

가상환경을 위한 파노라마 생성에 관한 연구

김태은*

요약

본 논문에서는 환경 탐색을 위한 영상 기반 환경을 자동으로 생성하기 위한 새로운 모자이크 기법을 제안 한다. 제안된 방법은 영상의 투영 변환 행렬을 이용하여 한 영상 위에 다른 영상을 합성 한다. 특징 모델을 기반한 정합을 이용하여 영상을 모자이크를 구현하고 구매펍을 통해 시각에 따라 변형되어 몰입감을 주는 영상 기반 환경을 생성한다, 특히, 취득한 영상간의 카메라 회전 성분에 강건한 기준 특징 벡터 필터링 방법을 제안하며 실험을 통해 그 유용성을 검증한다.

Panoramic Image Generation for the Virtual Environment

Tae-Eun Kim*

Abstract

This paper proposes a new algorithm that generates a panorama image for reconstruction of image-based virtual environment. Proposed algorithm can generate panorama image from input images using the projective transform based on the feature model and a fully immersive image-based virtual environment according to the user's view point by the sphere mapping. Especially, proposed algorithm supports robust methods of camera rotation and is applied to various images to estimate its performance.

Keywords : 가상환경(virtual environment), panorama image

1. 서론

최근 컴퓨터 하드웨어의 급속한 성장과 이에 따른 정보통신의 발전에 따라 인터넷 또는 컴퓨터 상에서의 가상환경 구현을 위한 다양한 기술들이 제안되고 있다. 가상 현실 시스템에서 3차원 세계를 표현하는 대표적인 두 종류의 방법은 3차원 모델을 이용하는 모델 기반 표현방법(model based method)과 2차원 영상을 이용하는 영상 기반 표현 방법(image based method)이다.

모델 기반 표현 방법은 가상환경내에 존재하는 모든 객체들에 대해서 3차원 모델을 생성하기 때문에 사실적인 표현이 가능하다. 그러나 3

차원 모델을 생성하기 위해서는 긴 처리 시간이 필요하고, 복잡한 모델은 표현할 수 있는 한계가 있다. 이에 반해, 영상 기반 표현(image-based method) 기술을 이용한 방법은 가상 환경을 표현하는데 있어서 더욱 빠른 속도로 실감 있는 환경을 표현해줄 수 있는 방법을 제시한다[1]. 영상 기반 환경을 표현하기 위해서는 카메라로 취득한 영상이 필요하고 그 영상들을 하나의 큰 영상으로 합성한 후 사용자가 보는 위치에서의 시선을 고려하여 영상을 변환하여 표현한다.

따라서 본 연구는 카메라로 취득한 영상을 모자이크 및 파노라마로 생성한 후 사용자의 시선을 고려하여 적절히 변형되어 몰입감 있는 영상 기반 환경을 생성하는 방법을 제시한다. 특히 모자이크 환경 영상을 생성하기 위하여 기준 특징 벡터 필터링을 사용하여 대응 관계를 검출 및 정합하였다.

기존 모자이크 방법으로는 크게 영상 내 색상 정보의 상관에 기반한 방법(correlation-based)과 특징점의 대응 관계에 기반(feature-based)한 방법 등이 있다. 상관 기반한 방법은 주로 비디오

제일저자(First Author) : 김태은

접수일자:2007년09월05일, 심사완료:2007년09월10일

* 남서울대학교 멀티미디어학과 교수

tekim@nsu.ac.kr

본 논문은 남서울대학교 교내 연구비지원에 의해서 이루어짐

연속(sequence) 영상을 대상으로 하여 작은 카메라의 모션을 대상으로 연구되어 왔다[1]. 카메라의 움직임이 큰 경우, 피라미드형의 계층적인 방법으로 반복 계산하는 방법과 푸리에(fourier) 변환에 따른 주파수 공간에서의 최대 임펄스를 이용하여 정합한 후 이를 구체화하는 방법 등이 있다[2]. 그러나 이러한 방법은 카메라의 좌, 우의 회전만을 고려하기 때문에 카메라의 광학축(optical axis)을 중심으로 회전하여 취득한 영상의 경우 정확한 정합을 찾아내기가 어려우며 또한 많은 양의 계산에 의한 시스템에 부하가 많이 걸리게 된다.

이에 반해 특징 기반 방법은 비교적 큰 움직임과 자유로운 회전을 극복하기 위해 연구되어 왔다[3,4]. 그러나 이러한 자유로운 회전을 극복하기 위해서는 강건한 특징점을 이용한 정확한 대응 관계(correspondence)가 필수적이다. Zoghlami[3]는 코너 모델을 이용하여 정합을 하였으며 T.J. Cham[4]은 베이지언 방법(Bayesian method)를 이용하여 특징점의 대응 관계를 찾아냈으며 대체로 좋은 결과를 보였다. 그러나 이러한 방법들은 예를 들면, 건물의 있는 창문들의 유사한 특성들에서 큰 오차가 발생하게 된다.

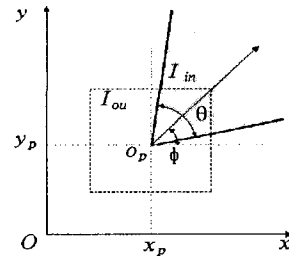
따라서 본 연구에서는 이를 개선하고자 정확한 특징 점 정합을 위한 효과적인 대응 관계 필터링 알고리즘을 제안한다. 유사한 특징에 의해 구해진 점들은 다대다의 대응관계를 가질 수 있으므로 이를 중간 벡터 필터링(Median vector filtering) 방법과 회전 상관 관계(rotational cross correlation)를 이용하여 정확한 일대일 대응 관계를 구한 후, 이 특징점들을 통해 평면 투영 변환 행렬(projective transform matrix)을 얻어낸다. 또한 영상 합성 과정에서 색상 차이에 따른 경계선을 없애고자 가우시안 분포 함수에 기반한 비선형 합성 함수를 적용하였다.

2. 대응 관계 검출

2.1 특징 모델 추출

카메라의 시점에서 큰 변화와 카메라 광학축(optical axis)에 대한 회전이 일어난 경우에도 정확한 정합을 이루기 위해서는 그러한 변화에도 강건한 영상 내의 특성이 필요하다. 그러한

특성으로는 모서리의 점을 중심으로 하여 일정한 윈도우 영역 내에서의 내부각, 내부 및 외부의 색상 정보가 있으며, 그 외 모서리 점의 좌표 및 모서리의 x축에 대한 상대적인 회전 각도 등이 좀 더 정확한 대응 관계를 찾아내기 위해 사용된다[3].



(그림 1) 특징 모델

위의 (그림 1)에서 볼 수 있듯이 6개의 파라미터 $P_i = (x_p, y_p, I_{in}, I_{out}, \theta, \phi)$ 에 의해 특징 모델을 정의하는데, 그 특징 모델의 중심이 되는 모서리 중심점은 국부(local) 영역 내의 곡률(curvature)값의 절대값이 최대인 점으로 하미식(1)과 같이 표현된다. 곡률을 계산하기에 위해서는 영상의 에지(edge)를 검출해야 하는데, 캐니 에지 검출기(canny edge detector)로 에지를 검출한 후 윤곽선에서의 끊어진 부분을 이어준다.

또한 곡률 스케일 공간(curvature scale space)[3]을 이용하여 σ 값을 변화시킴에 따라 더 정확한 모서리 점을 추출할 수 있다[5].

$$k(s) = \lim_{h \rightarrow 0} \frac{\phi}{h} \tag{1}$$

여기에서 ϕ 는 호의 길이 s 에 대한 탄젠트 벡터 $t(s)$ 와 $t(s+h)$ 사이의 각도를 나타내며, 식(2)는 이산 곡선(discrete curve)에 적용하기 위해 변형된 곡률 계산 식이다.

$$c_i = \sum_{s=0}^N \tan^{-1} \left(\frac{\partial x_s}{\partial y_s} \right) \tag{2}$$

얻어진 곡률의 미분을 취하면 부호가 바뀌는

점이 그 곡선의 곡률이 영점 교차(curvature-zero crossing)하는 부분이 되며 최종 특징 모델의 중심 좌표가 된다. 이 후 이 모서리 점들을 중심으로 모든 추출된 파라미터 값들을 얻어내어 기록한다.

2.2 대응관계의 검출

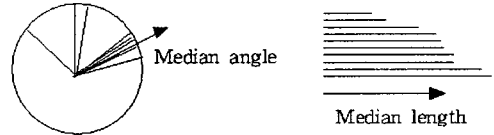
앞의 과정을 거쳐 추출한 영상 내 특징점을 통해 두 영상 간에 정확한 대응 관계를 찾아내는 것은 비중이 큰 중요한 부분이다. 특징점들간의 대응 관계를 찾기 위해 먼저 파라미터 값들을 비교하여 일정한 문턱치 값을 기준으로 추출한다. 비교 시 이용되는 파라미터들로는 특징점을 중심으로 일정 원도우 영역내의 내부각 θ , 내부, 외부 색상정보 I_{in}, I_{out} 를 사용한다.

특징 모델 성분만을 비교하여 얻어낸 대응관계는 정확한 대응점을 찾아내지 못할 확률이 높다. 예를 들면 건축 조형물에서의 창문들의 모서리들은 모두 유사한 파라미터를 가지고 있기 때문에 다대다 대응이 이루어질 수 있다. 이러한 문제를 해결하기 위해 중간 벡터 필터링 방법을 사용한다[6]. 이 방법은 겹치는 부분 내에 있는 일정한 국부(local) 영역 내의 특징점들은 특징 모델 중, 회전 각도($\phi_{des} - \phi_{ref}$)와 길이($\sqrt{(x_{pdes} - x_{pref})^2 + (y_{pdes} - y_{pref})^2}$) 값들의 유사성을 이용한다.

$$MVF_{local} = (\theta_{median}, l_{median}) \quad (3)$$

(그림 3)에서와 같이 일정한 원도우 내의 특징점들의 중간 각도(Median angle)와 중간 거리(Median length)값을 이용하여 이에 벗어난 특징점들을 제거한다.

최종 특징점들의 대응관계는 투영 변환 행렬에 큰 영향을 미치므로 최소 4쌍 이상의 정확한 대응점들만이 필요하다. 두 영상 내의 모든 특징점이 대응 관계를 지니고 있는 것이 아니라 겹쳐지는(overlapped) 부분에서만 특징점들의 대응



(그림 2) 중간 벡터 필터

관계가 있다. 즉 공통된 부분이 아닌 곳에서의 대응관계는 제거되어야 한다. 또한 윤곽선이 뚜렷이 나타나지 않는 부분 및 적은 수의 특징점을 지닌 영역에서는 위의 중간 벡터 필터로는 구별하기가 힘들다. 이 같은 문제를 제거하고 확실한 대응점을 얻기 위해 회전 상관 관계(rotational cross-correlation)를 이용한 필터링을 이용한다. 이 방법은 특징점을 중심으로한 영역을 그 특징점에 대해 중간 각도(median angle)만큼 회전시켜 상관 관계를 비교하는 방법으로써 정확한 매칭점을 찾을 수 있다.

3. 영상 합성

3.1 평면 투영 변환

고정된 시점에서 카메라를 회전하여 취득한 영상들은 투영 변환 행렬에 의해서 같은 평면위에 놓일 수 있다. XYZ-좌표계에서 한 점 $p(x, y, z)$ 은 투영 중심점에 대한 카메라의 회전(panning, tilting, rolling)에 의해 $p'(x', y', z')$ 로 변환되며 변환 행렬은 $det(\mathbf{R}) = \pm 1$, $\mathbf{R}^T = \mathbf{R}^{-1}$ 의 3×3 직교 회전 행렬로 표현된다.

$$(x', y', z')^T = \mathbf{R} \cdot (x, y, z)^T \quad (4)$$

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \quad (5)$$

변환된 점 $p'(x', y', z')$ 는 식(6)에 의해 $z' = f$ 의 2차원 평면의 한 점 $p_{img} = (x_{img}, y_{img})$ 에 투영된다.

$$\begin{aligned}
 x_{img} &= f \cdot \frac{x'}{z'} = f \cdot \frac{r_{11}x + r_{12}y + r_{13}f}{r_{31}x + r_{32}y + r_{33}f} \\
 y_{img} &= f \cdot \frac{y'}{z'} = f \cdot \frac{r_{21}x + r_{22}y + r_{23}f}{r_{31}x + r_{32}y + r_{33}f} \quad (6)
 \end{aligned}$$

결국, 투영 변환 행렬을 구하는 것이 관건인데 이는 두 영상 간의 최소 4쌍 이상의 대응관계를 알고 있으면 선형 연립방정식을 유도하여 구할 수 있다. 본 연구에서는 얻어진 특징점들의 좌표를 최소 자승 방법을 이용하여 중첩영역 내에서의 평면 투영 행렬을 구했다.

3.2 영상 합성

위의 투영 변환 행렬을 이용하여 한 영상 위에 다른 영상을 합성할 경우 색상의 차이에 따른 경계부분이 두드러지는 현상이 나타날 수 있다. 이는 블렌딩 기법으로 이를 해결할 수 있는데 그 중 가장 일반적인 방법은 영상의 중심에서 경계면까지의 거리에 따른 가중치에 색상 혼합비를 결정하는 선형 블렌딩하는 방법이다[1]. 그러나 이런 방법보다는 가우시안 분포 함수에 기반한 블렌딩 기법을 제안한다. 이 방법은 경계면과 두 영상이 겹쳐지는 중심 부분까지의 거리에 대한 가우시안 분포(Gaussian distribution) 가중치 w_g 를 두어 혼합하는 방법으로 보다 부드러운 결과를 보인다.

$$I_g = w_g \cdot I_A + (1 - w_g) \cdot I_B \quad (7)$$

4. 영상기반 환경 생성

모자이크 합성된 넓은 시야의 영상은 영상기반 환경에서 임의의 시각에서 보여질 경우에는 그 시각에서 바라보는 것처럼 느낄 수 있도록 보여줘야 한다. 전(omni-)방향에 대해서 몰입감을 줄 수 영상을 만들기 위해서는 구(sphere)맵핑이 이루어져야 한다. 기준 좌표계 $P = (X, Y, Z)$ 를 2D 구좌표계로의 변환식은 다음과 같다.

$$\begin{aligned}
 \theta &= \tan^{-1}(X/Z) \\
 \phi &= \tan^{-1}(Y/\sqrt{X^2 + Z^2}) \quad (8)
 \end{aligned}$$

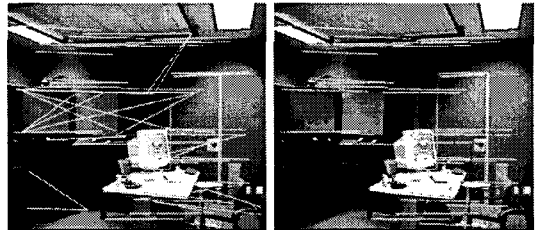
구맵핑을 하기 위해서는 카메라의 회전 각도 또는 카메라의 초점 거리를 알아야만 하는데 고정된 카메라이기 취득한 영상이기 때문에 투영 변환 행렬은 회전성분만을 갖게 되며, 이러한 직교 성질을 이용하여 초점 거리를 구할 수 있다.

$$M = \begin{pmatrix} r_{11} & r_{12} & r_{13}f_0 \\ r_{21} & r_{22} & r_{23}f_0 \\ r_{31}/f_1 & r_{32}/f_1 & r_{33}f_0/f_1 \end{pmatrix} \quad (9)$$

$$f_0^2 = \frac{r_{23}^2 - r_{13}^2}{r_{11}^2 + r_{12}^2 - r_{21}^2 - r_{22}^2} \quad (10)$$

$$f_0^2 = -\frac{r_{13} r_{23}}{r_{11} r_{21} + r_{12} r_{22}} \quad (11)$$

5. 실험 결과 및 결론



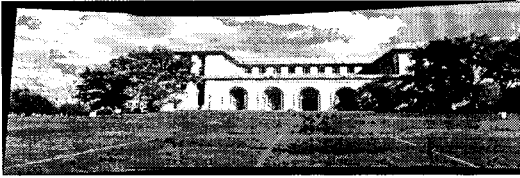
(그림 3) 대응 관계 필터링 처리 전(좌) 영상과 처리 후(우) 영상



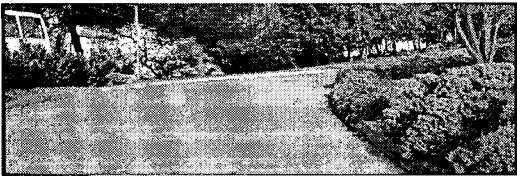
(그림 4) 블렌딩 처리 전(좌) 영상과 처리 후(우) 영상



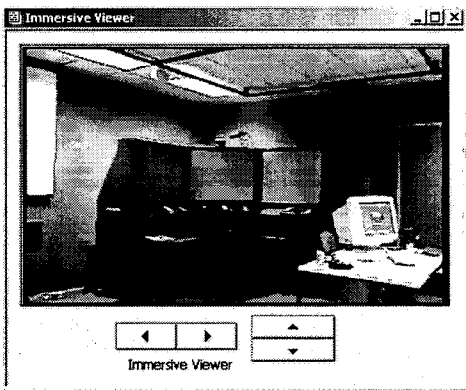
(그림 5) 파노라마 합성된 영상



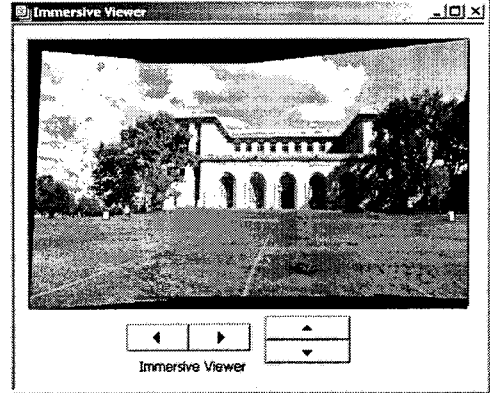
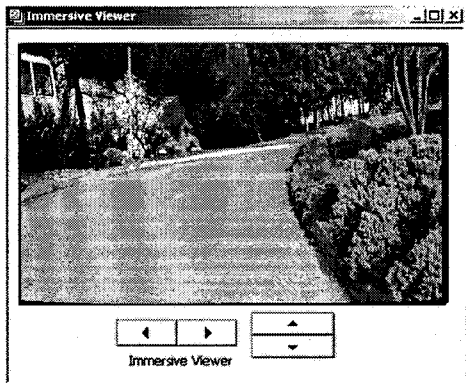
(그림 6) 파노라마 합성된 영상



(그림 7) 파노라마 합성된 영상



(그림 8) 영상 기반 몰입형 환경 뷰어



(그림 9) 영상 기반 몰입형 환경 뷰어

(그림 3)은 기준 벡터링 필터링을 이용하여 매칭되는 특징점들을 나타내며, (그림 4)는 블렌딩 함수를 이용하여 합성된 결과이며, (그림 5)는 10장의 영상을 가지고 파노라마 영상을 생성한 결과이다. 마지막으로 (그림 6)과 (그림 7)은 합성된 파노라마 영상을 사용자의 시각에 중심을 뒤서 변환하여 보여준 결과이다. (그림 8)과 (그림 9)는 영상기반 몰입형 환경뷰어를 나타내고 있다.

이미지 모자이크는 다양한 목적에서 영상간의 사영 관계를 이용해 다중의 영상을 결합하는 기법이다. 이 기법에서 영상간의 사영 관계를 유도하는 것은 가장 핵심이 되는 기술로 많은 알고리즘이 제안되어 왔다. 그러나 기존의 알고리즘은 기하학적인 특징이나 영상의 화소값을 이용하고, 최적화된 관계식을 유도하기 위해 반복연산을 수행하고 있어 많은 계산량을 요구하고 있다. 본 연구에서는 특징 모델을 기반한 정합을 이용하여 영상을 모자이크를 구현하고 구매펡을 통해 시각에 따라 변형되어 몰입감을 주는 영상기반 환경을 생성하였다. 기존의 방법으로는 수행하기 어려운 카메라의 자유로운 회전에서도 변함이 없는 특징을 사용하여 특징점을 정합하였다. 이에 특징의 유사성으로 인한 오차를 줄이고 정확한 일대일 대응 관계를 구하기 위해서 중간 벡터 필터링 방법과 회전 상관 관계를 이용하여 최종의 특징점을 추출하였으며 이를 실험으로 구현, 검증하였다. 또한 가우시안 분포가중치에 기반한 영상 합성을 이용함으로써 보다 자연스러운 합성을 구현하였다.

참 고 문 헌

[1] I. Zoghلامي, O. Faugeras, and R. Deriche, "Using geometric corners to build a 2D mosaic from a set of images", In Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pages 420-425, 1997.

[2] B. Srinivasa Reddy, B.N. Chatterji, "An FFT-Based Technique for Translation, Rotation, and Scale-Invariant image Registration", IEEE Trans. on Image Processing, Vol. 5, No.8, pages. 1266-1271, August 1996.

[3] R. Szeliski, "Image mosaicing for tele-reality applications", Technical Report CXRL94/2,DEC-CRL, May 1994.

[4] T.J. Cham and R. Cipolla, "A Statistical Framework for Long-Range Feature Matching in Uncalibrated Image Mosaicing", In Proc. of Conference on Computer Vision and Pattern Recognition, pages 442-447, 1998.

[5] F. Mokhtarian and R. Suomela, "Robust Image Corner Detection Through Curvature Scale Space", IEEE Trans. on Pattern Analysis and machine Intelligence, Vol. 20, No.12, December 1998.

[6] P. Smith, D. Sinclair, R. Cipolla and K. Wood, "Effective Corner Matching", British Machine Vision Conference, Vol II, pages 545-556, September 1998.

[7] A. Redert, E. Hendriks, and J. Biemond, "3D scene reconstruction with viewpoint adaptation on stereo displays,"h IEEE Trans. On Circuits and Systems for Video Technology, vol.10, no.4, pp. 550-562, June 2000.

[8] T. Kanade, P. J. Narayanan, and P. W. Rander, "Virtualized reality: concepts and early results,"hProc. of IEEE Workshop on Representation of Visual Scenes, pp. 69-76, 1995.



김 태 은

1989년 : 중앙대학교 전기공학과 졸업 (약전전공)

1992년 : 중앙대학교 대학원 전자공학과 졸업(석사)

1997년 : 중앙대학교 대학원 전자공학과 졸업(박사)

1993년~1995년 : 한국과학재단 참여연구원

1995년 : 삼성전자 휴먼테크 논문대상 은상수상

1997년~현재 : 남서울대학교 멀티미디어학과 교수

관심분야 : 증강현실, 영상인식, 컴퓨터비전