영상의 물리적 변화를 잘 나타낸다. 즉, 밝기 변화의 불연속은 실제 물체표면의 불연속을 의미하며 이러한 불연속은 표면의 깊이, 방향, 반사도, 조명 등에 의해 발생된다. 이러한 불연속으로부터 추출한 특징들을 이용하여 정합하므로 정합의 결과는 광학적 변화에 강한 특성을 가지게 된다. 또한 영상에서 특징들의 개수가 일반적으로 화소수보다 적으므로 전체적인 계산처리 시간도 영역기반 정합에 비해 빠르다. 그러

나 영상의 모든 점들을 정합하는 것이 아니므로 얻어지는 깊이 정보가 선기계 되어 깊이 정보를 구하지 못한 나머지 화소들에 대해서는 보간법을 이용하여 깊이 정보를 구해야 하는 단점이 있다. 그러나 이러한 단점에도 불구하고 장점이 갖는 특성이 훨씬 두드러지므로 많은 연구가 이루어져 왔다.

Marr와 Poggio^[6]는 인간의 시각구조에 근거하여 스테레오의 계산 구조 (computational structure)에 대한 이론적 기반을 수립하였다. 인간의 시각구조는 다음과 같은 과정을 거쳐 스테레오 문제를 해결한다고 보았다.

- (1) 스테레오 좌우 영상을 여러 크기를 갖는 창들에 의해 여러 방향으로 filtering 한다. 이 때 창의 형태는 DoG (Difference of Gaussian) 함수 형태로 한다.
- (2) 각 filtering된 영상으로부터 영교차 점들을 구한다.
- (3) 두 영상에서 각 창의 크기에 대해 같은 부호와 유사한 방향을 갖는 영교차점 쌍들을 정합한다. 이 때 모호한 정합쌍은 주변 영교차점의 정보를 이용하여 교정한다.
- (4) 큰 창을 사용했을 때 구한 정합 결과를 이용하여 작은 창을 사용할 경우의 정합에 사용한다.
- (5) 정합 결과를 2.5-D sketch라고 불리는 dynamic buffer에 저장한다.

Grimson^[7]은 Marr와 Poggio^[6]의 이론을 실제로 구현했는데 12개의 DoG 함수대신 하나의 LoG (Laplacian of Gaussian) 함수^[8]의 창을 이용하여 특징점을 추출하였다. 정합과정은 가장 크기가 큰 창으로 filtering된 영상부터 시작하여 가장 작은 창으로 filtering된 영상까지 반복 수행한다. 이 때 정합은 같은 라인상에서 수행되며 탐색 영역은 예상 변위와 창의 크기를 고려하여 정한다. 탐색 범위안에서 양쪽 영상의 영교차 점들이 서로 같은 부호를 갖고 30도 이내의 유사한 방향을 가지면 정합된 것으로 간주한다.

한편 Marr^[9]는 이치 (binary)의 random-dot stereograms을 정합하기 위한 구속 조건으로 다음과 같은 세 가지 조건을 제시하고 이를 이용한 cooperative 정합방법을 제안하였다.

- (1) 양립성 (compatibility): 검은 점은 검은 점끼리만 정합되어야 한다.
- (2) 유일성 (uniqueness): 한 영상의 검은 점은

다른 영상에서 항상 단 하나의 점과 정합될 수 있다.

- (3) 연속성 (continuity): 변위값은 깊이 정보가 불연속적인 곳을 제외하고는 서서히 변화한다.

Barnard와 Thompson^[10]은 연속성에 근거하여 각 특징점들간에 가능한 변위에 대한 예측치로부터 초기 정합 확률을 구하고 이를 반복적으로 갱신하는 이완 (relaxation) 기법을 적용하여 정합의 모호성을 줄이기 위한 방법을 제안하였다. Kim과 Aggarwal^[11,12]은 두 영상으로부터 영교차점을 구하고 연속성에 근거하여 수직 영교차점 연결 패턴을 정의하였다. 패턴과 빛의 밝기의 그래디언트 값의 유사도에 따라 초기 확률치를 구하고 정합의 모호성을 줄이기 위해 이완 기법을 적용하였다. 이 때 두 특징을 사용하여 구한 초기 확률치로부터 비교적 올바른 초기 정합 결과를 구할 수 있기 때문에 적은 반복횟수에도 빨리 수렴하는 결과를 구할 수 있다. Henderson 등^[13]은 연속적인 epipolar line들 사이에 에지의 연속성을 이용하여 정합의 모호성을 제거하려 하였다.

Baker와 Binford^[14]와 Ohta와 Kanade^[15]는 epipolar line 구속조건과 에지들 사이의 순서 구속 조건을 사용하여 여기에 동적계획 (dynamic programming) 기법을 적용하여 올바른 정합을 하고자 하였다. Lloyd 등^[16]은 epipolar 구속조건을 사용하여 좌우 영상의 같은 라인 상의 특징들을 정합시에는 동적계획 기법을 사용하고 epipolar line들 사이의 특징에 대해서는 이완 기법을 사용하는 방법을 제안하였다. Ohta와 Kanade^[15]와 Lloyd 등^[16]은 영상의 빛의 밝기, 에지들간의 길이 (length) 등을 특징으로 하여 스테레오 정합에 동적계획 기법을 사용하였다. 그러나 이들 방법에서는 영상취득시 기하학적 변화에 민감한 빛의 밝기, 에지들간의 길이를 특징점으로 사용하기 때문에 좌, 우 스테레오 영상에서 이들 특징들의 값에 차가 큰 경우에는 정합에 오류가 발생하는 단점이 있다. 최우영 등^[17,18]은 특징값으로 사용되는 영교차 에지가 각 스케일값에 따라 원영상의 실제 에지 위치와 달리 표현되어 정합하여 깊이 정보를 구했을 때 발생하는 오류를 줄이고자 스케일 공간상의 지문영상을 특징으로 사용한 정합 방법을 제안하였다. 스케일 공간상의 지문 영상을 특징으로 사용함으로써 모든 스케일의 영교차점을 고려할 수 있으며 기하학적 변화에 민감한 빛의 밝기 등을 사용하지 않으므로 정합의 정확도를 높일 수 있다.

Medioni와 Nevatia^[19]는 개개의 에지들을 정합하는 대신 연결된 에지들로부터 선소들을 추출하고 각 선소에 대해 정합가능한 범위안에 존재하는 이웃선소들로부터 구한 변위들간의 관계를 이용하여 정합하였다. 이들은 정합 수준을 에지 정합에서 한 단계 위인 선소들의 정합으로 끌어 올리는 데 기여했다. Lim과 Binford^[20]와 Kim 등^[21]은 에지로부터 시작하여 접합점 (junction), 곡선, 표면, 본체 (body)에 이르기까지 특징들을 계층별 구조로 하여 정합하는 방법을 제안하였다. Ayache와 Faverjon^[22]은 선소들 사이의 관계를 도출하여 이를 바탕으로 neighborhood graph를 구성하고 prediction-and-propagation이라는 heuristic을 사용하여 정합을 시도하였다. Horaud와 Skordas^[23]는 Ayache와 Faverjon^[22]의 방법보다 구체적으로 선소들 사이의 관계를 사용하여 relational graph를 만들고 이를 토대로 correspondence graph를 얻어 여기에 규칙 (rule)들을 사용하는 정합 방법을 시도하였다. Kim 등^[24]은 Horaud와 Skordas의 방법에서와 같이 복잡한 알고리즘을 사용하지 않고도 정합을 올바르게 하기 위하여 접합점, 선분, 면 등의 특징값들을 피지화하여 정합률을 높였다.

스테레오 정합의 특징으로 빛의 밝기 정보뿐만 아니라 색 (color) 정보를 사용하는 방법도 시도되었는데 이때 색정보는 정합의 효율이나 정확도를 높이기 위한 특징으로 사용하였다. Jordan과 Bovik^[25]은 정규화된 R, G, B 3색중 두 색의 차 R-G, G-B, B-R의 각 그래디언트 (gradient) 값과 방위 (orientation) 정보를 정합의 특징으로 추가 사용하여 정합의 정확도를 높이려 하였다. 이 방법은 영상에 존재하는 물체들간에 색상차가 있을 때 효과를 발휘할 수 있으며 빛의 밝기 정보만 이용할 때보다 R, G, B 각각에 대한 정보를 취득해야 하므로 많은 용량의 처리공간이 필요하다. Brockelbank와 Yang^[26]은 R, G, B 색 신호에 여러 색모델 (three-channel model, trichromatic model, opponent color model 등)을 적용하여 변환된 색 신호를 정합 특징으로 사용하였다. 이들 특징에 이완기법을 적용하여 정합을 했으며 실험 결과를 통하여 각 모델을 적용했을 때의 장단점에 대하여 서술하였다. ([26]의 표 V 참고) 이때 가장 좋은 결과를 얻을 수 있었던 모델은 보색 모델 (opponent color model)로 빛의 밝기 정보만을 사용한 경우보다 색 정보를 이용한 경우가 보다

정확한 깊이 정보를 구할 수 있음을 보였다.

Ayache와 Lustman^[27]은 카메라 세대를 사용하는 trinocular 스테레오 방법을 제안했는데 카메라를 한 대 더 사용함으로써 스테레오 방법보다 정합의 신뢰성과 정확성을 높였다. Pietikainen 와 Harwood^[28]도 수평 에지에 대한 미정합 문제와 가려진 물체에 의한 정합의 오류를 교정하기 위하여 trinocular 스테레오 방법을 사용하였다. Robert와 Faugeras^[29]는 trinocular 스테레오에서 정합특징으로 cubic B-spline으로 근사화된 곡선을 사용하여 특징을 직선 선소를 사용했을 때보다 정합률을 높이고 정합오차를 줄이고자 하였다. Okutomi와 Kanade^[30]는 스테레오 카메라 구조에서 베이스 라인 (baseline)이 짧을 때 구해지는 깊이 정보가 부정확해지고, 반대로 길 때 정합이 일어날 범위 즉, 변위 범위 (disparity range)가 넓어져 정합과정이 복잡해지거나 오류가 발생하게 되는 문제점을 해결하기 위해 여러 장의 스테레오 영상들을 이용하여 깊이 정보를 구하는 방법을 제시하였다. Kim과 Park^[31]은 영상취득시 양자화로 얻은 화소가 연결되어 나타나는 디지털 선분의 특성을 이용하여 화소단위 정합의 경우보다 확장된 형태의 양자화 오차의 영향을 분석하고, 이로부터 복원된 3차원 위치정보의 오차를 분석하였다.

3. 깊이 정보의 보간

스테레오 시스템에서 구해진 깊이 정보는 특징기반 정합 방법을 사용할 경우 특징점들에 대해서만 깊이 정보가 구해지게 되므로 조밀한 깊이 정보를 구하기 위해서는 보간이 필요하게 된다. 또한 영역기반 정합 방법을 사용할 경우에는 특징기반 정합보다 훨씬 더 조밀한 깊이 정보를 얻을 수 있지만 상호 상관관계를 구할 수 없는 영역에 대해서 구해지는 깊이 정보는 신뢰성이 떨어지므로 결국 이와 같은 영역에 대해서는 보간이 필요하게 된다.

가장 직접적으로 깊이 정보를 보간할 수 있는 방법은 희박하게 (sparse) 구해진 깊이 정보를 연속된 조밀한 깊이 정보로부터 표본화 (sampling)된 것으로 보고 이를 기존의 보간 방법을 사용하여 연속함수 (continuous function)로 근사화하는 것이다. 이것은 구해진 희박한 깊이 정보가 전체 깊이 정보에 대한 중요한 변화를 충분히 나타낼 수 있을 때 사용하기 적합한 방법이다. 따라서 항공 지형 영상과 같이 변화가 적은 경우에 사용하기 적합하다. 그러나 이러

한 방법은 물체들간에 가리움 (occlusion), 깊이 정보의 급격한 변화 등이 있는 경우에는 부적합하다.

이산적인 데이터 (discrete data)로부터 연속적인 함수를 보간하는 문제는 주어진 데이터가 희박하고 오차를 포함하며 추가적인 구속조건이 없을 경우 무한히 많은 표면 (surface)상의 점이 될 수 있으므로 이것은 불량 설정 문제 (ill-posed problem)이다. 불량 설정 문제를 풀기 위한 연구중에는 정칙화 기법 (regularization method)을 이용하는 방법이 있는데 이것은 다음과 같은 특성을 갖는 두항의 가중치 합을 최소화함으로써 해를 구하는 방법이다. 첫째 항은 해가 이미 알고 있는 정보와 얼마만큼 일치하는가를 측정하는 항이고 두번째 항은 해가 예상되는 조건에 얼마나 맞는지를 측정하는 항이다. 전자를 penalty 항이라고 하고 후자를 안정자 (stabilizer)라고 한다.

한편 Grimson^[32]은 영상에서 변화가 일어나는 특징점들사이의 표면에 대한 구속조건으로 영상 휘도 방정식 (image irradiance equation)으로부터 surface consistency constraints를 유도하였다. 이것은 영상을 구성하는 각 표면은 스테레오 방법 등에서 특징점들에 대해 구해진 깊이정보를 만족하고 이들 특징점사이에 존재하는 표면은 급격한 변화가 없어야 한다는 것이다. 또한 영상의 영역 R에서 깊이 정보를 알고 있는 기지점을 통과하는 최상의 표면은 기지점을 지나면서 다음 식을 최소화하는 표면임을 보였다.

$$\iint_R (f_u^2 + 2f_v^2 + f_w^2) dx dy \quad (4)$$

Grimson^[33]은 정칙화 기법을 사용하여 깊이 정보를 보간하는데 식 (4)를 안정자로 사용하였다. 보간하고자 하는 표면을 $f(x,y)$ 라고 할 때 다음 식을 최소화하는 표면을 구하므로써 깊이 정보를 보간하였다.

$$\begin{aligned} E(f) &= S(f) + P(f) \\ &= \iint (f_u^2 + 2f_v^2 + f_w^2) dx dy \\ &\quad + \beta \sum_i [f(x_i, y_i) - z(x_i, y_i)]^2 \end{aligned} \quad (5)$$

여기서 S는 안정화 범함수 (stabilizing functional)로 f의 변화를 측정하는 평활화 구속조건 (smoothness constraint)을 나타내고 P는 기지점의 정보와 보간되는 정보와의 차이를 측정하는 벌점

범함수 (penalty functional)이며, E는 이들 범함수로 이루어진 에너지 범함수 (energy functional)이다. 또 $z(x, y)$ 는 기지점에서의 깊이 정보값을 나타내며 β 는 보간 정보의 정확성과 평활화사이의 상대적 가중치를 나타낸다. Grimson은 식 (5)를 최소화하는 근사화 (approximation) 해를 위해서는 conjugate gradient방법을 사용하고 보간을 위해서는 gradient projection 방법을 사용하였다. Grimson의 방법은 반복 계산시 비교적 수치적으로 안정되며 병렬처리가 가능하다는 장점이 있다. 그러나 반복 계산식의 수렴이 느리고 수렴을 판단하기 위한 기준이 애매하며 영상전체에 대해 평활화 구속조건이 적용되므로 깊이가 불연속인 점에 대해서도 평활화가 되는 단점이 있다.

Terzopoulos^[34]는 Grimson 방법^[31]의 단점을 해결하기 위한 방법을 제시하였다. 먼저 깊이 정보와 면방향 (surface orientation)의 불연속점을 고려하기 위해 다음과 같은 연속성 제어 안정자 (controlled-continuity stabilizer)를 사용하였다.

$$S(f) = \iint \rho(x, y) [(1 - \tau(x, y))(f_x^2 + f_y^2) + \tau(x, y)(f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2)] dx dy. \quad (6)$$

이 때 (x, y) 와 (x, y) 의 범위는 $[0, 1]$ 사이값을 갖는다고 가정한다. 연속된 영역에 대해서는 $A(x, y) = \tau(x, y) = 1$ 로 하여 안정자를 thin plate spline이 되게 하여 C^1 연속성을 만족하는 표면이 생성되게 하고, 면방향의 불연속이 있는 영역에 대해서는 $A(x, y) = 1$, $\tau(x, y) = 0$ 로 정하여 안정자를 membrane 범함수가 되게 하여 C^0 연속성을 만족하는 표면을 보간하도록 한다. 깊이 정보에 불연속이 존재하는 경우에는 $A(x, y) = 0$ 로 하여 모든 연속성을 고려하지 않고 불연속 면을 포함한다. 또한 Grimson 방법에서 순환식의 수렴속도가 느린 단점을 보완하기 위해 multigrid 이완 기법이 제안하였다.^[35,36] 이 방법에서는 해상도가 낮은 단계 (coarse level)에서 계산된 결과를 해상도가 높은 단계 (fine level)에서 초기치로 이용하고 여기서 계산된 결과는 다시 이전 단계로 반환 (feedback)하여 이전 단계에서 계산된 결과를 개선하는 형태를 취하였다. 이와 같은 두가지 방식을 취함으로써 표면의 불연속을 처리할 수 있으며 계산 시간의 단축과 각 해상도별로 보간된 깊이 정보를 구할 수 있는 장점이 있다. 그러나 Grimson의 방법에

서와 같이 수렴을 결정하기 위한 판단기준이 모호하며 기지점과 멀리 떨어진 경우 수렴이 느리며 multigrid를 사용하므로 수렴의 안정성이 명확하지 않다. 또한 $A(x, y)$ 와 $\tau(x, y)$ 를 결정하는 문제는 깊이와 면방향의 불연속을 찾는 문제로서 결국 영역분할 (region segmentation) 문제가 되므로 이들 영역의 명확한 구분이 힘들게 된다. Terzopoulos^[37]는 불연속점을 찾기 위한 방법을 제시했으나 정규화 과정에서 불연속을 찾는 경우 평활화 구속조건에 의해 미묘한 불연속점들은 무시되며 표면 근사화의 각 순환과정마다 가중치 함수 $A(x, y)$ 와 $\tau(x, y)$ 가 결정되므로 Grimson^[31]방법에서와 같은 유일해를 구할 수 없는 단점이 있다. 또한 불연속 판단을 위하여 사용되는 임계치를 각 순환과정마다 결정하여야 하고 임계치를 정하는 기준을 정하기 어려운 단점이 있다.

Terzopoulos와 같이 multigrid 방법이나 불연속 검출 등에 의해 보간법의 효율을 높일 수도 있으나 Jou와 Bovik^[38]은 순환식에 의해 보간을 할 경우 초기치를 얼마나 잘 구할 수 있는가에 따라서도 계산의 효율을 높일 수 있다고 보고 구속조건 확장 (constraint expansion) 방법에 의해 초기치를 구하는 방법을 제시하였다. 또한 깊이와 면방향의 불연속을 찾기 위하여 원영상으로부터 구한 에지 정보를 이용한 방법을 제시하였다.

III. 스테레오 비전의 응용

3차원 정보를 얻기 위한 스테레오 비전은 II절에서 언급한 응용이외에도 자동 주행 (autonomous navigation), 항공 지도작성 (aerial cartography), 모델기반 인식 (model-based recognition), 산업 자동화 시스템 등에 응용되고 있다.

자동 주행 시스템^[39,40]에서는 스테레오 비전에 의해 취득된 3차원 정보를 이용하여 주행환경에 놓인 물체들을 검출 또는 측정하여 목적지까지 주행하게 되는 데 만일 물체가 장애물인 경우에는 그 물체를 피하여 주행하게 된다. Moravec^[41]은 단순한 내부, 외부정경에 대해 장애물을 피해 주행하는 시스템을 구성하였으며, Levine 등^[42]은 화성 탐사에 이용될 로봇의 주행 및 검사용 시각 시스템을 위하여 스테레오 비전을 사용하였다. Sato^[43]는 스테레오 비전을 이용하여

운전자의 차와 그 운전자의 앞차들간의 상대적인 속도와 거리를 측정함으로써 사고예방을 할 수 있는 시스템을 제안하였다.

항공 지도작성분야에서는 항공영상이나 위성영상으로부터 스테레오 비전을 이용하여 3차원의 고도 정보를 추출하여 자동적으로 수치 지형도 (Digital Elevation Model, DEM)를 작성하게 된다. Day와 Muller^[44]는 기존의 스테레오 비전 알고리즘을 SPOT 영상에 적용하여 수치 지형도를 구하는 시스템들을 비교하였으며 Kang 등^[45]은 항공영상의 스테레오 모델링을 통해 수치지형도를 만들어 DEM과 정합하였다.

모델기반 인식 시스템에서는 수작업이나 CAD 등에 의해 생성된 3차원 모델과 실제로 취득된 3차원 정보를 비교하여 물체를 인식하는 시스템으로 Burr와 Chien^[46]은 물체의 3차원 정보를 구하고 이를 저장된 모델과 비교하여 인식하는 시스템에 스테레오 비전을 이용하였다.

산업분야에서는 로봇의 시각 장치나 자동 검사 (automatic inspection) 장치 등에 활용되고 있는데 Yakimovsky와 Cunningham^[47]은 3차원 정보를 얻기 위한 로봇 시각 장치에 두 대의 TV 카메라를 이용한 스테레오 시스템을 구축하였다. Porrill 등^[48]은 에지를 특징으로 사용한 스테레오 비전에 의해 인식된 정보를 사용하여 로봇트가 물체를 들어 올려 그 위치를 이동시킬 수 있는 시스템을 구성하였다. Nishihara^[49]는 잡음에 강하고 산업용 로봇트에 적용 가능한 실시간 스테레오 영상 정합기를 개발하였다.

IV. 결론

지금까지 스테레오 비전의 처리과정과 그 응용에 대해 알아보았다. 수동적 깊이 추출 방법인 스테레오 비전을 사용함으로써 능동적 깊이 추출 방법에 비해 환경에 제약을 받지 않으므로 산업에서의 로봇트의 3차원 물체 인식장치, 자동항법 (auto-navigation), 사진측량기술 (photogrammetry) 등 비교적 범용적으로 응용될 수 있다.

스테레오 비전의 처리과정중 기존의 특징기반 스테레오 정합방법들의 연구 방향은 다음과 같이 크게 세 가지 추세로 볼 수 있다.

- (1) 정합에 사용되는 새로운 특징추출에 관한 연구
- (2) 올바른 정합 결과를 구하기 위한 정합 방법에 관한 연구
- (3) 정합특징과 정합방법의 계층적 처리 구조에 관한 연구

앞으로의 연구 방향으로 기존의 깊이 추출방법을 보다 개선하기 위해 여러 특징들간의 계층적 구조를 이용한 정합 방법과 이들 방법을 실시간 처리하기 위한 하드웨어의 개발이 필요하다. 또한 조밀한 깊이정보를 얻기 위하여 단순히 특징점들에 대해서 구해진 깊이 정보만을 이용하여 보간하지 않고 영상의 빛의 밝기와 같은 다른 특징들을 함께 사용한 보간 방법에 대한 연구가 필요하다. 또한 실제 응용분야에서 이용되기 위해서는 추출된 깊이 정보를 기반으로 하여 3차원 물체의 인식 및 표현에 관한 연구도 진행되어야 하겠다.

參 考 文 獻

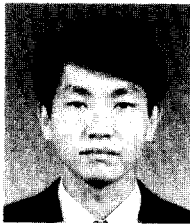
- [1] U. R. Dhond and J. K. Aggarwal, "Structure from stereo-A review," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-19, no. 6, pp. 1489-1510, Nov./Dec. 1989.
- [2] H. P. Moravec, "Towards automatic visual obstacle avoidance," in *Proc. Int. Joint Conf. Artif. Intell.*, Cambridge, MA, p. 584, Aug. 1977.
- [3] K. Mori, M. Kidode, and H. Asada, "An iterative prediction and correction method for automatic stereo comparison," *Comput. Graphics Image Process.*, vol. 2, no. 3/4, pp. 393-401, Dec. 1973.
- [4] L. H. Quam, "Hierarchical warp stereo," in *Proc. DARPA Image Understanding Workshop*, New Orleans, LA, pp. 149-155, 1984.
- [5] K. Ikeuchi, "Region-based stereo on needle maps," in *Proc. ICAR*, pp. 207-214, 1985.

- [6] D. Marr and T. Poggio, "A computational theory of human stereo vision," *Proc. Royal Soc. London*, vol. B204, pp. 301-328, 1979.
- [7] W. E. L. Grimson, *From Images to Surfaces*. MIT Press, Cambridge, MA, 1981.
- [8] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. Royal Soc. London*, vol. B207, pp. 187-217, 1980.
- [9] D. Marr, *Vision*. Freeman, San Francisco, CA, 1982.
- [10] S. T. Barnard and W. B. Thompson, "Disparity analysis of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 4, pp. 333-340, July 1980.
- [11] Y. C. Kim and J. K. Aggarwal, "Finding range from stereo images," in *Proc. Conf. Comput. Vision and Pattern Recognition*, San Francisco, CA, pp. 289-294, June 1985.
- [12] Y. C. Kim and J. K. Aggarwal, "Positioning three-dimensional objects using stereo images," *IEEE Journ. Robotics and Automation*, vol. RA-3, no. 4, pp. 361-373, Aug. 1987.
- [13] R. Henderson, R. Miller, and C. Grosch, "Automatic stereo reconstruction of man-made targets: Digital processing of aerial images," in *Proc. SPIE*, vol. 186, pp. 240-248, 1979.
- [14] H. H. Baker and T. O. Binford, "Depth from edge and intensity based stereo," in *Proc. Int. Joint Conf. Artif. Intell.*, Vancouver, B.C., Canada, pp. 631-636, Aug. 1981.
- [15] Y. Ohta and T. Kanade, "Stereo by intra- and inter-scanline search using dynamic programming," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-7, no. 2, pp. 139-154, Mar. 1985.
- [16] S. A. Lloyd, E. R. Haddow, and J. F. Boyce, "A parallel binocular stereo algorithm utilizing dynamic programming and relaxation labelling," *Comput. Vision, Graphics, Image Process.*, vol. 39, no. 2, pp. 202-225, Aug. 1987.
- [17] 최우영, 최종수, 이상욱, 박래홍, "스케일 공간상의 지문 영상을 이용한 스테레오 정합," 전자공학회 논문지, 제 28 권 B편 제 2 호, pp. 53-60, 1991년 2월
- [18] 최우영, 박래홍, "스케일 공간에서 동적 계획을 이용한 스테레오 정합," 전자공학회 논문지, 제 29 권 B편 제 8 호, pp. 44-53, 1992년 8월
- [19] G. Medioni and R. Nevatia, "Segment based stereo matching," *Comput. Vision, Graphics, Image Process.*, vol. 31, no. 1, pp. 2-18, July 1985.
- [20] H. S. Lim and T. O. Binford, "Stereo correspondence: A hierarchical approach," in *Proc. DARPA Image Understanding Workshop*, Los Angeles, CA, pp. 234-241, Feb. 1987.
- [21] D. H. Kim, K. B. Koo, W. Y. Choi, and R.-H. Park, "Stereo matching using hierarchical features for robotics applications," accepted, *Advanced Robotics*, vol. 8, 1994.
- [22] N. Ayache and B. Faverjon, "Efficient registration of stereo images by matching graph descriptions of edge segments," *Int. Journal Comput. Vision*, vol. 1, no. 2, pp. 107-131, 1987.
- [23] R. Horaud and T. Skordas, "Stereo correspondence through feature grouping and maximal cliques," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-11, no. 11, pp. 1168-1180, Nov. 1989.
- [24] D. H. Kim, W. Y. Choi, and R.-H. Park, "Stereo matching technique

- based on the theory of possibility," *Pattern Recongn. Letters*, vol. 13, no. 10, pp. 735-744, Oct. 1992.
- [25] J. R. Jordan, III and A. C. Bovik, "Computational stereo vision using color," *IEEE Control Systems Magazine*, vol. 8, no. 3, pp. 31-36, June 1988.
- [26] D. C. Brockelbank and Y. H. Yang, "An experimental investigation in the use of color in computational stereopsis," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-19, no. 6, pp. 1365-1383, Nov./Dec. 1989.
- [27] N. Ayache and F. Lustman, "Fast and reliable passive trinocular stereo vision," in *Proc. Int. Conf. Comput. Vision, London, England*, pp. 422-427, June 1987.
- [28] M. Pietikainen and D. Harwood, "Depth from three camera stereo," in *Proc. Conf. Comput. Vision and Pattern Recognition*, Miami Beach, FL, pp. 2-8, June 1986.
- [29] L. Robert and O. D. Faugeras, "Curve-based stereo: Figural continuity and curvature," in *Proc. Conf. Comput. Vision and Pattern Recognition*, Lahaina, Maui, Hawaii, pp. 57-62, June 1991.
- [30] M. Okutomi and T. Kanade, "A multiple-baseline stereo," in *Proc. Conf. Comput. Vision and Pattern Recognition*, Lahaina, Maui, Hawaii, pp. 63-69, June 1991.
- [31] D. H. Kim and R.-H. Park, "Analysis of quantization error in line-based stereo matching," *Pattern Recognition*, vol. 27, no. 7, pp. 913-924, July 1994.
- [32] W. E. L. Grimson, "Surface consistency constraints in vision," *Comput. Vision, Graphics, Image Process.*, vol. 24, no. 1, pp. 28-51, Oct. 1983.
- [33] W. E. L. Grimson, "An implementation of a computational theory of visual surface interpolation," *Comput. Vision, Graphics, Image Process.*, vol. 22, no. 1, pp. 39-69, Apr. 1983.
- [34] D. Terzopoulos, "Regularization of inverse visual problems involving discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 4, pp. 413-424, July 1986.
- [35] D. Terzopoulos, "Multilevel computational processes for visual surface reconstruction," *Comput. Vision, Graphics, Image Process.*, vol. 24, no. 1, pp. 52-96, Oct. 1983.
- [36] D. Terzopoulos, "Image analysis using multigrid relaxation methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 2, pp. 129-139, Mar. 1986.
- [37] D. Terzopoulos, "The computation of visible-surface representations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-10, no. 4, pp. 417-438, July 1988.
- [38] J.-Y. Jou and A. C. Bovik, "Improved initial approximation and intensity-guided discontinuity detection in visible-surface reconstruction," *Comput. Vision, Graphics, Image Process.*, vol. 47, no. 3, pp. 292-326, Sept. 1989.
- [39] D. B. Gennery, "A stereo vision system for an autonomous vehicle," in *Proc. 5th Int. Joint Conf. Artif. Intell.*, Cambridge, MA, pp. 576-582, Aug. 1977.
- [40] D. B. Gennery, "Object detection and measurement using stereo vision," in *Proc. 6th Int. Joint Conf. Artif. Intell.*, Tokyo, Japan, pp. 320-327, Aug. 1979.
- [41] H. P. Moravec, "Obstacle Avoidance and navigation in the real world by a seeing

- robot rover," *Stanford Artificial Intelligence Laboratory Memo AIM-340*, Sept. 1980.
- [42] M. D. Levine, D. A. O'Handley, and G. M. Yagi, "Computer determination of depth maps," *Comput. Graphics, Image Process.*, vol. 2, no. 2, pp. 131-150, Oct. 1973.
- [43] T. Sato, "Automotive stereo vision using deconvolution technique," in *Proc. 6th Int. Joint Conf. Artif. Intell.*, Tokyo, Japan, pp.763-765, Aug. 1979.
- [44] T. Day and J.-P. Muller, "Digital elevation model production by stereo matching spot image-pairs: A comparison of algorithms," *Image Vision Computing*, vol. 7 no. 2, pp. 95-101, May 1989.
- [45] M. S. Kang, R.-H. Park, and K.-H. Lee, "Recovering an elevation map using the aerial image sequence by stereo matching," *Optical Engineering*, vol. 33, no. 11, Nov. 1994.
- [46] D. J. Burr and R. T. Chien, "A system for stereo computer vision with geometric models," in *Proc. 5th Int. Joint Conf. Artif. Intell.*, Cambridge, MA, p. 583, 1977.
- [47] Y. Yakimovsky and R. Cunningham, "A system for extracting three-dimensional measurement from a stereo pair of TV cameras," *Comput. Graphics Image Process.*, vol. 7, no. 2, pp. 195-210, Apr. 1978
- [48] J. Porrill et al., "TINA: A 3D vision system for pick and place," in *3D Model Recognition from Stereoscopic Cues*. J. E. W. Mayhew and J. P. Frisby, Eds. MIT Press, Cambridge, MA, pp. 255-263, 1991.
- [49] H. K. Nishihara, "Practical real-time imaging stereo matcher," *Optical Engineering*, vol. 23, no. 5, pp. 536-545, Sept./Oct. 1984. 🌐

筆者紹介



崔祐榮

1961年 7月 1日生

1985年 2月 서강대학교 전자공학과 공학사

1987年 2月 서강대학교 대학원 전자공학과 공학석사

1992年 2月 서강대학교 대학원 전자공학과 공학박사

1985年 9月 ~ 1989年 8月 서강대학교 전자공학과 조교

1990年 9月 ~ 1991年 5月 생산기술연구원 HDTV 연구단 연구원

1992年 3月 ~ 현재 명지대학교 전자공학과 조교수

주관심 분야 : 영상신호처리, 컴퓨터 비전

筆者紹介



朴來弘

1954年 1月 1日生

1976年 2月 서울대학교 전자공학과 공학사

1979年 8月 서울대학교 전자공학과 공학석사

1981年 6月 미국 Stanford 대학 전기공학과 공학석사

1984年 6月 미국 Stanford 대학 전기공학과 공학박사

1978年 7月 ~ 1979年 6月 서울대학교 공과대학 정밀계측실 조교

1990年 1月 ~ 1991年 1月 미국 Univ. of Maryland, Computer Vision Lab. 객원부교수

1984年 9月 ~ 현재 서강대학교 전자공학과 교수

주관심 분야 : 영상신호처리, 컴퓨터 비전