

임의 시점 영상 합성을 위한 스테레오 영상 보정 기법에 관한 연구

“안진호”*, “장흥엽”**, “이재호”, 권용무*, 김상국*, 박상희*
*한국과학기술연구원, 정보전자연구부
**연세대학교 전기공학과

A Study on Rectification of Stereo Images for Synthesis of Arbitrary View Images

“Jin-Ho Ahn”*, Heung-Yeop Jang**, Je-Ho Lee*,
Yong-Moo Kwon*, Sangkuk Kim*, Sang-Hui Park**
*Division of Electronics & Information Technology, KIST
**Department of Electrical Engineering, Yonsei University

Abstract

This paper presents a new image rectification method using stereo image pairs to synthesize arbitrary viewing images. The image rectification is a preprocessing procedure that generates a vertically aligned images and makes epipolar lines collinear. The proposed rectification method has more robustness by considering the real situations such as the mismatch of pan angle, tilt angle and vertical shifting of cameras than previous method[5] which considers only pan angle. We evaluate the proposed method by comparison with previous method using synthesized images.

I. 서론

최근 가상 공간(virtual space)에 대한 관심이 증가함에 따라 많은 관련 기술이 연구되어지고 있다. 특히 실생활의 현실감을 영상 매체를 통해 그대로 표현하고자 하는 시도중 관찰자의 시점 변화에 대응하는 영상을 디스플레이상에 표시함으로써 실재감(real sensation)을 배가시키는 기술이 많은 유사분야에서 요구되고 있다. 그러나 관찰자의 변화에 대응하는 모든 영상을 카메라나 그 외 입력 장치를 통해 획득하는 것은 불가능하기 때문에 기존 영상 데이터를 이용해 임의의 시점 영상을 합성(Arbitrary View Generation)하는 기술의 필요성이 대두되고 있다[1-2].

기존의 임의의 시점 영상 합성 알고리즘으로는 2차원 영상의 3차원 모델링을 통한 재투영(reprojection) 방식이나 3차원 모델의 특징점을 이용하는 방식, 그리고 2차원 영상의 대응 관계를 이용하여 얻어진 변위나 깊이 정보를 사용하여 합성하는 방식이 있다. 이러한 알고리즘 대부분은 스테레오 정합을 이용한 정합점(corresponding point) 또는 고정점(invariant)의 정보를 이용하는데, 알고리즘을 효과적으로 구현하기 위하여 에피폴라 구속 조건(epipolar constraint)이 일반적으로 사용되고 있다. 대부분의 스테레오 정합 알고리즘은 에피폴라 구속 조건하에서 3차원상의 한 점에 대응하는 에피폴라 선은 y축의 좌표가 같다는 가정에서 시작한다. y축 좌표가 같기 위해서는 설치된 카메라의 광축(optical axis)이 평행해야 하는데 실제적으로 다시점 영상의 경우에는 일반적으로 원하는 물

체의 한 점에 카메라의 초점을 고정시키는 형태를 갖는다. 그러나 회전되거나 기울어진 카메라 구조하에 얻어진 영상은 정합점의 y축 좌표가 일치하지않기 때문에 기존의 정합 알고리즘을 맹목적으로 적용할 경우에는 중대한 에러를 유발한다. 따라서 광축이 평행한 상태의 영상으로 변환해주는 것이 필요한데 이 과정을 보정(Rectification)이라 한다.

지금까지의 대부분의 보정은 캘리브레이션(Calibration)을 통해서 얻어진 카메라의 내부 및 외부 파라미터의 정보를 이용하여 이루어졌다[3-4]. 그러나 캘리브레이션을 하기 위해서는 몇 개의 특징점의 3차원상의 실제 위치와 카메라에 의해 획득된 영상에서의 위치에 대한 정보가 필요하고 또한 실제적으로 정확한 카메라 캘리브레이션이 불필요하거나 불가능한 경우-예를 들면 object recognition이나 photogrammetry 응용 분야-에서는 이러한 방식의 적용이 어렵다. 따라서 캘리브레이션을 이용하지 않고 획득 영상의 특징을 추출하여 그 특징점의 대응 관계를 이용하는 방식의 보정 기법중 Papadimitriou와 Dennis에 의해 제안된 방식[5]을 기본으로 하여 실제 카메라 설치시 일어날 수 있는 다양한 구조에 의한 영상을 보정할 수 있는 있는 기법을 제안한다.

II. 보정의 원리

1. 카메라의 설치 구조

스테레오 영상에서 카메라 광축이 평행한 경우 카메라들의 공통 가시 영역(common viewing area)이 감소되므로 나타내하고자 하는 물체의 크기와 카메라사이의 거리상의 제약이 뒤따르게 된다. 카메라의 광축이 평행한 설치 구조는 주로 실외 영상(outdoor image)에 관련된 응용에 사용되는 구조로 실내 영상(indoor image)의 경우는 원하는 물체를 제대로 표현하기가 어렵기 때문에 실제 설치 구조에서는 카메라들의 초점을 일치시키는 구조를 갖는다. 이러한 구조에 의해 얻어진 영상은 카메라 광축의 회전 및 기울어짐, 그리고 기타 상하 등의 불균형으로 인하여 에피폴라 선이 평행하지 않게되므로 설치된 카메라 영상으로부터 임의의 시점의 영상들을 보간, 합성해내기 위한 기존 알고리즘의 맹목적인 적용은 중대한 에러를 유발하게 된다.

2. 영상 보정

영상 보정(image rectification)이란 카메라 광축이 평행하지 않은 경우의 영상들을 단순 기하학 구조의 영상들로 변환해주는 것을 의미한다. 여기서 단순 기하학 구조가 의미하는 것은 영상의 에피폴라 선이 영상의 가로 스캔 라인과 평행한 상태, 즉 에피폴라 선이 평행한 상태로 변위 추정(disparity estimation)시의 계산 시간을 감소하고 계산 과정 단순화 및 정합 에러를 감소하기 위하여 필수적인 영상 처리 과정이다.

III. 제안된 영상 보정 기법

기존의 보정식은 카메라의 기하학적 구조중 회전각만 고려함으로써 카메라의 경사각이 서로 다르거나 높낮이의 차이가 있는 카메라 구조에 의해서 발생하는 수직 변위에 대해서는 보정 영상이 왜곡되는 문제로 인하여 보정 효과가 크지 않다. 따라서 본 장에서는 기존의 보정식을 확장하여 카메라의 경사 및 상하 이동에 대한 파라미터를 첨가한 개선된 보정식을 제안한다.

그림 1은 카메라 광축이 회전된 형태의 구조이다. 그림에서 실선으로 표시된 영상(original image plane)은 카메라의 회전(pan)에 의해 획득된 원영상이고 점선으로 표시된 부분(normal image plane)은 카메라 광축이 평행한 경우의 영상이다. 3차원 좌표계의 한 점 P의 좌표를 (X, Y, Z)라 하였을 때 (x', y')와 (x'', y'')는 원영상에 투영된 P의 좌표이다. 왼쪽 카메라의 경우 두 점사이의 변위는 (d_L, d_r) = (x' - x'', y' - y'')로 정의할 수 있다. 원영상에서 점 P의 좌표인 (x', y')을 카메라 광축이 평행한 경우의 영상(그림 2에서 normal image plane)위로 투영했을 경우의 좌표를 (x_L, y_L)라 하였을 때 식 (1)과 같은 구조를 갖는다[6].

$$x_L = x_i = \frac{f(x'_i - x_b) \cos \theta_i - f_i \sin \theta_i}{(x'_i - x_b) \sin \theta_i + f_i \cos \theta_i} \quad (1)$$

$$y_L = y_i = \frac{f(y'_i - y_b)}{(x'_i - x_b) \sin \theta_i + f_i \cos \theta_i}$$

그림 2는 기준(reference) 카메라인 카메라 1에 대하여 camera 2가 a만큼 기울어졌을 경우의 기하학적 구조이다.

$$x' = \frac{f'(x'' - x''_b)}{(y'' - y''_b) \sin \alpha + f' \cos \alpha} \quad (2)$$

$$y' = \frac{f'[(y'' - y''_b) \cos \alpha - f' \sin \alpha - d]}{(y'' - y''_b) \sin \alpha + f' \cos \alpha}$$

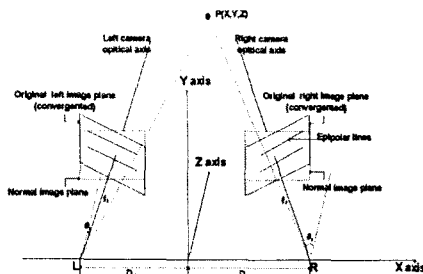


그림 1. 회전된 스테레오 카메라 구조[5]
Fig 1. Geometry of convergent stereo camera.

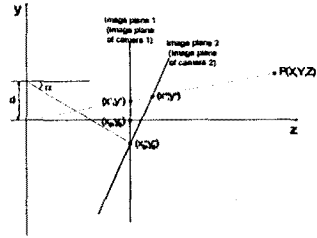


그림 2. 기울어진 카메라 구조
Fig 2. Geometry of camera tilted.

식 (1)의 경우와 같은 방식으로 식 (2)의 결과를 유도할 수 있다. 여기서는 캘리브레이션을 사용하지 않기 위해 새로운 방식으로 원영상의 정합점은 보정 영상에서의 y축 좌표값이 같다는 개념을 이용하였다. 즉 y_L = y, 또는 전개식으로 식 (3)처럼 나타낼 수 있다.

$$\frac{A_0 y_i'' + B_i}{C_0 x_i'' + D_0 y_i'' + E_i} = \frac{A_1 y_r'' + B_r}{C_1 x_r'' + D_1 y_r'' + E_r} \quad (3)$$

$$d_r = a_0 + a_1 x_i'' + a_2 y_i'' + a_3 x_r'' + a_4 x_i'' y_i'' + a_5 x_r'' y_i'' + a_6 y_i'' y_r'' \quad (4)$$

수직 변위 파라미터 y_L'' y_r''가 첨가된 식 (4)를 얻을 수 있다. 식 (4)에서 스테레오 정합 알고리즘을 사용하여 몇 개의 정합쌍 (x_L'', y_L''), (x_r'', y_r'')을 구한 이후에 linear least square 방법인 SVD를 통해 a₀, ..., a₆의 수직 변위 파라미터 값을 구할 수 있다. 최종적인 보정 영상은 식 (4)와 y_L - y_{r} = y_L'' - y_r'' - d_{r}를 결합하여 구한 새로운 보정식 (5)와 (6)을 이용하여 리샘플링과 보간을 통하여 구할 수 있다.}}

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} x_i'' \\ a_1 + b_1 x_i'' + c_1 y_i'' + d_1 x_i'' y_i'' + e_1 y_i'' \end{pmatrix} \quad (5)$$

$$a_1 = -a_0, b_1 = -a_1 + a_2 a_0, c_1 = 1 - a_2 + a_0 a_0, d_1 = -a_1, e_1 = -a_6$$

$$\begin{pmatrix} x_r \\ y_r \end{pmatrix} = \begin{pmatrix} x_r'' \\ y_r'' + b_r x_r'' + d_r x_r'' y_r'' \end{pmatrix} \quad (6)$$

$$b_r = a_3 + a_0 a_5, d_r = a_5$$

단 여기서 원영상의 좌표계만을 사용하기 위하여 식 (7)의 가정이 사용되었다.

$$x_r'' y_i'' = x_r'' (y_r'' + d_r) \approx x_r'' (y_r'' + a_0) \quad (7)$$

$$x_i'' y_r'' = x_i'' (y_i'' - d_r) \approx x_i'' (y_i'' - a_0)$$

$$y_i'' y_r'' = y_i'' (y_i'' - d_r) \approx y_i'' (y_i'' - a_0)$$

IV. 실험 및 결과 고찰

1. 실험

알고리즘 성능 평가 기준으로서는 수직 변위 보정 효과를 위한 부분으로 평균 오차(mean error)를 사용하였고 보정 영상의 왜곡(distortion) 정도의 평가를 위한 부분으로서 원영상과의 유사도를 비교하였다. 유사도 평가는 정규화된 상관 관계 함수(normalized correlation function)를 이용하였다.

표 1. 제작 영상 수직 변위
Table 1. Vertical disparity of synthesized images.

좌측 상단	-4
좌측 중단	-5
좌측 하단	-7
우측 하단	-3

표 2. 실험 결과
Table 2. Experimental results.

	기존의 보정식		제안된 보정식	
	왼쪽 카메라	오른쪽 카메라	왼쪽 카메라	오른쪽 카메라
a_0	-0.000699		0.000167	
a_1	-0.000203		0.000050	
a_2	-0.002937		0.000128	
a_3	0.002668		0.000121	
a_4	-0.000772		-0.000085	
a_5	0.000896		0.000074	
a_6	x		0.000181	
mean error	1.538916		1.683455	
standard deviation	1.407723		0.913707	
similarity (average)	0.591149		0.872782	

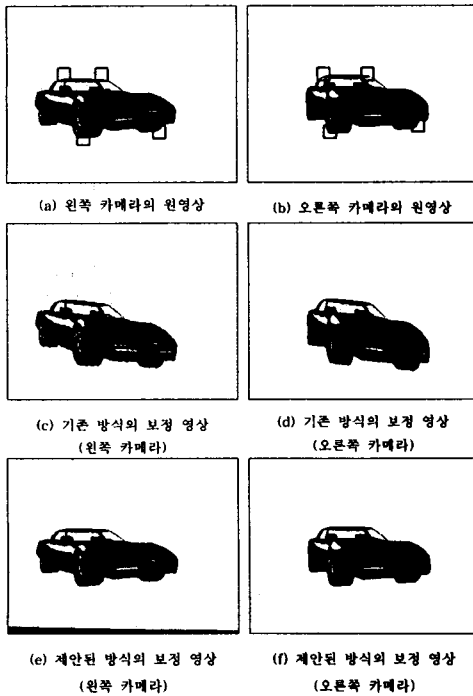


그림 3. 실험 결과

Fig 3. Experiment results.

2. 결과 고찰

제작 영상의 카메라 사이의 간격은 30cm이고 카메라 회전 각(pan angle = $\theta_1 + \theta_2$)이 12° , 경사각은 약 3° 로 배치되

었다. 제작된 원영상중 왼쪽과 오른쪽 카메라에 의한 원영상을 그림 3 (a)와 (b)에서 나타내었고 발생한 수직 변위는 표 1과 같다. 수직 변위 측정은 원영상의 정합점중 상하좌우의 가장 먼저 있는 것(그림에서 네모난 부분)의 수직 변위를 측정하였다. 그리고, 실험 영상은 320 x 240 크기로 하였다.

실험 결과는 표 2 (a)와 (b)에서 각각 나타내었다. 실험 결과를 보면 기존의 방식에 비해 제안된 방식의 수직 변위 보정 평균 오차가 근소하게 증가된 것을 확인할 수 있다. 이것은 추가된 파라미터에 의해 SVD에 의한 수직 변위 파라미터 값의 오차가 증가함으로써 발생한 수직 변위 오차의 폭이 파라미터를 추가함으로써 얻을 수 있는 보정 오차 감소의 폭보다 크기 때문에 일어난 현상이다. 그러나 제안된 방식의 변위 오차의 편차는 감소하여 전체적으로 고른 보정 효과가 있었고 또한 카메라의 경사각을 고려한 파라미터의 첨가로 인하여 보정된 영상의 왜곡도가 기존 방식에 비해 현저히 감소함을 확인할 수 있었다. 기존 방식의 보정 결과 영상은 그림 3의 (c)와 (d)에, 그리고 제안된 방식의 보정 결과 영상은 (e)와 (f)에 나타내었다.

V. 결론

본 논문에서는 임의 시점 영상 합성을 위한 전처리 과정으로서 확장된 수직 변위 파라미터를 사용하는 스테레오 영상 보정 기법을 제안하였다. 제안된 방식은 이전의 카메라의 회전만을 고려한 보정 기법을 확장하여 카메라의 기울어짐이나 수직 이동의 경우도 포함하여 실제 영상에 대한 보정이 가능하도록 하였다. 제작된 영상에 대한 실험 결과 제안된 방식은 기존 방식과 비슷한 수직 변위 보정 효과를 나타내는 반면 전체적으로 고른 보정이 가능하였고 또한 보정된 결과 영상은 기존 방식에 비해 왜곡이 없는 원영상과 유사하게 복원이 되는 우수한 보정 효과를 확인할 수 있었다.

참고문헌

- [1] A. Katayama, K. Tanaka, T. Oshino, H. Tamura, "A viewpoint dependent stereoscopic display using interpolation of multi-viewpoint images," *Proc. SPIE*, Vol. 2409, pp. 11-20, 1995.
- [2] N. L. Chang, A. Zakhor, "Arbitrary view generation for three-dimensional scenes from uncalibrated video cameras," *Proc. ICASSP*, Vol. IV, pp. 2455-2458, May 1995.
- [3] O. Faugeras et al, "Real time correlation-based stereo: algorithm, implementations and applications," Technical Report RR-2013, INRIA, 1993.
- [4] S. B. Kang, J. A. Webb, C. L. Zitnick, T. Kanade, "An active multibaseline stereo system with real-time image acquisition," Tech. Rep. CMU-CS-94-167, Dep. Computer Science, CMU, 1994.
- [5] D. V. Papadimitriou, T. J. Dennis, "Epipolar line estimation and rectification for stereo image pairs," *IEEE Trans. Image Processing*, Vol. 5, No. 4, pp. 672-676, April 1996.
- [6] A. Goshtagby, W. B. Gruver, "Design of a single-lens stereo camera system," *Patt. Recogn.*, Vol. 26, No. 6, pp. 923-937, 1993.