

에지 정보와 밝기 정보를 이용한 특징 기반 정합

김 정호^{*,}, 엄 기문^{*}, 이 켈희^{*}

* 서강대학교 전자공학과

Feature based matching using edge and intensity

Jung Ho Kim^{*,}, Gi Mun Um^{*}, Kwae Hi Lee^{*}

* Dept. of E.E. Sogang Univ.

Abstract

The methods for stereo matching are divided into two techniques: area-based matching and feature-based matching. To find corresponding points by area-based method, it takes a lot of time because there are many points to be matched. Feature-based matching algorithm is often used because with this method it matches only some feature points so that the processing time is fast even though it requires interpolation after matching.

In this paper, we propose the smart technique by which we makes features simpler than conventional methods to match an image pair by feature-based matching algorithm.

I. 서론

스테레오 비전의 목적은 3 차원 위치정보를 취득하는 것이며, 인간의 시각구조와 유사한 방식으로 얻어진다. 즉 두 눈이 서로 다른 위치에서 영상을 획득함으로써 원근감을 느끼는 것과 같이 스테레오 비전 역시 이 효과를 이용하여 3 차원 정보를 얻는다. 일반적으로 3차원 위치의 정보획득 단계는 다음과 같다.

- 1) 영상 취득
- 2) 영상 취득시의 카메라 모델링
- 3) 영상의 정합
- 4) 3 차원 정보의 추출

또한 스테레오 비전은 단계 1), 2), 3)에서의 형태에 따라 다음과 같이 분류된다.

- 1) 영상의 갯수 - 2장의 영상
2장 이상의 영상
- 2) 영상 취득시의 카메라의 자세 - 수직영상
사각영상
- 3) 영상 정합 방법 - 특징 정합
영역 정합

본 논문에서는 2장의 수직 영상을 대상으로 하여 특징 기반 정합을 수행하였다.

특징 기반 정합은 영상의 특징점들을 사용하여 특징점들 간의 관계를 조사하여 정합하는 방법이다. 이 특징점들은 영

상의 밝기 정보 자체를 그대로 이용하지 않고 에지 연산자나 기타 다른 연산자를 사용하여 추출하기 때문에 비교적 영상의 왜곡, 영상 획득시의 조명의 강도에 영향을 받지 않는다. 또한 특징점들만 정합을 수행하기 때문에 계산 시간이 적게 걸린다는 장점이 있다. 단점으로는 영상의 특징이 되는 에지와 같은 성분은 영상내에서 간격을 두고 존재하므로 정합의 정도가 조밀하지 못하다. 또한 자연 영상에서는 특징 추출이 어렵기 때문에 이 방법을 사용하기가 어렵지만 이러한 영상에 대해서 특징정합을 이용한 경우도 있다 [1].

특징 기반 정합을 하기 위해서는 우선 특징을 정해야 한다. 에지는 이러한 정합을 위해 널리 사용되는 여러 특징들의 기본이 된다. 이러한 에지는 정합에 직접 사용되지는 않지만 이를 이용하여 한 단계 높은 특징을 찾아낸다. LoG(Laplacian of Gaussian) 연산자와 체인 코드(chain code)를 이용하는 방법 [2], 에지를 선과 같은 segment 로 표현하여 이를 이용하는 방법 [3], 혹은 2차원 영상을 1차원으로 분할하여 스के일 공간상의 지문영상을 구해 이를 정합의 특징으로 이용하는 방법 [4]등이 이에 속한다.

본 논문에서는 공선 조건이 맞춰진 상태에서 보다 빠르고 많은 정합점들을 얻기 위해 에지 정보에 변형을 가하지 않고 이 정보와 빛의 밝기를 특징으로 하여 정합을 하는 방법을 제안하였다. 빛의 밝기 정보는 여러 요인에 의해 왜곡될 수 있어서 그대로 이용할 수는 없다. 따라서 본 논문에서는 사용한 특징점으로는 영상의 한 행(line)상의 미분값에 임의의 임계값을 주어 국부 최대값을 구하고 이로부터 얻은 에지부분을 경계로 구역을 나누어 밝기의 평균을 낸 다음 그 구역의 길이와 옆 구역과의 평균과의 차이를 구하여 이를 특징으로 이용하였다.

에지를 찾을 때는 단순한 영상이 아니면 잡음의 영향을 많이 받으므로 이를 없애기 위해 Perona와 Malik에 의해 제안된 에지 영역은 보존되고 다른 지역은 평활화하는 방법을 사용하였다 [5].

II. 본론

1. 특징의 추출

1.1 에지 추출

특징을 결정하기 위해서는 우선 에지를 찾아야 한다. 주로 사용하는 방법으로는 LoG 연산자를 이용하는 것이지만 국부

최대값을 결정하기가 어렵고 에지가 원래의 위치에서 벗어날 우려도 있으므로 국부 최대값을 결정하기 쉬운 그래디언트 에지 연산자를 이용하였다.

영상의 한 행에 해당하는 영상 밝기 단면에 대해 아래와 같은 그래디언트 연산자 (gradient edge operator)를 적용하여 에지의 크기를 구한다. (단 에지를 스텝 에지로 가정한다.)

$$e(i,j) = |p(i,j+1) - p(i,j-1)| \quad (1)$$

여기서 $p(i,j)$ 는 영상의 (i,j) 번째의 화소의 밝기이다.

에지 단면중에서 실제 에지에 해당하는 부분은 그 값이 문턱값보다 크고 국부 최대값을 갖는 부분에 대해서만 에지를 검출한다. 2 차원적으로 에지를 찾으면 국부 최대치를 구하기가 어려우므로 이를 구하기 쉬운 1 차원 에지 연산자를 이용하여 X 축 방향의 에지를 구한다.

1.2 에지의 개선

에지에 해당하는 부분은 강조되고 이외의 부분은 평활화 된다면 이는 특징을 찾을때 도움이 된다. 이를 위하여 본 논문에서는 Perona와 Malik이 제안한 anisotropic diffusion을 이용하였다 [5].

Anisotropic diffusion 식은 다음과 같이 주어진다.

$$I_t = \text{div} (c(x,y,t) \nabla I) = c(x,y,t) \Delta I + \nabla c \cdot \nabla I$$

여기서, div : divergence 연산자.

∇ : gradient 연산자.

Δ : laplacian 연산자.

$c(x,y,t)$: diffusion coefficient.

만일 $c(x,y,t)$ 가 상수이라면 수식은 $I_t = c \Delta I$ 가 된다.

우리가 원하는 것은 영역 경계내에서만 빛의 밝기가 평활화되는 것이다. 이를 위해 conduction coefficient를 영역내부에서는 1, 경계에서는 0으로 놓음으로써 얻을 수 있다.

그러나 우리는 각 scale에서의 영역 경계를 모르기 때문에 추정기 (estimator)를 사용한다.

$E(x,y,t)$ 가 추정기라 할 때 다음과 같은 성질을 가진다.

- $E(x,y,t) = 0$ (영역 내부)

- $E(x,y,t) = K e(x,y,t)$ (각 edge point)

e : unit vector K : local contrast

또한 conduction coefficient는 $c = g(\|E\|)$ 이 된다. 여기서 $g(\cdot)$ 는 그림 1에 나타난 바와 같이 $g(0) = 1$ 인 단조 감소 함수이다.

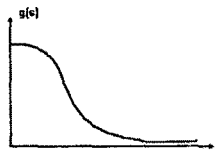


그림 1. $g(\cdot)$ 함수의 형태

여기서 $E(x,y,t) = \nabla I(x,y,t)$ 로 하면 conduction coefficient는 다음식과 같다.

$$c(x,y,t) = g(\|\nabla I(x,y,t)\|)$$

이를 실제로 구현하면 다음과 같다.

$$I_{ij}^{t+1} = I_{ij}^t + \lambda [c_N \cdot \nabla N I + c_S \cdot \nabla S I + c_E \cdot \nabla E I + c_W \cdot \nabla W I]_{ij}^t \quad 0 \leq \lambda \leq 1/4$$

∇ : nearest - neighbor difference

$$\nabla N I_{ij} = I_{i-1,j} - I_{ij}$$

$$\nabla S I_{ij} = I_{i,j-1} - I_{ij}$$

$$\nabla E I_{ij} = I_{i,j+1} - I_{ij}$$

$$\nabla W I_{ij} = I_{i,j-1} - I_{ij}$$

그리고 conduction coefficient는 다음과 같이 매번 그래디언트의 밝기로 갱신된다.

$$c_N^{t+1} = g(\|\nabla N I_{ij}^t\|)$$

$$c_S^{t+1} = g(\|\nabla S I_{ij}^t\|)$$

$$c_E^{t+1} = g(\|\nabla E I_{ij}^t\|)$$

$$c_W^{t+1} = g(\|\nabla W I_{ij}^t\|)$$

이식을 간단히 하면 다음과 같다.

$$c_N^{t+1} = g(1 \cdot \nabla N I_{ij}^t)$$

$$c_S^{t+1} = g(1 \cdot \nabla S I_{ij}^t)$$

$$c_E^{t+1} = g(1 \cdot \nabla E I_{ij}^t)$$

$$c_W^{t+1} = g(1 \cdot \nabla W I_{ij}^t)$$

이와 같이 t 번 반복 수행을 통하여 개선된 영상을 얻을 수 있게 된다.

$g(\cdot)$ 함수는 여러가지 방법으로 만들 수 있지만 본 논문에서는 아래식과 같이 지수함수의 형태를 이용하였다.

$$g(\nabla I) = e^{-(\|\nabla I\|/K)^2}$$

2. 정합

본 논문에서 제안한 정합 알고리즘은 다음과 같다.

공선 조건을 맞추고 두 영상의 같은 선상에서 에지를 구한 다음 이 에지들을 경계로 삼아 구간을 나누어 평균을 구한다. 이 평균을 이용하여 한 에지에서 왼쪽 구간의 길이와, 오른쪽 구간의 길이와, 왼쪽 구간과 오른쪽 구간과의 평균의 차이를 이용하여 이를 정합에 이용한다.

그림 2는 영상의 한 행의 밝기와 이 행에서 찾은 특징을 겹쳐서 나타내고 있다.



그림 2. 찾아진 특징

같은 선상의 에지 수를 세어 많은 쪽을 기준으로 삼고 다른 쪽을 정합해간다. 즉 영상 1의 에지 수를 ne_1 , 영상 2의 에지 수를 ne_2 , 영상 1의 i 번째 에지를 e_{1i} , 영상 2의 에지를 e_{2i} 라 하면

$ne_1 > ne_2$ 일때

$\text{margin} = ne_1 - ne_2$ 라 하면

e_{2i} 에 정합되는 영상 1의 에지는 $e_{1i-\text{margin}} \sim e_{1i+\text{margin}}$ 의 구간에서 찾도록 한다.

그리고 정합의 정도를 나타내는 척도로는 다음을 이용한 다.

$$M = K / (\alpha \times D_l + \beta \times D_r + \gamma \times D_m + 1) \quad (2)$$

D_L : 영상 1과 2에서 에지 왼쪽 구간의 길이 차이
 D_R : 영상 1과 2에서 에지 오른쪽 구간의 길이 차이
 D_m : 영상에서 얻은 에지 양 옆 구간의 평균차이의 차이
 K, α, β, γ : 상수

M 값이 클수록 정합 가능성은 높아진다.

정합시 제약 조건으로는 정합되는 점들이 겹치지 않도록 하였다.

3. 깊이 정보

본 논문에서 대상으로 하는 영상은 수직영상으로서 영상의 취득은 그림 3과 같이 이루어졌다.

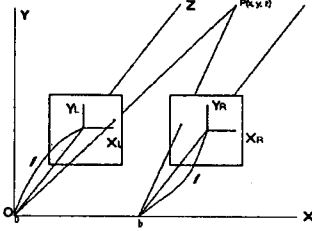


그림 3. 영상의 취득

여기서 양쪽 영상의 X 축 좌표는 $x_L = fx/z$, $x_R = f(x-b)/z$ 가 되고 Y축 좌표는 $y_L = y_R = fx/z$ 가 된다. 이러한 관계식으로부터 실제의 좌표는 다음과 같이 얻는다.

$$x = bx_L/(x_L - x_R) = bx_L/d$$

$$y = by_L/(x_L - x_R) = by_L/d$$

$$z = bf/(x_L - x_R) = bf/d$$

여기서, b : 카메라 간의 거리

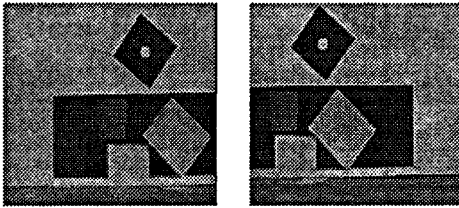
f : focal length

d : 시차

따라서 d 를 구하면 깊이 정보 z 를 구할 수 있게 된다.

III. 실험 결과 및 고찰

실험에 사용한 영상을 그림 4와 그림 5에 나타내고 있다.



a. 좌측 영상

b. 우측 영상

그림 4. 상자 영상



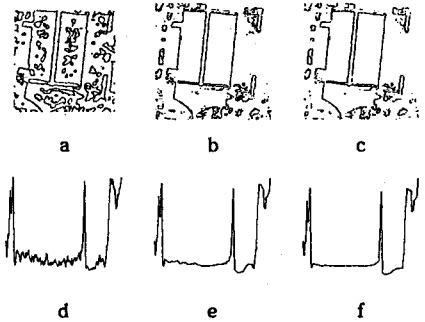
a. 좌측영상

b. 우측 영상

그림 5. 건물 영상

그림 4의 상자 영상은 CCD 카메라를 이용하여 얻은 256x256 영상이고 그림 5의 건물 영상은 비행기로부터 얻은 항공사진의 필름을 스캐너를 이용하여 얻은 4000x4000 영상 중 정합에 이용하기 위하여 필요한 부분을 자른 256x256 영상이다.

그림 6.a - 그림 6.c는 건물 좌측영상으로부터 찾은 에지를 보여 주고 있다. 그림 6.a는 LoG 연산자로 찾은 것으로 그림 6.b의 그라디언트 연산자가 더 좋은 결과를 보이고 있음을 알 수 있다. 그림 6.c는 Perona 와 Malik의 방법을 통해 평활화를 한후 찾은 에지를 보여 주고 있는데 직접 에지를 찾은 것보다 잡음이 많이 제거되어 더 좋은 결과를 보이고 있음을 알 수 있다. 그림 6.d - 그림 6.f는 이 영상의 40번 제행의 밝기를 보이고 있다. 이는 평활화 횟수를 증가시키에 따라 잡음은 감소되지만 밝기 변화가 심한 에지 부분에서는 평활화가 되지 않음을 알 수 있다.



a. 원영상의 LoG를 이용한 에지($\sigma=4$)

b. 1 차원 그라디언트를 이용한 에지

c. 평활화한 후의 에지(20회)

d~f. 영상의 한 행의 밝기

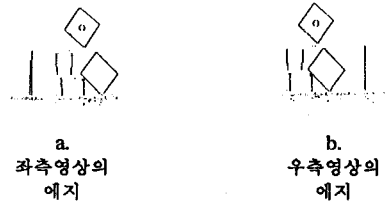
d. 원 영상

e. 평활화(20회)

f. 평활화(40회)

그림 6. 에지의 개선

그림 7은 식(1)의 그라디언트 연산자를 이용하여 찾은 상자 영상의 수직 성분의 에지이다.



a. 좌측영상의 에지

b. 우측영상의 에지

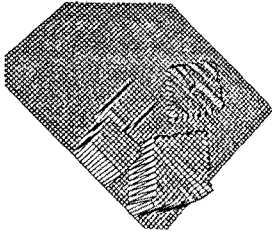
그림 7. 찾은 에지

그림 8은 그림 7을 정합하여 정합이 된 점을 표시하고 있다. 이 결과로 볼 때 평활화를 한 결과가 원 영상을 직접 정합한 것보다 정합이 잘 됨을 알 수 있다. 정합은 좌측 영상의 10 - 205 열, 100 - 205 행과 우측 영상의 0 - 195열, 10 - 150 행에 대해 수행하였다.

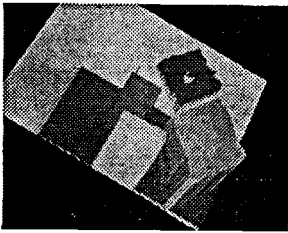
그림 9는 정합된 점의 시차를 이용하여 얻은 깊이 정보와 보간을 통한 결과를 perspective projection을 통해 3 차원적으로 재구성한 것이다.



a. 원 영상의 정합 결과 b. 평활화후의 정합 결과
그림 8. 정합된 결과.



a. mesh 형태의 그림



b. 원 영상을 입힌 그림
그림 9. 깊이 정보

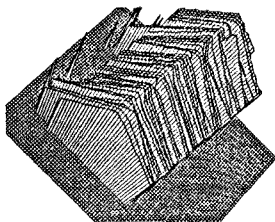
그림 10은 건물영상의 정합 결과를 보이고 있다. 정합은 좌측, 우측 모두 0 - 170 열, 20 - 185 행에서 수행하고 원래의 항공 영상 좌표로 환원시켰다. 이 두영상은 밝기가 서로 차이가 있지만 특징으로 사용한 것은 밝기가 아니라 에지 양 옆의 밝기의 차이를 이용하는 것이므로 두영상의 밝기가 달라도 정합에 큰 영향은 미치지 못하였다. 그렇지만 에지가 복잡한 부분에서는 정합이 잘 되지 않았다.



a. 좌측영상 b. 우측영상 c. 정합된 점

그림 10. 항공영상에 적용하여 정합된 결과

정합된 결과를 보간을 한 후 3 차원적으로 재구성하면 다음과 같다.



mesh 형태
그림 11. 깊이 정보

IV. 결론

본 논문에서는 특징 기반 정합을 하기 위한 특징을 간단하게 설정하였다. 특징이 간단하므로 구현하기가 쉽고 수행도 빠르게 되었다. 또 특징이 간단한 반면 잡음의 영향을 받으므로 이를 개선하기 하기위해 에지 부분이 아닌 곳은 평활화 하는 방법을 알아 보았다. 그리고 이를 이용하여 정합을 해 보았다. 그 결과 평활화를 한 경우가 보다 좋은 결과를 얻었음을 알 수 있었다.

본 논문에서 수행한 실험은 정합시 제약 조건으로서 단순히 정합점들이 겹치지 않도록 하였다. 그러므로 이완 정합 기법을 이용하거나 동적 계획 기법과 같은 좀더 발전된 방법을 이용한다면 보다 좋은 결과를 얻을 수 있으며 또 에지가 복잡한 경우에 있어서는 제한한 특징과 더불어 영역 기반정합의 방법을 결합한다면 더 좋은 결과를 얻을 수 있으리라 기대된다.

참고 문헌

- [1] J. S. Grenfeld, "Experiments with Edge-based Stereo Matching," *American Society for Photogrammetry & Remote Sensing*, vol. 55, no. 12, pp. 1771 - 1777, Feb. 1992.
- [2] Jeffrey J. Rodriguez and Aggarwal, "Matching Aerial Images to 3-D Terrain Maps," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI- 12, no. 12, pp. 1138 - 1149, Dec. 1990.
- [3] G. Medioni and R. Nevatia, "Segment-based Stereo Matching," *Comput. Vision, Graphics, Image Process.*, vol. 31, pp. 2 - 18, July 1985.
- [4] 최 우영, 최 중수, 이 상욱 박 래홍, "스케일 공간상의 지문 영상을 이용한 스테레오 정합," *전자공학회 논문지*, 제 28권, B편, 제 2호, pp. 53-60, 1991년.
- [5] Petro Perona and Jitendra Malik, "Scale-Space and Edge detection Using Anisotropic Diffusion," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI- 12, no. 7, pp. 629-639, Jul. 1990.
- [6] Yuan C. Hsieh, David M. McKeown, Jr. and Fredric P. Perlant, "Performance Evaluation of Scene Registration and Stere Matching for Cartographic Feature Extraction," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-14, no. 2, pp. 214 - 237, Feb. 1992.
- [7] William Hoff and Narendra Ahuja, "Surface from Stereo: Integration Feature Matching, Disparity Estimation, and Contour Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-11, no. 2, pp. 121 - 136, Feb. 1989.