

객체 분할을 위한 에피폴라 Rectification

강성숙*, 정승도**, 김수선***, 최병욱*

*한양대학교 정보통신대학원

**한양대학교 전자통신전파공학과

***한양여자대학 전산정보계열

e-mail : s_polaris@ihanyang.ac.kr

An Epipolar Rectification for Object Segmentation

Sung-Suk Kang*, Seung-Do Jeong**, Su-Sun Kim***, Byung-Uk Choi*

*Dept. of Information and Communications, Hanyang University

** Dept. of Div. of Electrical and Computer Engineering, Hanyang University

***Div. of Computer Science and Information Processing, Hanyang Women's College

요 약

Rectification 은 서로 다른 시점에서 얻은 두 영상을 동일한 평면에 투영시키는 호모그래피를 적용하여 원 영상의 에피폴라 라인을 수평방향으로 정렬시키는 과정이다. 이때 호모그래피에 의해 변환된 영상 사이에서도 에피폴라 제약조건이 성립해야 하며, 이를 만족시키는 호모그래피가 여러 개 존재하므로 제약조건을 추가하여 응용에 적합한 호모그래피를 구하게 된다. 본 논문에서는 Rectification 결과를 객체분할에 이용하기 위해, 원 영상과 유사한 결과를 얻을 수 있는 호모그래피를 구하도록 하였다. 기존의 방법은 복잡한 과정을 거쳐 최종적으로 구한 호모그래피를 적용한 후에도 결과 영상을 얻기 위해서 입력 영상에 따라 크기와 위치를 재조정해야 한다는 문제가 있는 반면, 제안한 알고리즘은 크기나 위치에 대한 조정 과정이 필요 없으므로 어느 영상에나 동일하게 적용할 수 있다. 제안한 방법으로 원 영상과 유사하고 원 영상 사이의 변위값을 유지하고 있는 결과 영상을 얻을 수 있었으며 이를 객체분할에 적용하였다.

1. 서론

서로 다른 시점에서 얻은 두 장의 영상을 이용하면 물체의 3 차원 정보를 얻을 수 있다.

가장 간단한 예로, 스테레오 카메라로부터 얻은 두 영상에서 3 차원상의 한 점은 동일한 수평선상에 존재하게 되므로 두 영상에서 일치하는 점을 쉽게 찾을 수 있다. 이때 두 대응점 사이에 수평 값의 차를 변위 값(disparity)이라 정의하며 깊이(depth) 정보를 의미한다.

하지만 보통의 경우 이러한 스테레오 영상을 얻기 힘들며, 이 두 영상에서는 깊이 정보를 구하기에 충분히 큰 변위값을 얻기 어렵다. 따라서 동일한 물체에 대해 다른 방향과 각도에서 찍은 영상을 이용하는 것이 바람직하다. 이 경우 첫번째 영상의 한 점과 대응하는 점은 두번째 영상의 에피폴라 라인(epipolar line) 상에 존재한다는 에피폴라 제약조건을 따르게 된다.

일반적으로 에피폴라 라인은 좌표축에 평행하지 않

아 대응점을 찾는 데 많은 시간이 소요되므로 에피폴라 라인을 수평 방향에 평행하도록 하여 문제를 간단히 할 수 있으며, 이 과정을 Rectification 이라 한다 [1][2].

기존의 Rectification 방법 중 카메라가 보정(calibration)된 상태[3]나 혹은 두 카메라의 상대적인 위치와 각도를 알고 있는 상태에서 접근하고 있는 방법[4]은 현실적인 적용이 어렵다. 카메라 보정 없이 또는 에피폴라 지오메트리 없이 Rectification 하는 방법[2][5][6]도 제시되고 있으나 계산하는 과정에서 오차에 민감하고, 수식적으로도 복잡한 과정을 거쳐야 한다. 그리고 대부분의 경우 두 영상을 동일한 평면에 투영시키는 호모그래피(homography)를 구한 후에 영상에 따라 크기나 위치를 조정해야 하는 문제가 있으므로 어느 영상에나 적용할 수 있는 방법을 제시하지 못하고 있다.

본 논문에서는 서로 다른 시점에서 얻은 두 영상으

로부터 에피플라 지오메트리를 계산하여 카메라 보정 없이도 어느 영상에나 동일하게 적용할 수 있는 Rectification 방법을 제안하고자 한다. 이때 원 영상과 유사한 결과를 얻을 수 있는 호모그래피를 적용하여, Rectification 후 변환된 두 영상에서 변위값을 구하고 이를 이용하여 영상 내 특정 객체를 분할하는 데 적용해 보고자 한다.

2. 이론적 배경

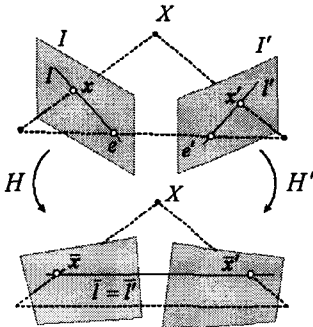


그림 1. 에피플라 지오메트리와 Rectification

그림 1 에서와 같이 좌, 우의 원 영상을 각각 I, I' 이라 하고, 3 차원 상의 점 X 가 I, I' 에 투영된 점을 x, x' , 에피폴을 e, e' , 에피플라 라인을 l, l' 그리고 H, H' 에 의해서 Rectification 된 영상에서의 점을 \bar{x}, \bar{x}' 라 한다.

$$x'^T F x = 0 \tag{1}$$

두 영상은 (1)의 에피플라 제약조건을 만족하게 된다. 이때 원 영상에서의 기본행렬(fundamental matrix), F 는 (2)와 같이 e' 과 임의의 행렬, M 의 곱으로 표현 할 수 있다[7].

$$F = [e']_x M \tag{2}$$

Rectification 은 동일한 기준 평면으로 두 영상의 에피플라 지오메트리를 변환하는 과정이라 볼 수 있다. 이러한 변환은 두 영상의 각 에피폴을 미리 정해진 한 점으로 옮기는 사영 변환(projective transformation), 혹은 호모그래피를 적용함으로써 가능하며, 보통 에피폴이 무한대의 점, $(1,0,0)^T$ 이 되도록 한다.

Rectification 후 결과 영상 간의 에피플라 지오메트리를 나타내는 기본행렬은 (3)으로 표현된다.

$$\bar{F} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \tag{3}$$

이때 변환된 영상은 (4)의 에피플라 제약조건을 만족하게 된다.

$$\bar{x}'^T \bar{F} \bar{x} = 0 \tag{4}$$

$$x'^T H'^T \bar{F} H x = 0 \tag{5}$$

$$F = H'^T [e']_x H \tag{6}$$

(4)에서 $\bar{x} = Hx$, $\bar{x}' = H'x'$ 이므로 (5)와 같이 쓸 수 있고, (1)과 (5)로부터 기본행렬은 (6)이 된다. 이것을 만족시키는 H 와 H' 은 여러 개가 존재하므로 제약조건을 두어 이를 만족하는 H 와 H' 을 구한다.

3. Rectification

3.1 과 3.2 에서 기존의 방법 중 대표적인 두 방법에 대해 고찰하고 3.3 에서 제안하는 방법에 대해 소개하고자 한다.

3.1 원 영상의 왜곡 최소화

[2]의 방법은, (6)을 만족하는 H 와 H' 을 각각 H_s, H_p, H'_p 와 H'_s, H'_p, H_p 로 재구성 해 나가는 방식이다. 먼저 e, e' 을 무한대 점으로 보내 각 영상의 에피플라 라인끼리 서로 평행하도록 하는 사영 변환, H_p, H'_p 을 구한다. 이때 H_p, H'_p 이 가능한 어파인 변환(affine transformation)이 되도록하여 영상의 왜곡을 최소화한다. 그 다음 무한대의 에피폴을 $(1,0,0)^T$ 방향으로 회전시키고, 둘 중 하나의 영상을 y 방향으로 이동시켜 두 영상의 에피플라 라인을 x 축에 평행하면서 서로 일치되도록 하는 유사 변환(similarity transformation), H_s, H'_s 을 적용한다. 여기서 H_s, H'_s 은 둘 중 한 영상에 대해 맞춰진 것이기 때문에 두 영상을 다 고려하는 비틀림(shearing) 변환, H_s, H'_s 을 적용하여 최종 결과 영상을 얻는다. 이때 H_s, H'_s 은 y 좌표는 그대로 유지한 채 x 좌표만 변화시키면서 어느 한쪽 영상이 커지거나 작아지지 않도록 해 준다.

그러나 최종 결과 영상을 얻기 위해서는 크기와 위치를 다시 조정해야 하는 문제가 있다.

3.2 결과 영상간의 변위값 최소화

[7]의 방법에서는, 먼저 오른쪽 영상의 에피폴을 무한대 점으로 보내는 H' 을 선택하고 (7)이 최소화되도록 하는 H 를 찾는다.

$$\sum_i d(Hx_i, H'x'_i)^2 \tag{7}$$

(2)와 H' 이 주어졌을 때, 임의의 벡터 a 에 대해 H 의 필요충분 조건은 (8)과 같다.

$$H = (I + H'e'a^T)H'M \tag{8}$$

(8)에서 H' 은 e' 을 $(1,0,0)^T$ 로 보내는 것이므로 $I + H'e'a^T = I + (1,0,0)^T a^T$ 가 되므로 H_a 의 형태가 된다.

$$H_a = \begin{bmatrix} a & b & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{9}$$

따라서 H 를 (10)과 같이 쓸 수 있다.

$$H = H_a H_0 \quad (10)$$

여기서 $H_0 = H'M$ 이고 $\hat{x}'_i = H'x'_i$, $\hat{x}_i = H_0x_i$ 라 하면, (7)은 (11)로 나타낼 수 있다.

$$\sum_i d(H_a \hat{x}_i, \hat{x}'_i)^2 \quad (11)$$

\hat{x}_i 와 \hat{x}'_i 를 각각 $(\hat{x}_i, \hat{y}_i, 1)^T$ 와 $(\hat{x}'_i, \hat{y}'_i, 1)^T$ 라 할 때, H' 과 M 은 이미 알고 있으므로 대응점, $x'_i \leftrightarrow x_i$ 로부터 \hat{x}_i 와 \hat{x}'_i 를 계산 할 수 있다. 결과적으로 (11)은 (12)를 최소화 하는 문제와 동일하다.

$$\sum_i (a\hat{x}_i + b\hat{y}_i + c - \hat{x}'_i)^2 \quad (12)$$

(12)를 만족하는 a 와 b , c 를 구하면 H' 에 대응하는 H 를 구할 수 있다.

그러나 H 는 결과 영상의 대응점, $\hat{x} \leftrightarrow \hat{x}'$ 의 차이가 최소화 되도록 하는 것이므로, H 에 의해 변환된 영상은 H' 으로 변환 된 영상과의 변위값이 최소화 된 결과로 나타나기 때문에 3 차원 깊이 정보를 잃어버릴 가능성이 높다.

3.3 제안하는 Rectification 방법

본 논문에서 Rectification 결과를 객체 분할에 이용하고자 세운 기본적인 가정은, 결과영상이 원 영상과 유사해야 하고 변위값은 유지 되어야 한다는 것이다

이 가정을 호모그래피를 구하는 제약조건으로 두고, 좌, 우 영상 모두 만족하는 결과를 얻을 수 있는 H 와 H' 을 구하였다.

제안하는 방법의 전체적인 과정은 그림 2와 같다.

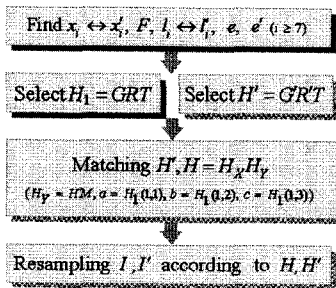


그림 2. 알고리즘 흐름도

조건을 만족시키도록 하기 위해 두 영상에 (13)을 적용하였다[8].

$$H = GRT \quad (13)$$

(13)에서 T 는 주어진 점 x_0 를 원점으로 이동시키

는 변환이고, R 은 원점을 중심으로 에피폴을 x 축 위의 점 $(f, 0, 1)^T$ 로 회전시키는 변환이다. 그리고 $(f, 0, 1)^T$ 는 (14)의 G 에 의해 $(1, 0, 0)^T$ 로 이동된다.

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/f & 0 & 1 \end{pmatrix}, |x/f| < 1 \quad (14)$$

G 는 x_0 에서 단위행렬(identity matrix)로 근사화 되는 특징이 있다.

대부분의 영상에서는 특정 객체가 영상의 중앙에 나타나므로 x_0 를 영상의 중심점으로 하고, 왼쪽 영상에는 (15), 오른쪽 영상에는 (16)을 적용하면 원 영상과 유사하면서 에피폴라 라인들이 x 축에 평행한 영상을 얻을 수 있다.

$$H_1 = GRT \quad (15)$$

$$H' = G'R'T \quad (16)$$

그림 3의 아래 두 영상은 원 영상 A, A' 가 각각 (15)와 (16)에 의해 변환된 영상이다.

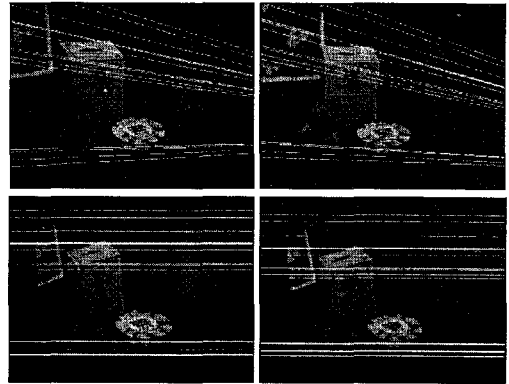


그림 3. 영상 A, A' (위)와 (15), (16)에 의해 변환된 영상(아래), 흰 선은 에피폴라 라인을 나타냄

다음으로 두 영상에서 서로 대응하는 에피폴라 라인이 일치 되도록 하기 위해, 오른쪽 영상의 H' 에 대응하는 H 는 (17)~(19)에 의해 구한다

$$H = H_X H_Y \quad (17)$$

$$H_Y = H'M \quad (18)$$

$$H_X = \begin{pmatrix} a & b & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (19)$$

H_Y 는 두 영상의 y 좌표에 대해서 정확히 일치시켜 주는 변환이다. 이것은 오른쪽 영상이 H' 에 의해 변환되었고 $F = [e']_x M$ 이므로 얻게 되는 결과이다.

H_X 는 x 좌표에만 영향을 주는 변환이며 a 와 b , c

의 값은 앞에서 구한 H_1 의 첫 번째 행벡터를 이용하여 원 영상과 유사하고 변위정보를 유지하는 H_1 의 특성을 H_x 에 부여할 수 있다.

마지막으로 구한 H, H' 을 각각 영상 I, I' 에 적용하여 에피폴라 라인이 정확히 일치된 결과 영상을 얻을 수 있다.

4. 실험결과

제안한 방법으로 구한 호모그래피를 적용한 결과, 기존의 방법과 달리 크기나 위치 조정 없이도 원 영상과 유사한 결과를 얻을 수 있었고, 변위값이 최소화 되어 3 차원 정보를 잃어버릴 수 있는 문제를 해결 할 수 있었다.

그림 4의 (a)와 (b)는 제안한 알고리즘을 적용하여 얻은 결과 영상이다.

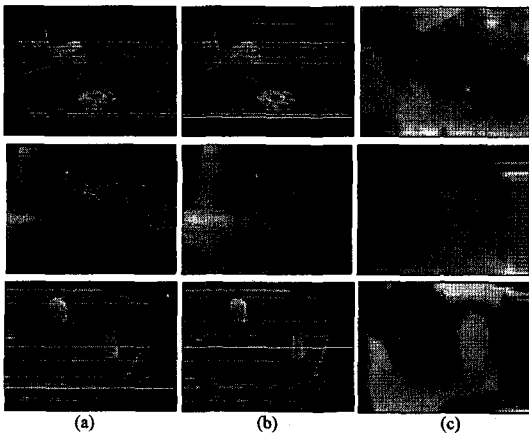


그림 4. A, A'(위) B, B'(중간) C, C'(아래)
Rectification 결과와 변위값 영상

그림 4의 (c)는 두 결과영상의 x 값의 차로 구한 변위값을 영상으로 표현한 것이며 카메라와 가까이 있는 곳일수록 밝게, 반대의 경우일수록 어둡게 나타난다.

Rectification 결과는 그림 5와 같은 과정을 통해 기존의 영상 분할 알고리즘 결과와 함께 객체 분할에 활용될 수 있다.

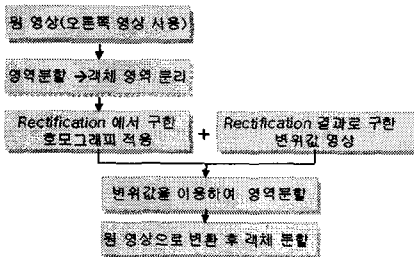


그림 5. 객체분할 방안

그림 6의 (a)는 영상 A와 B, C를 HSI 색상 공간(color space)에서 I 값을 기준으로 영역을 분할한 결과

이고, (b)는 영상에서 객체가 존재하는 영역에 호모그래피를 적용한 후 그림 4의 변위값 영상과 결합하여 얻은 결과이며, (c)는 객체에 존재하는 변위값을 기준으로 영상을 분할한 후 원 영상으로 다시 변환한 결과이다. 색상을 이용한 영상 분할 결과에서는 동일한 객체라도 색상이 다를 경우 여러 영역으로 나뉘어 지는 한계가 있으나 영상 내 객체의 3차원 정보를 이용하여 동일한 객체를 하나의 영역으로 분리 할 수 있다.

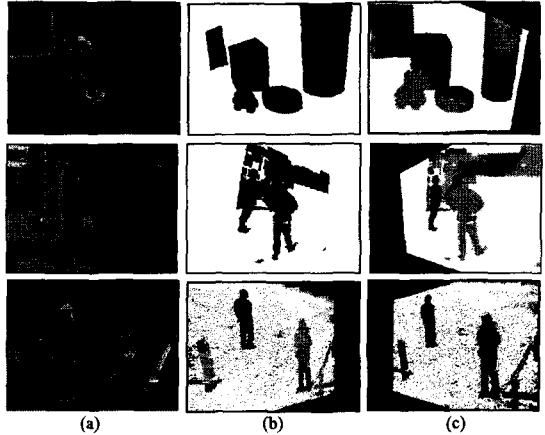


그림 6. 영역분할결과와 Rectification 결과를 이용한 객체분할결과

5. 결론

본 논문에서는 서로 다른 시점에서 얻은 두 영상으로부터 에피폴라 지오메트리를 계산하여 카메라 보정 없이도 어느 영상에나 동일하게 적용하여 원 영상과 유사한 결과 영상을 얻을 수 있는 Rectification 알고리즘을 제안하였다.

또 이 과정에서 얻은 결과를 영상 분할 알고리즘과 함께 이용하여 영상 내 특정 객체를 분할 할 수 있음을 보였다. 그러나 이 결과를 객체 분할에 적용하고자 할 때 원래의 두 영상이 서로 큰 차이가 나지 않는다는 가정이 필요하며, 변위값을 기준으로 영역을 묶을 때 원 영상의 스케일에 대한 고려가 있어야 하므로 이에 대한 추가적인 연구가 필요하다.

참고문헌

- [1] Linda G Shapiro and George C. Stockman, *Computer Vision*, Prentice-Hall, 2001.
- [2] Charles Loop and Zhengyou Zhang, "Computing Rectifying Homographies for Stereo Vision", CVPR, 1999.
- [3] D. V. Papadimitriou and T. J. Dennis, "Epipolar Line Estimation and Rectification for Stereo Image Pairs", IEEE Transaction on Image Processing, 1996.
- [4] A. Fusiello, E. Trucco, and A. Verri., "A compact algorithm for Rectification of stereo pairs", Machine Vision and Applications, 2000.
- [5] Richard Hartley, "Stereo from Uncalibrated Cameras", ECCV, 1992.
- [6] Francesco Isgro and Emanuele Trucco, "Projective Rectification without Epipolar Geometry", CVPR, 1999.
- [7] Richard Hartley, "Theory and Practice of Projective Rectification", International Journal of Computer Vision, 1998.
- [8] Richard Hartley and Andrew Zisserman, *Multifocal View Geometry in Computer Vision*, Cambridge univ. press, 2000.