

스테레오 영상 보정 알고리즘에 기반한 새로운 중간시점 영상합성 기법

정회원 박창주*, 고정환*, 김은수*

A New Intermediate View Reconstruction Scheme based-on Stereo Image Rectification Algorithm

Chang-Ju Park*, Jung-Hwan Ko*, Eun-Soo Kim* *Regular Members*

요약

본 논문에서는 비교정 상태의 스테레오 입력영상에 영상보정 알고리즘을 적용한 새로운 중간시점 영상합성 기법을 제시하고 그 성능을 분석하였다. 제시된 방법에서는 먼저, 좌, 우 스테레오 영상의 각 화소 값들에 대한 유사도 및 모서리 검출을 통해 특징점을 추출한 다음, 이들 특징점을 이용하여 스테레오 영상간의 움직임 벡터와 에피폴라 선을 검출하였다. 그리고, 스테레오 영상간의 수평선을 일치시킴으로써 좌, 우 스테레오 영상을 보정하고 최종적으로 적응적 변위추정 기법을 이용하여 최적화된 중간시점 영상을 합성하였다. CCETT의 'Man' 영상과 스테레오 카메라를 사용하여 촬영한 '사람' 및 '자동차' 영상을 사용한 중간영상 합성 실험결과 본 논문에서 제안된 보정기법으로 교정된 스테레오 영상의 경우가 비교정 상태에 비해 'Man' 영상은 3.6dB, '사람' 및 '자동차' 영상은 2.59dB, 1.47dB의 PSNR이 각각 개선됨이 분석됨으로써 본 논문에서 새로이 제시한 스테레오 영상 보정 알고리즘 기반의 중간시점 영상합성 기법의 실질적 응용 가능성을 제시하였다.

Key Words : rectification; stereo image; intermediate views reconstruction; epipolar geometry.

ABSTRACT

In this paper, a new intermediate view reconstruction method employing a stereo image rectification algorithm by which an uncalibrated input stereo image can be transformed into the calibrated one is suggested and its performance is analyzed. In the proposed method, feature points are extracted from the stereo image pair through detection of the corners and similarities between each pixel of the stereo image. And then, using these detected feature points, the moving vectors between stereo image and the epipolar line is extracted. Finally, the input stereo image is rectified by matching the extracted epipolar line between the stereo image in the horizontal direction and intermediate views are reconstructed by using these rectified stereo images. From some experiments on synthesis of the intermediate views by using three kinds of stereo images; a CCETT's stereo image of 'Man' and two stereo images of 'Face' & 'Car' captured by real camera, it is analyzed that PSNRs of the intermediate views reconstructed from the calibrated images by using the proposed rectification algorithm are improved by 2.5dB for 'Man', 4.26dB for 'Face' and 3.85dB for 'Car' than those of the uncalibrated ones. This good experimental result suggests a possibility of practical application of the proposed stereo image rectification algorithm-based intermediate view reconstruction scheme to the uncalibrated stereo images.

I. 서론

최근, 양안 시차로 국한되는 기존의 스테레오

3D 디스플레이 방식의 단점을 해결하기 위한 접근 방법으로 여러 형태의 다시점 스테레오 3D 디스플레이 시스템에 관한 연구들이 활발히 진행되고 있

* 광운대학교 전자공학과 국가지정 3차원 영상미디어 연구실
(cjpark@explore.kw.ac.kr, misagi@explore.kw.ac.kr, eskim@daisy.kw.ac.kr)

※ 본 연구는 정보통신부 ITRC 사업 및 과학기술부 국가지정연구실(NRL) 사업의 지원 아래 수행되었습니다.
논문번호 : #040045-0202, 접수일자 : 2004년 2월 2일

다^[1]. 이러한 시스템은 다안식 입체 카메라를 통해 다시점의 영상획득과 디스플레이가 이루어지기 때문에 시점의 개수가 증가함에 따라 지역이 확대되고 보다 자연스러운 3D 디스플레이가 가능하다. 그러나 시점의 개수가 증가됨에 따라 영상 데이터도 동시에 증가하기 때문에 실시간적 영상처리 프로세서 및 초고속, 광대역의 전송채널이 또한 요구되게 된다^[2].

따라서, 이러한 문제점을 해결할 수 있는 새로운 접근방법으로 제한된 개수의 스테레오 카메라에서 입력된 영상으로부터 원하는 수만큼의 다시점 영상을 디지털적으로 합성해 낼 수 있는 중간시점 영상합성(IVR: intermediate view reconstruction) 기법이 제안되어 지역이 확대된 자연스러운 3D 입체 디스플레이의 구현 가능성을 제시하고 있다^[3].

그러나, 기존의 중간시점 영상합성 기법에서는 교정(calibration)이 필요 없는 이상적인 디지털 표준 영상을 사용하는 경우에는 큰 문제가 없지만^[4] 왜곡되거나, 실제 카메라로부터 획득된 스테레오 영상을 사용하는 경우에는 합성오류가 발생하는 문제점을 안고 있다^[5]. 즉, 자연환경으로부터 스테레오 카메라를 사용하여 획득된 실사영상의 경우 스테레오 카메라간의 기하학적 구조차이 및 오정합 등으로 입력된 스테레오 영상은 왜곡되고 결과적으로 이를 이용하여 합성된 중간영상의 경우 합성오류와 함께 PSNR의 저하로 불완전한 3D 디스플레이의 결과가 나타나게 된다.

최근, 이러한 기존 중간영상합성 기법의 한계를 극복하기 위한 새로운 접근방식으로 비교정된(uncalibrated) 스테레오 영상에 대한 영상 보정(rectification) 및 교정 기법이 제시되고 있다. Hartley는 서로 다른 시점의 스테레오 영상에 대한 수학적 보정기법을 제시하고^[6], Loop와 Zhang은 스테레오 영상의 보정기법으로 컴퓨터 비전에서의 상종관계(homography)에 따라 주변 대응점들과의 상호 거리 정보를 이용하여 잘못된 후보 점들을 제거하는 기법을 제시하였으나^[7], 이는 큰 오차를 가지는 하나의 후보 점에 대해서 주변의 대응 후보 점들이 모두 제거된다는 비안정적인 결과를 나타내고 있다. 또한, Papadimitriou와 Dennis는 수렴 기하학에 의한 스테레오 영상보정 알고리즘을 제시하였고^[8], Ayache와 Hansen은 교정상태에서 3시점의 영상에 대한 보정 알고리즘을 제안하였다^[9]. 그러나 이러한 기법들은 카메라의 구조를 복잡한 수학적 모

델로 분석해야하는 어려움이 있으며, 이에 따라 교정된 카메라를 통해 획득된 스테레오 영상에 국한되는 한계점을 가지고 있다.

따라서, 본 논문에서는 비교정된 스테레오 영상에 대해서 에피폴라 구조조건(epipolar constrain), 특징점 추출에 따른 보정정보 및 이에 따른 변환 파라메타 등을 이용하여 영상을 효과적으로 보정할 수 있는 새로운 스테레오 영상 보정 기법을 제안하고자 한다. 또한, 자연환경에서 획득된 비교정 상태의 스테레오 영상에 본 논문에서 제시된 스테레오 영상 보정 기법이 적용된 새로운 중간시점 영상합성 기법을 제시하고자 한다. 즉, 새로이 제시된 중간시점 영상합성 기법에서는 먼저, 좌, 우 스테레오 영상의 각 화소 값들에 대한 유사도 및 모서리 검출을 통해 특징점을 추출한 다음, 이들 특징점을 이용하여 스테레오 영상간의 움직임 벡터와 에피폴라 선을 검출하게 된다. 그리고, 스테레오 영상간의 수평선을 일치시킴으로써 좌, 우 스테레오 영상을 보정하고 최종적으로 적응적 변위추정 기법을 이용하여 중간시점 영상을 합성하게 된다. 그리고 제안된 기법의 효용성을 검증하기 위한 실험으로 CCETT (Centre Commund'Etudes de Telecommunic ations et de Telediffusion)의 'Man' 영상, 스테레오 카메라로 촬영된 '사람' 및 '자동차' 영상에 대해서 교정 전의 영상과 본 논문에서 새로이 제시된 기법으로 교정된 후의 영상을 이용하여 각각 합성된 중간시점 합성결과를 PSNR 관점에서 비교 분석하고 이를 통해 본 논문에서 제안된 스테레오 영상 보정기법에 기반한 새로운 중간시점 합성 기법의 실질적 응용 가능성을 제시하고자 한다.

II. 제안된 스테레오 영상보정기법

스테레오 영상 보정이란 좌, 우 영상간의 스테레오 영상 정합 및 합성을 위한 선행 작업으로 영상의 관찰 대상인 특징점들을 하나의 선으로 연결하여 이를 수평적으로 맞추어주는 선 정합(line matching) 과정을 말한다. 따라서 스테레오 영상보정을 위해서는 선 정합에 따른 정확한 에피폴라 선의 검출과 이에 따른 정규화된 카메라 좌표계의 수학적 정립이 필수적이다.

그림 1은 스테레오 영상 보정 및 중간시점 영상합성 알고리즘의 흐름도와 각 알고리즘 수행에 필요한 파라메타 변환과정을 각각 나타낸 것이다.

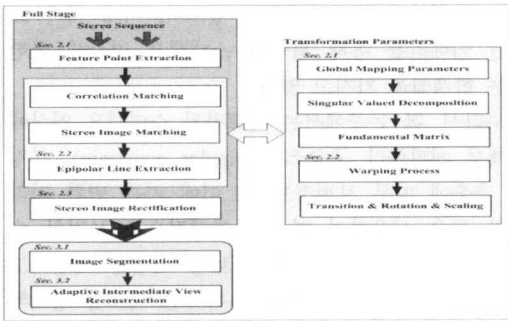


그림 1. 스테레오 영상보정 및 중간시점 영상합성 알고리즘의 흐름도

즉, 좌, 우 스테레오 영상의 특징점을 검출한 다음, 기준 영상과 비교하여 좌, 우 영상의 대응점을 기준으로 상대적인 움직임 형태를 판단하게 된다. 또한, 일반적으로 항상 정확한 대응점을 구하는 것은 불가능하기 때문에 일정한 변환 계수를 이용하여 움직임을 추정한 후, 각 프레임 영상을 최종 스테레오 영상에 맞도록 변환할 수 있는 와핑(warping) 파라미터의 검출을 통해 좌, 우 영상 간에 존재하는 에피폴라 선을 수평적으로 정합함으로써 비교정된 스테레오 입력영상을 보정하게 된다.

1 스테레오 영상 정합을 위한 좌, 우 영상의 특징점 검출

스테레오 영상 정합이란 좌, 우 영상간의 유사성을 측정하여 영상 내에 존재하는 물체의 기하학적인 특성을 찾아낸 후, 일치되는 한 점을 검출해 내는 과정을 말한다.

본 논문에서는 스테레오 영상 정합의 전처리 과정인 일치된 대응점 탐색기법으로 Harris의 모서리 탐색 과정과 상관성 매칭 기법을 사용하였으며^[10], 식 (1)과 같이 영상의 휘도 차이에 대한 자기상관 함수를 이용하여 임의 영상의 관심영역에 대한 모서리를 검출하였다.

$$\delta I(\delta x, \delta y) = \sum_{ij \in patch} (I_1(i + \delta x, j + \delta y) - I_1(i, j))^2$$

$$(i = 1, 2, 3 \dots, j = 1, 2, 3 \dots) \quad (1)$$

식 (1)은 스테레오 영상 중 임의의 영상 I_1 에 대해 x 방향과 y 방향으로 각각 한 픽셀씩 이동해 가면서 모서리를 찾아내는 방식으로, 식 (1)을 Taylor 급수로 전개하면 식 (2)와 같이 주어지고 식 (2)에

서 N 은 식 (3)과 같이 정의된다.

$$\delta I(\delta x, \delta y) = (\delta x, \delta y) N \left(\frac{\delta x}{\delta y} \right) \quad (2)$$

$$N(x, y) = \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} \quad (3)$$

식 (3)에서 N 의 x, y 는 영상의 관심영역에 대한 두 가지 고유값(eigen value)으로, 한 영상의 그 자신에 대한 픽셀의 차이를 고유값으로 나타내는 회전 불변의 성질을 포함하고 있으며, 픽셀의 변화량을 이용하여 특징위치의 모서리를 찾는데 이용된다^[11].

따라서, 어느 한 영상의 관심영역에 대한 픽셀 차이를 고유치로 나타낸 뒤, 픽셀의 변화량을 이용한 특정 위치의 모서리에 대한 세기(strength) 신호를 표현하면 식 (4)와 같이 나타낼 수 있다.

$$\Phi(x, y) = | N(x, y) | - k Tr^2 N(x, y) \quad (4)$$

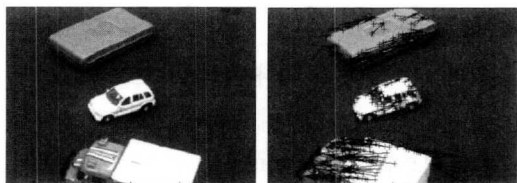
여기서, k 는 카메라로 투영된 2차원 영상의 일부 분인 패치(patch) 크기를 결정짓는 행렬 비례상수 값으로 0.04로 주어지는 데 이 값은 Harris에 의해 실험적으로 증명된 바 있다^[11]. 그리고 Tr 은 행렬을 이루고 있는 대각 요소의 합(trace of matrix)을 의미하며, 검출된 모서리 크기는 2차원 영상의 조합인 $\Phi(x, y)$ 의 패치 크기로 결정된다.

일반적으로, 스테레오 영상보정을 위해서는 우선적으로 좌, 우 영상의 특징점을 검출하여 기준 영상과 비교한 뒤, 기준 영상에 대하여 상대적으로 어떠한 움직임이 있었는지를 계산하는 대응점 추정과정이 필요하다. 따라서 본 논문에서는 대응점 추정을 위한 상관기법으로 식 (4)를 통해 검출된 모서리 탐색 영역을 기반으로 식 (5)와 같은 상관함수를 통해 좌, 우 영상의 특징점을 비교하여 초기의 대응되는 후보 점을 검출하게 된다^[8].

$$C = \sum_{ij \in patch} (I_2(i, j) - I_1(i, j))^2 \quad (5)$$

이와 같은 과정으로 대응점을 검출한 뒤 그림 2

와 같이 좌, 우 영상 각각의 특징점을 한 영상에 나타내면 특징점들 사이의 변화정도를 검출할 수 있게 된다.



(a) 좌 영상 (b) 우 영상
그림 2. 좌, 우 영상간의 특징점 변위 검출

2 에피폴라 선 검출

스테레오 영상 정합을 위한 검출된 좌, 우 영상의 특징점들은 그 변화정도가 각각 틀리기 때문에, 각 영상에서 검출된 특징점들을 구체화하기 위해서는 스테레오 영상에 존재하는 에피폴라 선을 검출하는 것이 필요하다.

그림 3은 대응되는 한점을 기준으로 투영된 스테레오 영상간의 에피폴라 기하학의 관계를 나타낸 것이다.

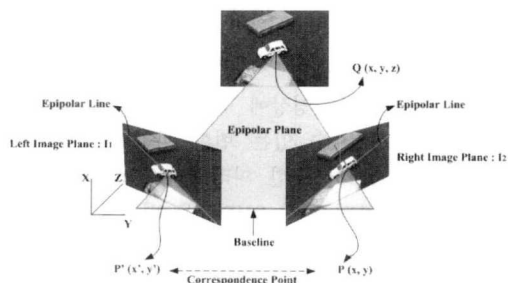


그림 3. 스테레오 영상간의 에피폴라 기하학

그림 3에서 알 수 있듯이 스테레오 카메라를 통해 한 점 Q의 좌표를 나타내는 좌, 우 영상의 평면을 같은 차원의 평평한 면에 투영시키기 위해서는 에피폴라 선을 이용하여 스테레오 영상을 평행할 수 있도록 만들어야 한다. 또한, 좌, 우 영상에서 검출된 특징점들을 구체화하기 위한 에피폴라 선의 검출은 에피폴라 평면의 기하학 정보를 나타내는 기본 행렬(fundamental matrix)을 이용하여 각 영상의 특징점인 에피폴을 찾음으로써 검출할 수 있으며, 기본 행렬은 일반적인 정방 행렬(square matrix)에서 고유벡터(eigen vector)를 산출하는 SVD(singular valued decomposition) 이론을 적용하여 검출하게 된다^[8].

식 (6)은 좌, 우 영상인 I_1 과 I_2 의 일치점을 이용하여 SVD 형태인 $n \times 9$ 행렬로 나타낸 것이다.

$$D_i = [uu', vv', ww', uv', vv', ww', uw', vw', ww'] \quad (6)$$

즉, 좌, 우 영상간의 동차(homogeneous) 좌표로 표시된 어느 한 일치점 (x_i, x'_i) 에 대해, $x_i = [u, v, w]^T$ 및 $x'_i = [u', v', w']^T$ 일 때 D 의 i 번째 행인 D_i 를 계산한 것이며, 이를 통해 SVD의 형태는 식 (7)과 나타낼 수 있다^[8].

$$D = UWV^T \quad (7)$$

또한, 식 (7)에서 W 의 마지막 값이 최소 특이값(singular value)을 갖는다고 하면 기본행렬 F 는 V 의 마지막 열이 되고, 영상 I_1 에서의 에피폴 e 는 $F \cdot e = 0$ 인 단위벡터로 둘 수 있게 된다. 따라서 이러한 단위벡터는 행렬 $F^T F$ 의 최소 고유치에 해당하는 단위 고유벡터를 계산함으로써 검출할 수 있다^[8]. 따라서 기본행렬의 해들은 각 영상에서 나타나는 특징점들의 상태를 반영하고 있으므로 원 영상에서 나타나는 특징점들의 2차원 패턴을 보여주며, 이상의 과정을 통해 검출된 특징점들은 이후 전개될 스테레오 영상보정의 정보로 이용된다.

3 스테레오 영상의 보정

2에서 좌, 우 영상간의 수학적 모델링의 결과로 얻어진 좌, 우 영상의 기본 행렬은 스테레오 영상의 특징점에 대한 정보만을 내포하고 있기 때문에 특징점들을 통해 검출된 좌, 우 영상의 에피폴라 선을 수평으로 맞추는 보정작업을 수행할 수 있는 함수가 요구된다. 따라서, 이를 위해서는 영상간의 전이, 회전, 확대 및 축소와 같은 기하학적 변환을 통해 기준 영상에 맞도록 변형할 수 있는 변환 파라메타가 필요하며, 투영된 2차원 상에 있는 영상의 한 점을 3차원으로 옮긴 후에 그 영상에 적용된 전역 움직임의 반대 방향 움직임으로 영상을 역변환한 후, 다시 2차원으로 투영시키는 과정이 필요하다.

그림 4는 이러한 과정을 해석하기 위해 카메라 초점(Z축)이 동일하다는 가정 하에 설정된 스테레오

영상간의 2차원 위치 관계를 나타낸 것이다. 즉, 영상의 한 가운데 점을 영상의 중심점이라 할 때, 두 영상의 중심점이 직각 좌표계에서 z좌표가 초점거리 위치에 있는 경우 그 위치 관계를 그림으로 나타낸 것이다.

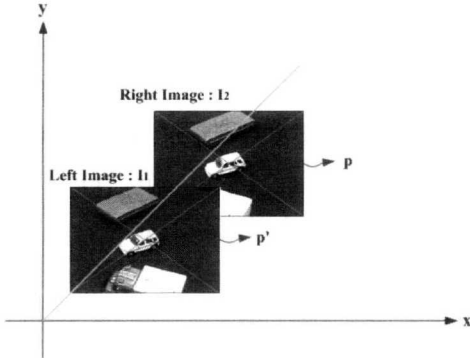


그림 4. 스테레오 영상간의 2차원 위치 관계

즉, 좌측 영상 I_1 을 기준 영상으로, 우측 영상 I_2 를 비교 대상 영상으로 보면, I_2 의 한 점 p 에 대응하는 I_1 점을 p' 라 했을 때, 전이, 회전 변환을 거친 p 의 변환 p' 은 식 (8)과 같이 나타낼 수 있다.

$$P' = TRP \tag{8}$$

이때 T 는 전이 변환 행렬을, R 은 회전 변환 행렬을 각각 나타내며, 이를 통해 얻어진 p' 의 좌표를 초점 거리 변환을 통해 확대 혹은 축소 역 변환을 거친 후의 좌표는 p'' 와 같아야 한다. 따라서,

의 좌표를 p'' 이라 하면, p'' 과 p 의 관계는 식 (9)과 같이 나타낼 수 있다.

$$\begin{aligned} p''_x &= \lambda_1 \frac{p'_x}{p'_z} \\ p''_y &= \lambda_1 \frac{p'_y}{p'_z} \end{aligned} \tag{9}$$

식 (9)에서 P'_x, P'_y, P'_z 각각에 식 (8)의 결과를 대입하여 이를 행렬식으로 정리하면 식 (10)과 같은 형태가 된다.

$$\begin{bmatrix} p''_x \\ p''_y \\ 1 \end{bmatrix} = \begin{bmatrix} k_1 & k_2 & k_3 \\ k_4 & k_5 & k_6 \\ k_7 & k_8 & k_9 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ 1 \end{bmatrix} \tag{10}$$

각각의 k_i 는 변환 파라미터들의 관계식으로 구성되어 있으며, 모든 k_i 를 구하게 되면 이 관계식을 통해 역으로 각 파라미터들을 산출할 수 있으며 이를 통한 변환 행렬은 식 (11)과 같이 표현될 수 있다.

$$\begin{bmatrix} k_1 k_2 k_3 \\ k_4 k_5 k_6 \\ k_7 k_8 k_9 \end{bmatrix} = \begin{bmatrix} \sum x_j^2 & \sum x_j y_j & \sum x_j \\ \sum x_j y_j & \sum y_j^2 & \sum y_j \\ \sum x_j & \sum y_j & n \end{bmatrix}^{-1} \begin{bmatrix} \sum x_j p'_j & \sum x_j y'_j & \sum x_j \\ \sum y_j p'_j & \sum y_j y'_j & \sum y_j \\ \sum x'_j & \sum y'_j & n \end{bmatrix} \tag{11}$$

식 (11)과 같이 구해진 행렬의 각 요소 k_i 와 전역 움직임의 파라미터들 간의 관계식으로 전역 대응 파라미터를 구할 수 있으며, 변환 행렬 $[]$ 는 스테레오 영상의 원 영상에 특징점들의 위치의 벡터들을 변화 시킬 수 있는 변환 파라메타가 되는 것이다. 식 (11)에서 구한 파라메타는 좌, 우 영상의 특징점으로 부터 구한 SVD값을 변환할 수 있는 변환식이 되며, 이를 통해 x축이 되는 에피폴의 수평선 값에 평행하게 하고, x 와 x' 간의 거리를 일정하게 유지시켜주게 되면 최종적으로 영상보정이 이루어지게 된다.

III. 보정된 스테레오 영상 기반의 중간시점 영상합성

중간시점의 영상을 합성하는 방법에는 크게 외삽법(extrapolation)과 내삽법(interpolation)^[10-11]이 있는데 이 중, 가중 평균값으로 내삽하는 기법이 외삽법의 경우보다 자연스러운 중간시점의 영상을 합성할 수 있기 때문에 본 논문에서는 중간시점에 해당하는 영상값은 가중평균값으로 내삽하였다.

그림 5는 보정된 스테레오 영상으로부터 검출된 특징값의 크기에 따라 정합될 정합창의 크기를 적응적으로 선택하는 알고리즘의 흐름도를 나타낸 것

이다^[4].

으로 전, 후 배경이 동시에 존재하고 있다.

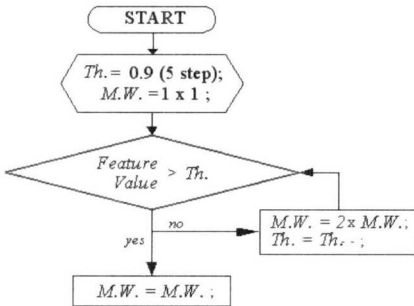


그림 5. 적응적 정합 알고리즘 흐름도

식 (12)는 그림 5와 같이 비용함수를 이용하여 추출된 특징값과 임계값의 비교를 통해 5가지의 정합창의 크기를 선택하는 수식을 나타낸 것이다.

$$\frac{1}{N_x N_y} \sum_{i=1}^N \sum_{j=1}^N |I_L(i, j) - I_R(i + d, j)|^2 > Th_n(Step[n]) \Rightarrow 2^{5-n} \quad (12)$$

그리고 식 (13)은 시점의 위치에 따라 가중 평균값으로 내삽하는 경우를 나타낸 식이다^[4].

$$I_p(i, j) = (1 - \alpha)I_L(i - d_{ij}(i, j), j) + \alpha I_R(i - d_{ij}(i, j), j) \quad (13)$$

여기서, d_{ij} 는 탐색범위 내의 변위값을, I_p 는 좌, 우 영상의 변위가 교차되는 지점에 합성된 중간시점 영상의 위치를 각각 나타낸 것이다. 따라서 적응적 정합창 선택 알고리즘은 스테레오 입력영상에서 추출된 특징값의 크기를 설정된 임계값과 비교, 판단하여 적응적으로 적절한 정합창을 설정하게 된다.

IV. 실험 및 결과분석

본 실험에서는 640×480 픽셀크기의 CCETT의 좌, 우 스테레오 영상인 그림 6 (a), (b)의 ‘Man’ 영상과 카메라로 실제 촬영한 좌, 우 스테레오 영상인 그림 6 (c), (d)의 ‘얼굴’, 그리고 그림 6 (e), (f)의 ‘자동차’ 스테레오 영상을 이용하였다. 즉, 그림 6 (c), (d)는 단순 배경의 얼굴 영상으로 CCETT의 ‘Man’ 영상에 비해 배경 그림자(silhouette)가 있는 차이점이 있고, 그림 6 (e), (f)는 스테레오 물체추적 시스템에서 주시각이 제어된 이후의 좌, 우 영상

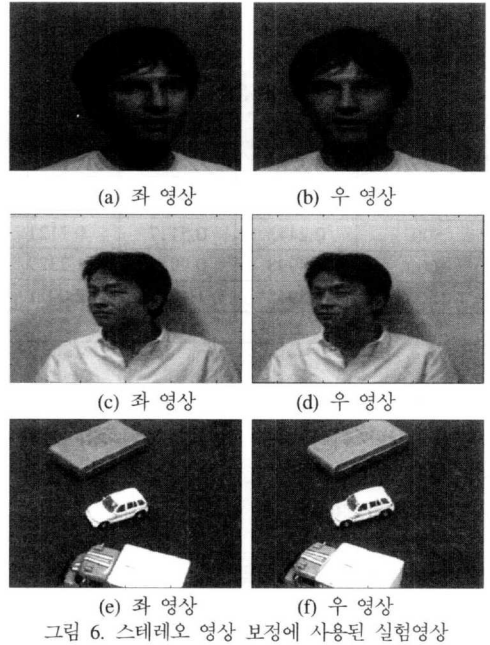


그림 6. 스테레오 영상 보정에 사용된 실험영상

1 스테레오 영상의 특징점 추출

그림 7은 기존의 모서리 검출방식 중 원 영상에 자기상관 기법을 적용한 Birchfield/Tomasi 방식^[11]을 이용하여 실험에 적용한 각 영상들에 대해 200개의 특징점을 검출한 결과를 나타낸 것이다.



그림 7. 실험영상을 이용한 특징점 검출결과

표 1은 본 실험에 사용된 각기 다른 조건의 영상들에 대해 특징점의 증가에 따른 검출 시간을 비교한 것이다.

표 1. 특징점 검출시간 (단위: 초)

특징점 수 \ 영상	Man 영상	얼굴 영상	실험실 영상
200	0.2002	0.3557	0.4527
500	0.2443	0.5117	0.7121
1000	0.8998	0.9917	1.2329
1500	1.5945	1.4821	2.0021
2000	2.1806	2.4897	3.2002

표 1에서 특징점의 개수를 늘리면 검출 소요시간은 증가되지만 이는 더욱 확실한 에피폴라 선의 검출을 위한 선행 작업으로 입력영상의 복잡성에 따라 적응적으로 변화될 수 있다.

2 스테레오 영상의 정합

그림 8은 스테레오 영상 간의 특징점을 검출한 다음 각각의 대응점에 대해 변위위치의 움직임 벡터를 우 영상에 나타낸 것이다.



그림 8. 스테레오 영상의 움직임 벡터 검출

즉, 스테레오 영상에서 검출된 특징점을 변위정보로 이용하여 좌, 우 영상간의 움직임 벡터로 표현한 것이다. 또한, 우 영상에 표현된 선은 보정정보를

위한 대응점의 거리를 보여줌과 동시에 스테레오 영상의 정합에 대한 후보 특징점의 경로를 제시해 주며 이는 최소자승법을 통해 X 좌표에 이르는 최단거리인 수평선을 찾는 후보거리가 된다.

그림 8 (e), (f) '실험실' 영상의 경우 좌, 우 카메라의 위치가 인간 시각 시스템과 현격히 다르기에 따라 수직 시차가 크게 나타나기 때문에 스테레오 정합을 위한 변위 벡터에 큰 차이가 나타남을 알 수 있다.

3 스테레오 영상의 보정 정보를 위한 에피폴라 선 검출

그림 9는 2.1.2에서 검출한 변위벡터와 SVD를 이용하여 각 영상에 대해 200개의 특징점을 부여하여 검출된 스테레오 영상의 에피폴라 선을 각각 나타낸 것이다. 그러나, 그림 9 (a), (b)의 경우, 배경과 사람간의 휘도값이 비슷하면서 나타날 수 있는 특징점 검출 오류로 인해 에피폴라 선의 검출에도 수직 오류가 나타남을 알 수 있다.

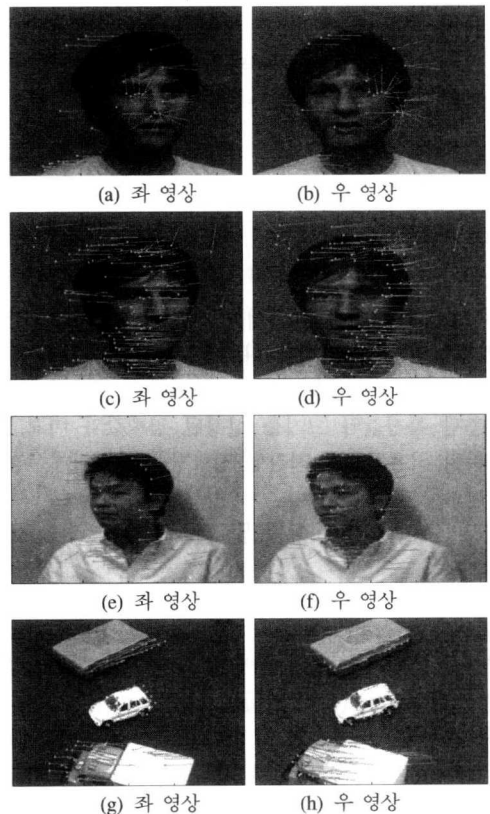


그림 9. 에피폴라선 검출

따라서, 그림 9(c), (d)와 같이 특징점의 수를 500 개로 증가시킴으로써 보다 정확한 에피폴라 선의 검출이 가능하였다. 이는 표 1에서도 언급을 했듯이, 특징점의 수를 늘림으로써 계산시간은 더 걸리나 보다 더 정확한 특징점 검출이 가능함을 알 수 있다.

그림 10(a), (b), (c)는 비 교정된 스테레오 영상 간의 보정 정보를 추정하기 위해 2.1절에서 얻은 움직임 변위벡터와 2.2절에서 검출된 에피폴라 선을 각각 우 영상에 표현한 것이다. 또한, 그림 10(d), (e), (f)는 그림 10(a), (b), (c)에서 획득한 영상의 보정 정보를 3차원 공간상에 각각 나타낸 것이다. 즉, 그림에서 x, y축은 스테레오 카메라를 통해 획득된 영상좌표를, z축은 스테레오 카메라의 초점거리를 각각 나타낸다. 그리고, 붉은색으로 표현된 변위정보는 변환 파라미터를 이용하여 추정된 움직임 벡터의 변위를, 연두색은 두 영상의 에피폴라 선의 의미를 나타낸다. 따라서, 붉은색 변위정보들은 연두색으로 표현된 에피폴라 선을 기준으로 두 영상간의 존재하는 에피폴라 선을 만들어내며, 이는 두 영상간의 최종적인 보정 정보로 이용될 수 있다. 그림 10(d), (e), (f)에서 붉은색으로 표현된 보정정보는 변위정보가 존재하는 스테레오 영상의 특징 정보를 나타내는 것으로, 그 특징점의 수가 많을수록 유사성이 높음을 의미하고, 특징점의 간격이 벌어질 수록 각 변위에 대한 보정의 필요성이 크다는 것을 의미한다.

따라서 그림 10(d), (e)의 경우, 붉은색으로 표현된 흩어진 특징 정보의 수렴 영역이 연두색의 에피폴라 선을 중심으로 존재하는 것을 알 수 있으며, 이는 움직임 벡터간의 최소 이동벡터에 대한 추정 정보가 명확하다는 것을 알 수 있다. 또한, 그림 10(f)의 경우는, 수렴 영역이 중앙에 위치한 자동차에 대해서만 존재하고, 전, 후 배경에 존재하는 물체들에 대해서는 존재하지 않는 것으로 보아, 현재 중앙에 위치한 자동차를 중심으로 주시각이 제어되어 수직, 수평 시차가 거의 존재하지 않는 것을 알 수 있다.

4 중간시점의 영상합성

그림 13은 4.4에서 획득한 보정 정보와 변환 파라미터를 이용하여 최종적으로 중간시점 영상합성 기법에 적용한 후의 결과영상을 나타낸 것이다. 즉, 그림 13(a), (c)는 비교정된 상태의 영상을 기존의 블럭기반의 시차추정기법을 이용하여 합성된 중간시점 영상을 나타낸 것이고, 그림 13 (b), (d)는 본 논문에서 새로이 제시한 기법으로 보정된 영상을 블럭기반의 시차추정 기법을 이용하여 합성된 중간시점 영상을 나타낸 것이다.

그림 13 (a), (b)의 'Man' 표준영상의 경우는 보정 전, 후의 영상으로 각각 합성된 중간시점 영상결과에 별 차이가 없음을 알 수 있다. 그러나 그림 13 (c), (d)의 '얼굴' 영상 및 (e), (f)의 '자동차' 영상의 경우는 보정 전, 후의 중간시점 영상합성 결과에 많은 차이가 있음을 볼 수 있다.

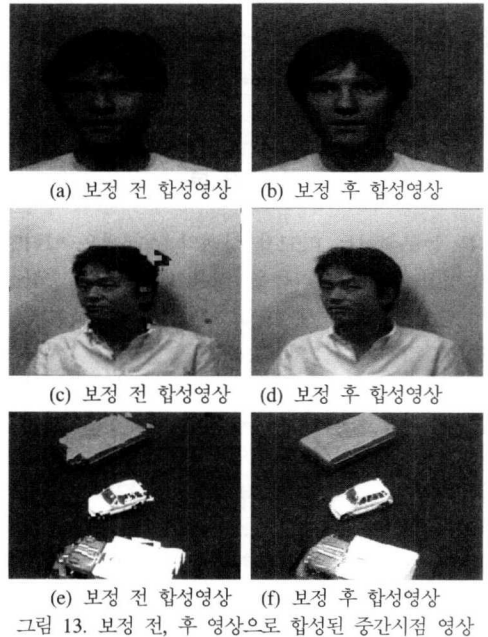
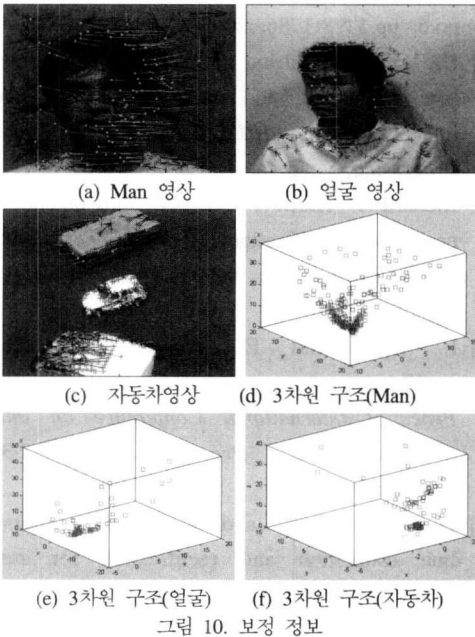


표 2는 'Man', '얼굴' 및 '자동차' 영상에 대해 원 영상을 기준으로 보정 전, 후의 영상으로 각각 합성된 중간시점 영상의 PSNR을 각각 나타낸 것이다. 표 2에서 보면, CCETT의 표준 영상인 'Man'의 경우에도 영상보정 후에 PSNR이 3.6dB 증가됨을 알 수 있으며, 카메라로 실제 촬영된 스테레오 '얼굴' 영상의 경우는 2.59dB의 PSNR 개선이 이루어짐을 알 수 있다. 또한, '자동차' 영상의 경우도 영상보정 후에 1.47dB의 PSNR이 개선되었는데, 이는 중앙에 위치한 자동차를 중심으로 주시각이 제어된 입력영상이라 할지라도 전, 후 배경에 존재하는 물체에 대해서도 영상보정으로 오류없는 중간시점 영상합성이 이루어짐을 알 수 있다.

표 2. PSNR 결과 비교

영상 보정 여부	PSNR (dB)		
	'Man' 영상	'얼굴' 영상	'자동차' 영상
보정 전	26.25	18.18	16.47
보정 후	29.85	20.77	17.94

따라서, 이상의 실험결과, 본 논문에서 새로이 제시된 영상보정 기법으로 교정된 영상으로 합성된 중간시점 영상의 경우 PSNR이 평균 2.55dB 개선됨이 분석됨으로써 본 논문에서 새로이 제시한 스테레오 영상 보정 알고리즘 기반의 중간시점 영상 합성 기법의 실질적 효용성과 이를 이용한 보다 효과적인 다시점 3D 디스플레이 시스템의 구현 가능성을 제시하였다.

V. 결론

본 논문에서는 비교정 상태의 스테레오 입력영상에 영상보정 알고리즘을 적용한 새로운 중간시점 영상합성 기법을 제시하고 그 성능을 분석하였다. 제시된 방법에서는 먼저, 좌, 우 스테레오 영상의 각 화소 값들에 대한 유사도 및 모서리 검출을 통해 특징점을 추출한 다음, 이들 특징점을 이용하여 스테레오 영상간의 움직임 벡터와 에피폴라 선을 검출하였다. 그리고, 스테레오 영상간의 수평선을 일치시킴으로써 좌, 우 스테레오 영상을 보정하고 최종적으로 적응적 변위추정 기법을 이용하여 최적화된 중간시점 영상을 합성하였다. CCETT의 'Man' 영상과 스테레오 카메라를 사용하여 촬영한 '사람'

및 '자동차' 영상을 사용한 중간영상 합성 실험결과 본 논문에서 제안된 보정기법으로 교정된 스테레오 영상의 경우가 비교정 상태에 비해 'Man' 영상은 3.6dB, '사람' 및 '자동차' 영상은 2.59dB, 1.47dB의 PSNR이 각각 개선되어 평균적으로 2.55dB가 개선됨이 분석됨으로써 본 논문에서 새로이 제시한 스테레오 영상 보정 알고리즘 기반의 중간시점 영상합성 기법의 실질적 효용성과 이를 이용한 보다 효과적인 다시점 3D 디스플레이 시스템의 실질적 응용 가능성을 제시하였다.

참고 문헌

- [1] A. Redert, E. Hendriks and J. Biemond, "Correspondence estimation in image pairs", *IEEE Signal Processing Magazine*, vol.16, no.3, pp.29-46, 1999.
- [2] T. Emanuele and V. Alessandro, *Introductory Techniques For 3-D Computer Vision*, Pearson Higher Education, 1998.
- [3] O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, 1993.
- [4] K. H. Bae, J. J Kim, E. S. Kim, "New disparity estimation scheme based on adaptive matching window for intermediate view reconstruction", *Optical Engineering*, vol.42, no.6, pp.1778-1786, 2003.
- [5] D. V. Papadimitriou, T. J. Dennis, "Epipolar line estimation and rectification for stereo image pairs", *IEEE Trans. on Image Processing*, vol.5, no.4, 1996.
- [6] Richard I. Hartley, "Theory and practice of projective rectification", *International Journal of Computer Vision*, vol.35, no.22, pp.115-127, 1999.
- [7] F. Isgro and E. Trucco, "On robust rectification for uncalibrated images", *IEEE Proc. of International Conference on Image Analysis and Processing*, pp.297-302, September 1999.
- [8] D. Papadimitriou and T. Dennis. "Epipolar line estimation and rectification for stereo image pairs", *IEEE Trans. on Image Processing*, vol.5, no.4, pp.672-676, 1996.

[9] N. Ayache and C. Hansen, "Rectification of images for binocular and trinocular stereovision", *Proc. of International Conference on Pattern Recognition*, vol.1, pp.11-16, November 1988.

[10] http://www.ri.cmu.edu/people/kanade_takeo.html

[11] A. Mancini, *Disparity estimation and intermediate view reconstruction for noble applications in stereoscopic video*, Master thesis, McGill University, 1998.

[12] J. S. Lee, J. H. Ko, E. S. Kim, "Real-time stereo object tracking system by using block matching algorithm and optical binary phase extraction joint transform correlator", *Optics Communication*, vol.191, pp.191-202, 2001

[13] A. Redert, E. Hendriks and J. Biemond, "Correspondence estimation in image pairs", *Signal Processing Magazine*, vol. 16, no.3, pp.29-46, May 1999

[14] J. Shao and C. Fraser, "Rectification and matching of trinocular imagery", *Geomatics Research Australasia*, vol.71, pp.73-86,1999.

[15] A. Fusiello, E. Trucco, and A. Verri. "A compact algorithm for rectification of stereo pairs", *Machine Vision Applications*, vol.12, no.1, pp.16-22, 2000.

[16]http://www.vision.caltech.edu/bouguetj/calib_doc/index.html

[17] S. Avidan, "Trajectory triangulation: 3D reconstruction of moving points from a monocular image sequence", *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol.22, no.4, 2000.

[18] Z. Chen, C. Wu, H. T. Tsui, "A new image rectification algorithm", *Pattern Recognition Letters*, vol.24, no.1-3, pp.251-260, 2003.

[19] <http://research.microsoft.com/~zhang/>

[20] D. J Coombs and C. M. Brown, "Cooperative gaze holding in binocular vision", *IEEE Control System*, pp.24~33, 1991.

[21] V. S. Grinberg, G. W. Pondar and M. W. Sigel, "Geometry of binocular imaging", *Proc. of SPIE*, vol.2177, pp.56~65, 1994.

[22] E. Izquierdo, "Stereo matching for enhanced

telepresence in three-dimensional video communications", *IEEE Trans. on Circuits and Systems for Video Technology*, vol.7, no.4, pp.629~643, 1997.

[23] C. H. Yang, "Geometric models in stereoscopic video", *Rapport technique de l'INRS-Tel'ecomunications*, vol.9, pp.5~12, 1995.

박 창 주(Chang-ju Park)

정회원



2002년 2월 : 호남대학교
전파공학과 학사
2002년 3월 : 광운대학교
전자공학과 석사과정

<관심분야> 3D 영상처리, 영상통신

고 정 환(Jung-Hwan Ko)

정회원



1999년 2월 : 광운대학교
제어계측공학과 학사
2001년 8월 : 광운대학교
전자공학과 석사
2003년 8월 : 광운대학교
전자공학과 박사수료
2004년 3월~현재 :
동아방송대학 정보통신
계열 겸임교수

<관심분야> 3D 로봇비전, 스테레오 카메라, 영상통신

김 은 수(Eun-Soo Kim)

정회원

광운대학교 국가지정 3차원 영상 미디어 연구실
(<http://3dmedia.kwangwoon.ac.kr>)
한국통신학회지 Vol.28, No. 4C 참조