

특집논문-01-06-2-05

아웃라이어 배제에 기초한 일반화된 파노라마 영상 재구성

서종열*, 박종현**, 강문기**

Generalized Panoramic Scene Reconstruction from Video Sequences Based on Outlier Rejection

Jong-Yeul Suh*, Jong Hyun Park** and Moon Gi Kang**

요 약

본 논문에서는 동영상 정보를 파노라마로 합성하는 자동화된 알고리즘을 제안한다. 제안한 카메라 모델은 일반화된 파노라마 영상의 특성을 반영하며, 기존의 방법과는 달리 영상의 확대 및 축소를 고려하여 보다 개선된 파노라마 합성 영상을 제공한다. 또한 통계적 특성을 이용하여 움직이는 물체 또는 기타 여러 잡음(noise)의 영향을 배제하는 전역 움직임 추정 기법을 사용하여 기존 방법의 문제점을 보완한다. 이를 통해 파노라마 영상으로부터 움직이는 물체 또는 잡음을 제거하여 더 깨끗하고 완벽한 배경 영상을 얻을 수 있다. 제안한 방법은 전역 움직임 및 전체 파노라마 합성 과정에서 수동 작업을 배제하여 파노라마 합성 과정을 일반화시켰다고 할 수 있다. 성능 평가를 위해 실제 방송에서 사용되는 축구 경기 영상에 제안한 알고리즘을 적용하고 기존의 알고리즘과 비교를 통해 그 우수성을 검증한다.

Abstract

In this paper, we propose a new practical motion model that can exploit the general properties of camera motion in constructing a panorama, accounting for panning, tilting, and even the change in focal length of the camera. We also present an efficient algorithm to handle moving objects or noise in the scene based on outliers rejection. Spatial and temporal statistical properties of motion field are exploited to detect the outliers. The proposed algorithm removes moving objects or noise from the panoramic image so that more clear and complete view of the background image can be obtained. This method does not require assumptions or a priori knowledge of the scene. The entire process is fully automatic as this method does not require any manual correction in the process of constructing a panorama. The proposed algorithm is tested on the broadcasting images of soccer games. Our simulation result shows that this method is superior to conventional image mosaicing algorithms.

I. 서 론

영상 모자이크, 혹은 파노라마는 다양한 시점에서 획득한 여러 장의 영상들로 단일 파노라마 영상을 만들어내는 기법을 말한다. 우리는 이 기법을 이용하여 영상의 크기

나 범위에 관계없이 영상을 재구성할 수 있으며, 영상 획득 장치의 화각이라는 물리적 한계를 극복할 수 있다. 현재 이러한 기법은 가격 면에서 효율적인 가상 환경 구성^{[1][2][3]}, 초해상도 영상 구성^[4], 그리고 위성 사진이나 항공 사진의 관심 있는 영역에 대한 넓은 지역 관찰^{[5][6][7]} 등에 응용되고 있다. 또한 축구경기 방송 시스템에서도 이 기술을 통해 전체 구장의 광경을 재구성하여 시청자에게 더 넓은 시야를 제공할 수 있다^[8].

지금까지 다양한 목적에 적용할 수 있는 파노라마 합성

* LG전자
LG Electronics

** 연세대학교 전기전자공학과
Dept. of Electrical and Electronic Eng., Yonsei Univ.

방법들이 제안되었다. 특징 기반 위치 지정(feature-based registration) 방법을 사용한 영상 모자이크 알고리즘^{[6][9]}은 영상에 대한 사전 정보가 필요하고, 특징(feature)을 정확하게 맞추기 위해서는 어느 정도 사용자의 개입이 필요하다. 이를 통해서 시스템의 성능을 보다 향상시킬 수는 있지만^{[10][11]}, 그 보다는 특징을 사용하지 않는 위치 지정(featureless registration) 방법을 사용하여 파노라마 합성을 자동화하는 것이 바람직하다. 기존의 알고리즘^{[1][2][11][12]}은 움직이는 물체나 잡음 등 일반적인 상황에 대한 해결책을 제시하지 못했고, 카메라와 실제 피사체간의 기하학적 관계를 적절히 묘사하지 못했다. 또한 기존의 방법은 정적인 광경의 경우만을 고려했기 때문에, 다수의 이동물체가 있는 영상에서는 문제가 있다. 그리고 실제적인 경우 전역 또는 지역에서 밝기의 변화가 일어나는데 기존의 방법들은 이러한 밝기 변화에 자동적으로 대응하지 못한다. 한편 기존의 방법들에서는 카메라의 움직임에 대한 3차원 정보가 필요하기 때문에 경우에 따라서는 추가적인 수작업이 필요할 때도 있다.

본 논문에서는 지금까지 언급한 기존 방법들의 문제점을 개선한 비디오 시퀀스의 파노라마 합성 방법을 제시한다. 이 방법으로 우리는 프레임 수에 관계없이 파노라마 영상을 합성할 수 있다. 제안한 방법은 카메라의 상하좌우 이동, 초점거리의 변화 등 카메라의 움직임이 무엇인지에 관계없이 효율적으로 동작한다. 또한 제안한 방법은 파노라마 합성에 필요한 전역 움직임 변수(global motion parameter)를 자동적으로 구할 수 있다. 그리고 이 방법은 아웃라이어 배제(outliers rejection) 방법에 기초한 움직임 벡터 필터링(motion vector filtering)을 사용하여 움직이는 물체가 있는 경우에도 파노라마 영상을 만들어 낼 수 있다. 더 나아가 본 논문에서 우리는 물체의 잔상을 배경 영상으로 대체할 수 있는 새로운 알고리즘을 제시한다. 이를 통해 제안한 알고리즘은 기존의 파노라마 알고리즘에 비해 보다 더 정확하고 깨끗하며 자연스러운 영상을 제공한다.

II. 이론적 배경

영상 모자이크 알고리즘은 크게 세 단계로 구성된다. 먼저 인접 영상간의 전역 움직임을 추정하여, 전역 대응 변수를 추정한다. 그 이후에는 전 과정에서 얻은 변수를 통해 워핑 변수를 계산한다. 워핑 변수는 최종 합성된 변수에 있어 각 프레임이 어디에 위치하는가를 정의해주는

변수이다. 워핑 변수 값은 크기 관계(scaling factor)와 삼차원 좌표를 이차원에 연결짓는 기타 변수에 의해 영향을 받는다. 마지막으로, 추정된 변수를 통해 파노라마 영상을 합성한다.

이를 효과적으로 수행하기 위해 본 논문에서는 네 개의 변수로 구성되는 새로운 카메라 모델을 제시한다. 그리고 움직이는 물체에 의한 영향을 제거하는 방법도 제시한다.

1. 제안한 카메라 모델

기존 영상을 이용하여 어떤 영상의 위치를 지정(register)하기 위해서는 카메라 움직임 모델이 필요하다. 일반적인 카메라 움직임 모델은 어파인(affine), 투시(perspective) 모델^[13] 등이 있다. 어파인 모델은 가장 일반적으로 사용하는 움직임 모델로, 카메라의 회전(rotation), 전이(translation), 그리고 줌(zoom)을 고려한 움직임 모델이다. 어파인 모델^{[4][14]}은 식 (1)과 같이 정의한다.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \\ m_4 & m_5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} m_3 \\ m_6 \end{bmatrix} \quad (1)$$

어파인 모델을 더 단순화시킨 모델이 식 (2)의 4-변수 모델이다.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2)$$

투시 모델은 삼차원 세계를 이차원 평면으로 투사할 때 일어나는 움직임을 고려한 것으로, 카메라에서 일어날 수 있는 패닝(panning), 기울임(tilting), 회전, 전이, 줌 등의 모든 움직임을 고려한 모델이다. 앞에서 언급한 모델들은 파노라마 영상 합성에 사용될 영상 시퀀스(sequence)의 특성을 잘 반영하지 못하므로 개선이 필요하다.

그림 1은 본 논문에서 제안한 카메라 모델이다. 본 모델은 카메라의 패닝, 기울임, 초점거리 변화 등의 움직임을 기초로 했다. 카메라의 회전 움직임은 파노라마 합성을 위해 영상을 얻을 때 거의 일어나지 않는 움직임이므로 고려하지 않았다. 이 모델을 기반으로 할 때 영상을 변환하는 데 필요한 변수는 4개이다. 이는 기존의 모델

이 필요로 하는 변수 수보다 적은 것이다. 이것은 영상 내에 특징 수가 많지 않은 경우에도 움직임을 추정할 수 있는 것을 의미하므로, 영상에 화면결(texture)이 많지 않은 경우에 이 모델은 매우 유용하다.

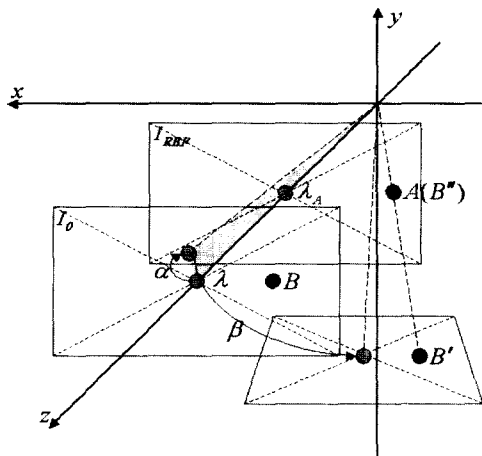


그림 1. 제안한 카메라 모델
Fig. 1. Proposed camera model

그림 1에서 α 는 제안한 모델에서 기울임 된 정도를 나타낸다. β 는 팬 움직임 정도를 나타낸다. λ 는 기준 영상에 위치지정(registration)을 해야 할 입력 영상의 초점거리이며 λ_A 는 기준 영상의 초점거리이다. 만약 입력 영상 I_0 의 한 점 B 가 기준 영상의 점 A 와 일치한다고 하면, 그것은 서로 다른 영상에 있는 두 점이 3차원 세계 좌표계에서는 동일한 물리적인 점이며, 그 점이 카메라의 움직임에 따라 각각의 2차원 영상에 투영되었다고 해석할 수 있다. 따라서, 만약 우리가 카메라의 움직임을 추정할 수 있다면, 우리는 모든 영상 시퀀스에 대한 대응관계(correspondence)를 찾을 수 있고, 그런 후에 모든 영상을 기준 영상에 맞춰 정렬할 수 있다.

영상 I_0 에 있는 점 B 는 카메라의 패닝과 기울임에 따라 다음과 같은 관계식에 의하여 B' 으로 옮겨진다.

$$\begin{bmatrix} b_{i,x} \\ b_{i,y} \\ b_{i,z} \end{bmatrix} = \begin{bmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} b_{i,x} \\ b_{i,y} \\ \lambda \end{bmatrix} \quad (3)$$

$$= \begin{bmatrix} \cos \beta & -\sin \alpha \sin \beta & -\cos \alpha \sin \beta \\ 0 & \cos \alpha & -\sin \alpha \\ \sin \beta & \sin \alpha \cos \beta & \cos \alpha \cos \beta \end{bmatrix} \begin{bmatrix} b_{i,x} \\ b_{i,y} \\ \lambda \end{bmatrix}$$

다시, B' 은 초점 거리 변화의 결과로 다음과 같은

등식에 의해 B'' 으로 옮겨진다.

$$\begin{aligned} b''_{i,x} &= \lambda_A \frac{b_{i,x}}{b_{i,z}} \\ &= \lambda_A \frac{b_{i,x} \cos \beta - (b_{i,y} \sin \alpha + \lambda \cos \alpha) \sin \beta}{b_{i,x} \sin \beta + (b_{i,y} \sin \alpha + \lambda \cos \alpha) \cos \beta} \end{aligned} \quad (4)$$

$$\begin{aligned} b''_{i,y} &= \lambda_A \frac{b_{i,y}}{b_{i,z}} \\ &= \lambda_A \frac{b_{i,y} \cos \alpha - \lambda \sin \alpha}{b_{i,x} \sin \beta + (b_{i,y} \sin \alpha + \lambda \cos \alpha) \cos \beta} \end{aligned}$$

B'' 은 기준 영상 I_{REF} 에 있는 점 A 와 일치해야 하기 때문에, 다음의 관계식이 성립한다.

$$\begin{aligned} a_{i,x} &= b''_{i,x} \\ a_{i,y} &= b''_{i,y} \end{aligned} \quad (5)$$

등식 (5)를 사용하여, 영상의 일치하는 두 점이 주어졌을 때 연립 일차방정식을 풀어서 우리는 4개의 변수, α , β , λ , λ_A 를 구할 수 있다.

2. 위치 지정(Registration)

제안한 영상 모자이크 알고리즘에서 위치지정은 다음의 세 단계로 구성된다.

- (1) 연속되는 프레임간의 움직임 추정
- (2) 추정된 움직임 벡터의 아웃라이어 배제(outliers rejection)
- (3) 전역 대응 변수 계산

여기서 움직임을 추정하는 과정과 전역 대응 변수를 계산하는 것은 서로 다른 두 가지 과정이라는 것을 유의해야 한다. 움직임 추정은 연속된 두 프레임 간의 일치 관계를 결정하는 것이고, 전역 대응 변수들은 이 일치에 대한 정보를 바탕으로 계산되는 것이다. 이 경우에는, 각 영상을 정사각형의 블록으로 나누고, 각 블록과 일치하는 부분을 바로 이어지는 영상에서 찾는다. 그러면 전역 대응 변수는 일치하는 블록들의 움직임 벡터를 이용하여 계산할 수 있다.

2.1 움직임 추정

전역 대응 변수 α , β , λ , λ_A 를 찾기 위한 자료로 쓰일 두 영상간의 전역 움직임을 계산하기 위해 우리는 연속된 영상 사이에 일치하는 부분을 찾아야 한다. 이 논문에서, 우리는 특징 기반 접근법(feature-based method)보

다는 특징을 사용하지 않는 위치 지정(featureless registration)을 하였다. 그 이유는 실제적인 경우 영상에는 보통 유용한 특징(feature)이 얼마 되지 않으며, 화면결(texture)이 없는 배경에는 종종 특징이 아예 없기도 하기 때문이다. 제안한 방법에서, 우리는 특징 배제 위치 지정 방법의 하나인 블록 매칭 알고리즘(Block Matching Algorithm)을 사용했다. 이 알고리즘을 적용하기 위해 우리는 우선 각 영상을 일정한 크기의 정사각형 블록으로 나눈다. 그 후에 첫 번째 영상에 있는 각 블록에 대하여, 각 화소간의 차의 제곱의 합이 최소인 블록을 두 번째 영상에서 찾는다. 그러면, 그 블록이 바로 첫 번째 영상의 블록과 가장 잘 일치하는 블록이고, 움직임 벡터는 두 블록간의 변위이다.

2.2 아웃라이어 배제(Outliers rejection)

바로 전 단계에서 얻은 움직임 벡터들 중 일부는 움직이는 물체나 잡음에 의한 것으로, 정적인 배경의 움직임과는 무관하므로, 전역 움직임 변수를 계산하는 데 쓰이면 안 된다. 따라서 이 논문에서는 이와 같은 움직임 벡터들을 추정과정에 있어서 좋지 않은 자료인 아웃라이어(outlier)로 간주한다. 아웃라이어들은 전역 움직임 변수를 계산하기 전에 시간적, 공간적인 관점에서 제거하였다.

공간적인 기준으로, 우리는 움직임 벡터의 각과 크기를 모두 사용하였다. 먼저, 우리는 모든 움직임 벡터의 각을 계산하였다. 만약 우리가 움직임 벡터를 M 으로 정의하면, 그 각 $f_A(M)$ 은

$$f_A(M) = \tan^{-1}(m_y/m_x) \quad (6)$$

이다. 여기서 m_x 와 m_y 는 각각 움직임 벡터의 가로와 세로 성분이다. 그런 후에, 모든 각들은 $\pi/12$ 간격으로 양자화 된다. 그러면 우리는 이 과정에 의해 24개의 각의 집합을 얻을 수 있고, 그 중 가장 수가 많은 것이 주요 각이 된다. 그런 후에, 미리 정의된 한계값(threshold) 이상 그 지역의 주된 각의 값을 벗어나는 모든 움직임 벡터들은 아웃라이어로 분류하였다. 그런 다음, 다시 크기에 대한 중앙값 필터링을 수행하여 지역적인 아웃라이어들을 제거하였다.

한편 초점거리 변화, 혹은 줌의 경우에 있어서는, 일부 유효한 움직임 벡터들이 바로 전에 언급한 공간적 다수 필터링(spatial majority filter)에 의해 제거될 수 있다. 그림 2의 영상에서 전역 줌이 일어나는 점선으로 표시된 부

분은 좋은 예이다. 여기서, 이 영역에 속한 움직임 벡터들은 서로 다른 방향이지만, 영상의 줌을 고려하기 위해서는 제거되면 안 된다. 따라서 본 논문에서는 현재 영상의 움직임 벡터를 앞, 뒤 프레임에서 같은 위치에 있는 벡터와 비교한 뒤, 앞, 뒤 프레임과 다른 움직임 벡터들은 제거한다. 만약 움직임 벡터들이 바로 뒤따라오는 프레임과는 같은 각을 갖지만 전 프레임과는 다르다면, 그것은 현재의 프레임이 새로운 형태의 전역 움직임의 시작점이라고 해석할 수 있다. 결론적으로, 시공간적인 테스트 양쪽 모두를 통과하지 못한 벡터들이 아웃라이어로 취급되어, 전역 움직임 변수 계산 전에 제거된다.

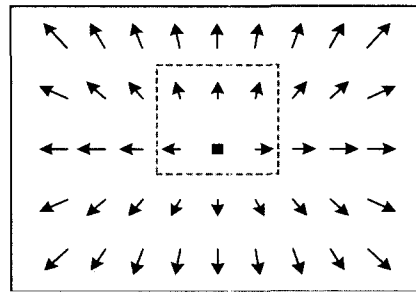


그림 2. 전역 움직임이 일어나는 영상의 움직임 필드
Fig. 2. Motion field of an image where global zooming occurs

2.3 전역 움직임 변수의 계산

파노라마 합성은 프레임간(frame-to-frame) 접근법에 기초한다. 즉, 어떤 기준 영상이 주어지면, 이 기준 영상과 나머지 영상간의 전역 대응 변수(global mapping parameter)가 계산되며, 각 영상은 파노라마 영상이 구축될 때 기준 영상의 좌표계로 정렬된다. 이 때, 결과 모자이크 영상의 좌표계는 기준 영상의 좌표계를 따른다.

식 (4)와 (5)로부터, 첫 번째 영상에서 2차원 영상의 점 (x, y) 를 두 번째 영상의 대응점 (x', y') 로 대응시키는 것은

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda_A \cos \beta & -\lambda_A \sin \alpha \sin \beta & -\lambda_A \lambda \cos \alpha \sin \beta \\ 0 & \lambda_A \cos \alpha & -\lambda_A \lambda \sin \alpha \\ \sin \beta & \sin \alpha \cos \beta & \lambda \cos \alpha \cos \beta \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (7)$$

와 같다. 4개의 변수 $\alpha, \beta, \lambda, \lambda_A$ 는 전에 언급한 바와 같이 두 쌍의 일치점으로 구성할 수 있는 선형 연립 방정식의 해를 구함으로써 계산할 수 있다. 그러나, 전 영상의 전역 움직임을 정확히 표현하는 두 쌍의 일치점을 찾는 것은 실제적으로 어렵다. 따라서, 본 논문에서 우리는 2쌍

이상의 모든 유효한 일치점들을 사용하여 최소 자승법으로 움직임 변수들을 추정한다.

n 쌍의 일치점들에 대하여, 식 (5)는

$$\begin{bmatrix} x_1' & x_2' & \dots & x_n' \\ y_1' & y_2' & \dots & y_n' \\ 1 & 1 & \dots & 1 \end{bmatrix} = \begin{bmatrix} k_1 & k_2 & k_3 \\ k_4 & k_5 & k_6 \\ k_7 & k_8 & k_9 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \\ 1 & 1 & \dots & 1 \end{bmatrix} \quad (8)$$

의 형태이다. 최소 자승법을 적용하여 변환행렬의 원소 k_i 를 얻으면, 우리는 식(5), (6)의 관계를 통해 움직임 변수 $\alpha, \beta, \lambda, \lambda_A$ 를 추정할 수 있다. 그러나 파노라마 영상을 합성하기 위해 연속된 2차원 영상간의 전역 대응 변수를 구하는 것이 이 단계의 목적이라면 그것은 $[k_1, k_2, \dots, k_9]^T$ 를 아는 것만으로도 충분하다.

3. 모자이크 변수의 추정

모자이크 변수는 각 프레임을 최종적으로 합성되는 파노라마 영상에 어떻게 정렬할 것인지를 결정한다. 제안한 방법은 프레임간 접근 방법에 기반하고 있고 우리는 시퀀스의 마지막 영상을 기준 영상으로 정했다. 그러면, 시퀀스 내의 임의의 영상간의 움직임 변수는 변환행렬의 연속적인 곱에 의해 결정된다. 만약 두 영상 I_k 와 I_{k+1} 간의 변환을 3×3 행렬 \mathbf{T}_k 로 정의하면, I_k 와 I_{k+n} 간의 변환은 $\mathbf{T}_k \mathbf{T}_{k+1} \mathbf{T}_{k+2} \dots \mathbf{T}_{k+n-1}$ 이 된다. 그림 3은 이 과정을 나타내고 있다. A 로 표시된 블록은 두 연속된 영상간의 움직임을 정의하는 전역 대응 행렬을 가리키고, T 는 모자이크 변수 행렬을 가리키며, S 는 최종적인 결과인 파노라마 영상이다. 만약 파노라마 영상의 해상도가 시퀀스의 각 프레임의 해상도와 동일하게 설정된다면, 이 추정 과정 뒤에 줌 과정이 추가된다.

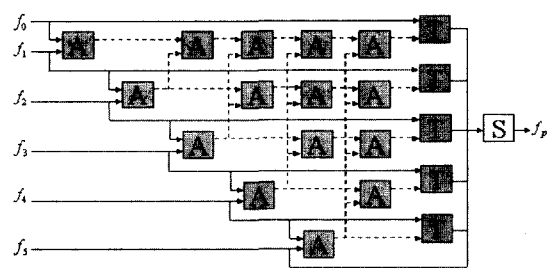


그림 3. 모자이크 변수 추정의 개념도
Fig. 3. Block diagram of mosaic parameter estimation

4. 영상 모자이크 합성하기

이 절에서는 입력 영상의 움직이는 물체들이나 노이즈를 제거한 후 하나의 파노라마 영상으로 합성하는 방법을 제안한다. 기존의 방법은 움직이는 물체를 다루는 데 있어서 비효과적이었다. 제안한 알고리즘은 합성 과정에서 움직이는 물체를 효과적으로 제거하기 위해 영상 시퀀스 전체에 걸친 해당 화소들을 이용하여 각각의 화소나 화소들의 그룹의 시간적인 도수분포도를 구한다. 전역 움직임 변수를 사용하여, 영상 시퀀스에 걸친 각 점의 위치를 추적하는 것이 가능하다. 그러면, 여러 장의 영상에 걸쳐 있는 동일한 점의 밝기 값의 목록을 얻을 수 있다. 만약 그 값이 시퀀스 전체에서 동일하게 유지된다면, 그 화소는 배경의 일부뿐이라고 할 수 있다. 그러나 해당 지점에서 밝기 값이 달라지면, 그 부분은 움직이는 물체나 잡음이라고 판단할 수 있으므로, 합성과정에서 이 화소는 시간적 도수분포도의 최대값으로 교체된다.

III. 실험 결과

경기장 전체의 확장된 시야를 제공하기 위해 축구 경기 방송 영상 시퀀스에 제안한 알고리즘을 적용하였다. 그림 4는 시험 시퀀스 "축구"의 처음과 마지막 프레임을 보여 주고 있다. 이 시험 영상은 선수나 공과 같은 움직이는 물체를 임의로 많이 가지고 있는 반면 정적인 요소는 적기 때문에, 영상을 파노라마 영상으로 합성하기에 매우 어렵다. 따라서 제안한 알고리즘을 시험하기에 알맞다. 모의실험에 사용된 모든 시퀀스 내의 각 영상의 크기는 720×486 이다. 연속된 영상을 위치 지정하기 위해서, 블록 매칭 알고리즘을 적용했다. 각 블록의 크기는

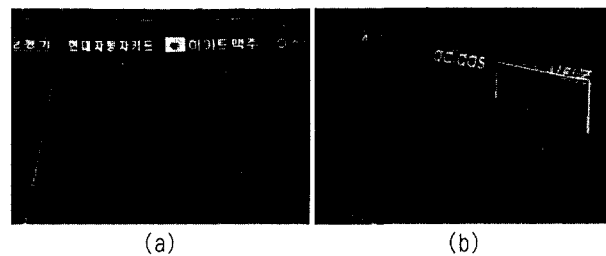


그림 4. (a) "축구" 시퀀스의 첫 번째 프레임 (b) "축구" 시퀀스의 마지막 프레임
Fig. 4. (a) the first frame of "Football" sequence (b) the last frame of "Football" sequence

32×32이다. 탐색 영역은 가로 방향 ±64, 세로 방향 ±32로 정했다. 각 블록에 대한 움직임 벡터를 얻은 후에, 시간적 필터링에 기초한 아웃라이어 배제를 수행하였다.

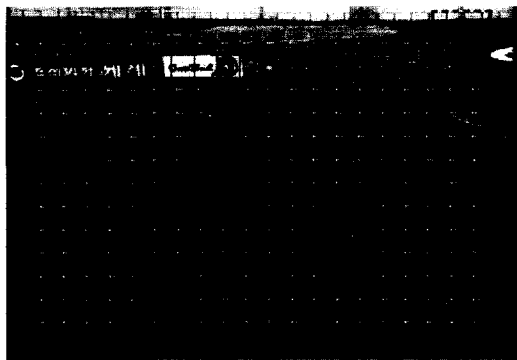


그림 5. 아웃라이어 배제의 결과
Fig. 5. the result of outliers rejection

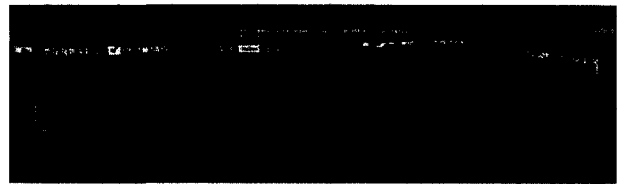
그림 5는 "축구"시퀀스의 185번째 프레임에 대한 아웃라이어 배제의 결과이다. 어두운 블록은 그들의 움직임 벡터가 아웃라이어로 탐지되었음을 가리킨다. 이를 통해 선수들의 움직임에 의한 움직임 벡터들이 효과적으로 제거되었음을 볼 수 있다. 그림 6은 그 움직임 벡터들을 제거하지 않은 파노라마 영상을 보여준다. 그림에서 보듯이 결과 영상에는 심한 왜곡이 있다. 이것은 영상 시퀀스 후반부의 줌이 일어나는 영상에 있는 아웃라이어들 때문이라고 할 수 있다. 그림 7의 파노라마 영상은 제안한 아웃라이어 배제 알고리즘을 사용하여 얻었다. 그림 7은 제안한 방법이 모자이크 변수를 계산하는 데 있어서 효과적이고 강력하다는 것을 보여준다.

파노라마 영상으로부터 움직이는 물체를 제거하여 배경 자료를 더욱 완전하고 깨끗하게 볼 수 있다. 움직이는 물체들, 이 경우에 있어서 선수들은 영상 시퀀스에서 해당되는 영상 화소들의 시간적 도수분포도를 이용하여 배경 영상으로 교체된다. 제거는 블록을 기본으로 하여 수행되었고, 블록의 크기는 8×8이었다. 전역 움직임 변수들을 사용하여, 각 블록을 다음 프레임들의 해당 블록들과 비교한다. 시간적인 아웃라이어 블록들을 찾기 위해서, 블록 내의 화소의 밝기에 대한 평균값이 사용되었다. 이 방법은 각 프레임을 1/8로 부표분화한 후, 부표분화한 영상의 각 화소를 시간적인 방향에서 비교한 것으로 볼 수 있다. 그림 7(b)는 제안한 방법을 사용하여 선수들이 제거된 파노라마 영상을 보여준다. 그림 7(a)와 비교해 볼 때, 선들과 기타 정적인 배경은 영상에서 잘 보존되어 있는 반면 선

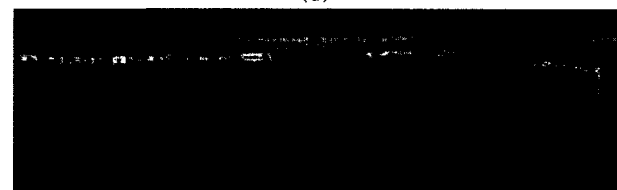
수들은 성공적으로 제거된 것을 볼 수 있다.



그림 6. 아웃라이어 배제를 사용하지 않은 "축구" 시퀀스의 파노라마 영상
Fig. 6. Panoramic image of "Football" sequence without outliers rejection



(a)



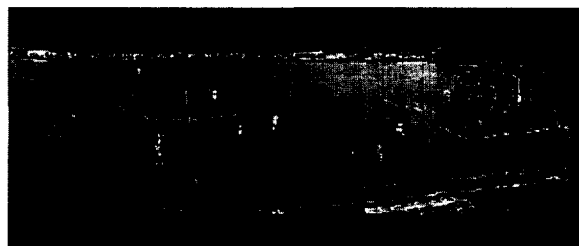
(b)

그림 7. 아웃라이어 배제를 사용한 "축구" 시퀀스의 파노라마 영상 (a) 선수들 지우지 않음 (b) 선수들 제거
Fig. 7. Panoramic images of "Football" sequence using the outliers rejection (a) players not erased (b) players removed

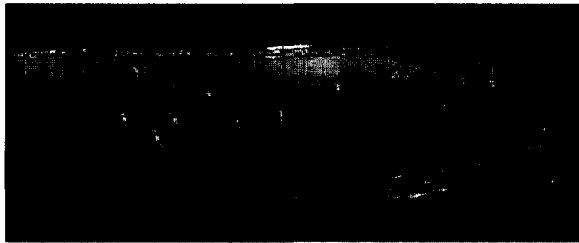
제안한 카메라 모델의 효과성을 증명하기 위해, 기존 영상 변환 모델을 사용하여 만든 파노라마 영상과 비교해 보자. 이 실험에서 사용된 "축구 2"시퀀스는 관중의 손 움직임에 의한 심한 잡음 등이 존재하고 후반부에는 카메라의 줌이 일어나기 때문에 "축구 1" 시퀀스에 비하여 파노라마 영상을 합성하기 더욱 어렵다. 그림 8(a), (b), (c)는 각각 제안한 모델, 6-변수 어파인 움직임 모델, 4-변수 모델을 사용하여 만든 파노라마 영상을 보여준다.

움직임 모델들을 비교하는 데 있어서, 공평한 비교를 확실히 하기 위해 모든 조건은 동일하게 설정하였다. 동일한 조건이란 비교를 위해 아웃라이어 배제를 적용하고, 모든 세 가지 움직임 모델에 동일한 움직임 벡터를 사용한 것을 가리킨다. 기존 움직임 모델을 사용한 파노라마 영상의 왜곡은 줌을 잘못 다루었기 때문일 수 있다. 두 기존 모델에 있어서 줌은 효과적으로 기술되지 않기 때문에, 영

상 시퀀스의 뒷부분에서 일어나는 전역 대응 에러는 점점 더 커지고 합성 과정에서 에러는 시퀀스의 첫 프레임에까지 쌓인다. 결과적으로, 두 기존 모델을 사용한 파노라마 영상은 더 많은 왜곡을 포함하고 있다. 제안한 모델이 카메라의 페닝, 기울임과 초점 거리의 변화를 효과적으로 나타내기 때문에 제안한 방법에 의해 만들어진 파노라마 영상에는 기하학적 왜곡이 거의 없다.



(a)



(b)



(c)

그림 8. "축구 2" 시퀀스의 파노라마 영상 (a) 제안한 모델 사용 (b) 6-변수 움직임 모델 사용 (c) 4-변수 움직임 모델 사용
Fig. 8. Panoramic images of "Football 2" sequence using (a) proposed model (b) 6-parameter model (c) 4-parameter model

면을 다루기에 적절한 방법을 제공해 주지 못했다. 그 방법들은 완전히 자동화되지 않았다. 또한, 복잡한 추정 과정이 필요한 비 선형 방정식 시스템이 되는 카메라 모델에서는 해를 얻기 힘들었다.

이 논문에서는, 완전 자동적인 방법으로 파노라마 영상을 합성할 수 있는 새롭고 효과적인 영상 모자이크 방법을 제공했다. 제안한 방법에서는, 카메라 변수의 추정을 용이하게 하면서도 큰 정확도를 달성하는 단순화된 카메라 모델을 제시했다. 가로와 세로 움직임이 주된 부분을 차지하는 반면 화면의 회전은 거의 일어나지 않는 일반적인 파노라마 영상의 특성을 활용했다. 이 모델은 2차원 영상 평면과 카메라 움직임과의 관계가 단순하면서도 더 정확도가 높았다. 제안한 모델은 연립 1차 방정식을 통해 쉽게 구할 수 있는 4개의 변수만으로 카메라의 페닝, 기울임, 줌을 다룰 수 있었다. 우리는 또한 움직이는 물체들과 잡음이 전역 움직임 변수 추정에 영향을 주는 것을 방지하기 위한 좋은 위치 지정 방법을 제안했다. 전역 움직임 벡터를 구할 때 움직이는 물체들과 잡음에 의한 움직임은 아웃라이어 배제 방식에 따라 제거하였다. 제안한 아웃라이어 배제 방식은 움직임 벡터들의 시간적, 공간적인 정보를 둘 다 활용한다. 실험적인 결과로 움직이는 물체들, 지역적 움직임, 잡음 등이 있는 상황에서도 영상이 올바르게 위치 지정된다는 것을 알 수 있었다. 그리고, 움직이는 물체에 의한 시각적인 결함(artifact)을 배경 영상으로 대체하는 새로운 알고리즘도 제안했다. 제안한 알고리즘을 사용함으로써, 영상 합성과정에서 생기는 움직이는 물체들에 의한 고스트 효과와 심한 번짐 현상과 같은, 여러 가지 시각적인 잡음을 성공적으로 제거할 수 있었다.

제안한 알고리즘을 축구 경기 방송영상에 적용하였고 알고리즘이 영상 모자이크 혹은 파노라마 영상을 합성하는 데 매우 효과적임을 보였다. 제안한 방법은 또한 다른 운동경기나, 위성, 항공 영상 합성 등 낮은 해상도의 입력 영상으로부터 넓은 공간적인 정보를 필요로 하는 다른 여러 종류의 영상 처리 시스템 등에도 적용할 수 있다.

IV. 결 론

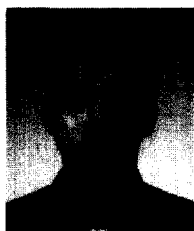
영상 모자이크는 모든 입력 영상들을 전 영상 시퀀스가 포함하고 있는 시야를 담고 있는 하나의 넓은 영상으로 합성하여 배경에 대한 확장된 시야를 제공해 준다. 기존의 영상 모자이크 알고리즘은 정적인 영상에만 초점을 맞추었기 때문에 움직이는 물체들이나 잡음이 있는 동적인 화

참 고 문 헌

- [1] R. Szeliski, "Video Mosaics for Virtual Environments," *IEEE Computer Graphics and Applications*, Vol. 16, No. 2, pp. 22-30, March 1996
- [2] A. Matos, J. Gomes, A. Parente, H. Siffert, and L. Velho, "The Visorama System : a Functional

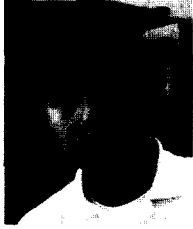
- Overview of a New Virtual Reality Environment," *Proc. of Computer Graphics International*, pp. 205-212, 1997
- [3] S. E. Chen, "Quicktime VR - An Image-Based Approach to Virtual Environment Navigation," *Proc. of ACM Computer Graphics (SIGGRAPH'95)*, pp. 29-38, Aug. 1995
- [4] M. Irani and S. Peleg, "Improving Resolution by Image Registration," *Graphical Models and Image Processing*, Vol. 53, No. 3, pp. 231-239, May 1991
- [5] D. M. Nazarenko, G. Mitchell, and G. C. Staples, "Use of RADARSAT ScanSAR Products for Regional Mapping in Southeast Asia," *IEEE International Geoscience and Remote Sensing, IGARSS'97*, Vol. 2, pp. 1060-1062, Aug. 1997
- [6] S. Saatchi, F. De Grandi, M. Simard, and E. Podest, "Classification of JERS-1 Image Mosaic of Central Africa using A Supervised Multiscale Classifier of Texture Features," *Proc. of International Geoscience and Remote Sensing Symposium, IGARSS'99*, Vol. 5, pp. 2765-2767, 1999
- [7] M. Holm and S. Rautakorpi, "Automatic creation of image mosaics and digital surface models using airborne digital camera data," *Proc. of the Videometrics*, Vol. 6, pp. 139-150, Jan. 1999
- [8] D. Yow, B. L. Yeo, M. Yeung and B. Liu, "Analysis and Presentation of Soccer Highlight from Digital Video," *Proc. of Second Asian Conference on Computer Vision*, Vol. II, pp. 499-503, 1995
- [9] S. R. Yhann and T. Y. Young, "Boundary Localization in Texture Segmentation," *IEEE Trans. on Image Processing*, Vol. 4, No. 6, pp. 849-856, June, 1995
- [10] Y. Xiong and K. Turkowski, "Registration, Calibration and Blending in Creating High Quality Panoramas," *Proc. of Fourth IEEE Workshop on Applications of Computer Vision, WACV'98*, pp. 69-74, 1998
- [11] Y. Gong, G. Proietti, and D. LaRose, "A Robust Image Mosaicing Technique Capable of Creating Integrated Panoramas," *Proc. of 1999 IEEE International Conference on Information Visualization*, pp. 24-29, 1999
- [12] K. Mase and H. Nishira, "Computing the Field-of-View of a Stitched Panorama to Create FoV Sensitive Virtual Environments," *Proc. of the 13th International Conference on Pattern Recognition*, Vol. 1, pp. 151-155, 1996
- [13] Lisa G. Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys*, vol. 24, No. 4, December 1992
- [14] J. Koenderink and A. Van Doorn, "Affine structure from motion," *Journal of the Optical Society of America*, Vol. 8, 1991

저 자 소 개



서 종 열

1994년 2월 : 연세대학교 전자공학과 졸업 (공학사)
 1996년 8월 : 연세대학교 대학원 전자공학과 졸업 (공학 석사)
 2000년 8월 : 연세대학교 대학원 전기·컴퓨터 공학과 졸업 (공학 박사)
 2000년 9월~현재 : LG전자 선임연구원
 주관심분야 : 디지털 비디오 처리, 비디오 시스템 설계, 파노라마 합성



박 종 현

2001년 2월 : 연세대학교 기계전자공학부 졸업(공학사)
 2001년 3월~현재 : 연세대학교 대학원 전기전자공학과 석사과정
 주관심분야 : 영상 복원, 영상 향상, 영상 모자이크



강 문 기

1986년 2월 : 서울대학교 전자공학과 졸업(공학사)
 1988년 8월 : 서울대학교 대학원 전자공학과 졸업(공학석사)
 1994년 2월 : Northwestern University 졸업(공학박사)
 1994년 3월~1994년 10월 : Northwestern University Research Fellow
 1994년 11월~1997년 2월 : University of Minnesota, Duluth 조교수
 1997년 3월~현재 : 연세대학교 전자공학과 부교수
 주관심분야 : 영상복원, 초해상도 영상복원, 디지털 비디오 처리