

3차원 데이터베이스 구축을 위한 시점 합성

View Synthesis for 3D Database

*이종원^o, *이광연, *김성대

*한국과학기술원 전기및전자공학과

*Jong-Won Yi^o, *Kwang-Yeon Rhee, *Seong-Dae Kim

*Dept. of Elec. Eng. KAIST

요약

시점 합성(View synthesis)이란 일반적으로 주어진 기준 영상들(reference viewpoint images 혹은 key frames)으로부터 임의의 시점에서 바라본 영상을 예측하여 합성하는 것을 의미한다. 이러한 시점 합성은 영상 기반 렌더링(Image-based rendering) 기법의 일종으로 컴퓨터 그래픽스 분야에서 실시간 렌더링이나 가상 현실(Virtual reality) 구현을 위한 매우 강력한 접근 방식이다. 시점 합성은 물체를 표현하기 위한 3차원 모델링이 필요가 없으며, 렌더링 시간이 물체의 복잡도와는 무관한 장점이 있다.

이러한 시점 합성 기법으로 물체의 3차원 표현과 효과적인 데이터베이스 구축에 응용이 가능하다. 본 논문에서는 물체의 3차원 데이터베이스 구축을 위해 영상 기반 렌더링 기법에 근거한 간단하고, 효과적인 시점 합성 방법을 제안한다.

1. 연구 배경

가상 현실이나 물체의 3차원 데이터베이스 구축, 또는 고속 포스트 렌더링(Fast post-rendering) 시스템과 같은 분야에서 3차원의 실제 물체를 다루는 요구가 점점 늘어나고 있다. 그러나, 3차원 물체를 실제와 같이 표현하

기 위해서는 완벽한 기하학적인 모델링이 요구되고 렌더링에 막대한 계산량과 시간을 요구한다. 이러한 기하학 기반 렌더링(Geometry-based rendering)의 단점을 극복하기 위해 영상 기반 렌더링 기법이 대두되게 되었고, 많은 연구가 활발히 이루어지고 있다.

영상 기반 렌더링에 근거한 시점 합성은 많은 응용 분야를 지닌다. 가상 박물관이나 인터넷 가상 쇼핑과 같은 가상 현실 구현이나 프레임을 증가시키거나 low-latency 시스템을 구현, 물체의 3차원 데이터베이스 구축등이 그것이다. 본 논문에서는 3차원 데이터베이스 구축에 초점을 맞추어 적합한 시점 합성 기법에 대하여 서술한다. 기존의 시점 합성 기법들은 크게 영상 워핑(Image Warping)에 의한 방법들[1][2][3][4][5], lumigraph[7][8]나 light field 렌더링[6] 기법들, Faugeras의 기법[9]들로 나눌 수 있다. 그러나, 영상 워핑 기반 기법들이나 Faugeras의 기법은 각 화소간의 정확한 대응 관계를 알고 있어야 하고, 화소들의 depth 정보들도 알아야 되므로 일반적으로 적용이 힘들고, 기준 영상과 함께 변이 지도(disparity map)를 저장하기 때문에 기준 영상이외에도 저장해야 하는 정보량이 많아 데이터베이스 구축에 적합하지 않다. 또한, lumigraph나 light field 렌더링은 화소 대응 관계나 depth 정보를 알 필요가 없는 대신에 상당히 많은 기준 영상

을 저장해야 한다는 단점이 있다.

본 논문에서는 3차원 데이터베이스 구축을 고려하여, 3차원 움직임 파라미터 추정과 역방향 워핑(Backward warping)에 의한 간단하고 실용적인 시점 합성 기법을 제안한다.

2. 제안한 시점 합성 기법

제안한 시점 합성은 3차원 움직임 파라미터의 추정과 이 파라미터로부터 다시 계산하여 보정된 변이장을 얻고, 영역을 분할하고 분할된 영역별로 변이장을 affine 변환으로 근사화하여 역방향 워핑을 수행하는 과정으로 구성되어 있다.

2.1. 변이장(Disparity field) 추정

변이장의 계산은 간단하면서 효과적인 이동 블럭 방법(Sliding block method)을 사용하였다. 블럭의 크기는 16×16 으로 설정하였고, 조밀한 변이장을 얻을 수 있게 8 화소씩 이동하면서 계산을 하였다.

2.2. 3차원 움직임 파라미터 추정

3차원 움직임 모델은 다음 식(1)과 같이 8 파라미터 모델을 사용한다. 8 파라미터 모델은 하나의 평면에 대하여 3차원 회전과 평행이동을 근사적으로 표현할 수 있으므로 물체와 카메라가 멀리 떨어져 있다고 가정한다면 실제 3차원 움직임의 움직임을 잘 묘사할 수 있다.

$$\begin{cases} v_X = a_1 + a_2x + a_3y + a_7x^2 + a_8xy \\ v_Y = a_4 + a_5x + a_6y + a_7xy + a_8y^2 \end{cases} \quad (1)$$

앞 절에서 구한 변이장의 벡터를 $v_X(x_i, y_i)$, $v_Y(x_i, y_i)$, ($i = 0, 1, \dots, N$)이라고 하면 N 은 8 파라미터 추정에 사용되는 모든 벡터들의 개수를 의미하고, 이 벡터들에 대응하는 영상의

공간좌표를 (x_i, y_i) 라고 한다. 3차원 움직임 파라미터에 의해 추정된 각 블럭의 속도 벡터 $\hat{v}_X(x_i, y_i)$, $\hat{v}_Y(x_i, y_i)$ 가 직접 구한 변이장 벡터 $v_X(x_i, y_i)$, $v_Y(x_i, y_i)$ 에 가장 잘 정합이 되게 파라미터를 추정한다.

몇몇 개의 outlier 벡터들을 제외하면 잡음은 모두 평균이 0인 가우시안 분포라고 가정할 수 있으며 Gauss-Markov 정리에 의하여 least-squares 추정은 maximum likelihood 의미에서 바이어스되지 않은 최소의 분산(unbiased and minimum variance)을 가지는 최적 추정기(optimum estimator)가 된다. 이러한 방법을 linear regression이라고 한다. 다음과 같은 오차 함수를 정의하고 이 정의된 오차 함수가 최소가 되도록 파라미터를 구한다.

$$E(\mathbf{a}) = \sum_{i=1}^N \left\{ [v_X(x_i, y_i) - \hat{v}_X(x_i, y_i)]^2 + [v_Y(x_i, y_i) - \hat{v}_Y(x_i, y_i)]^2 \right\} \quad (2)$$

식(2)에 식(1)을 대입하여 각 파라미터에 대하여 편미분하여 그 값을 0으로 두어 다음과 같이 구한다. 즉, $\partial E(\mathbf{a})/\partial a_i = 0$, ($i = 1, 2, \dots, 8$)이 된다. 편미분한 식들을 잘 정리하여 방정식을 풀면 원하는 파라미터의 값을 얻을 수 있다. 이러한 과정을 outlier들을 제거하면서 파라미터의 값들이 수렴할 때까지 반복 계산한다.

2.3. 변이장 후처리

올바른 시점 합성을 위해서는, 극부적으로 잡음 성분이 섞여 있는 변이 벡터보다는 물체의 전체적인 3차원 움직임을 대표할 수 있는 3차원 움직임 파라미터를 이용하는 것이 바람직하다. 이를 위해 앞 절에서 구한 3차원 움직임 파라미터로부터 새로운 변이장 벡터들을 계산하여 원래 값을 대체하고, 새로 대체된 벡터들을 이용하여 역방향 워핑을 수행한다.

2.4. 영역 분할과 역방향 워핑

합성될 시점에서의 어떤 위치에서의 화소 값을 추정하기 위해서는 기준 영상에서의 어떤 화소에 대응되는지 알 필요가 있다. 그러나, 8 파라미터 모델에서는 역변환에서의 해를 구하는 것이 쉽지 않다. 그림1에서 보듯이 물체를 32 × 32 크기의 블록으로 영역 분할하고, 각각의 분할된 영역은 충분히 6 파라미터의 affine 변환으로 선형 근사화할 수 있기 때문에 쉽게 역방향 워핑이 가능하게 된다.

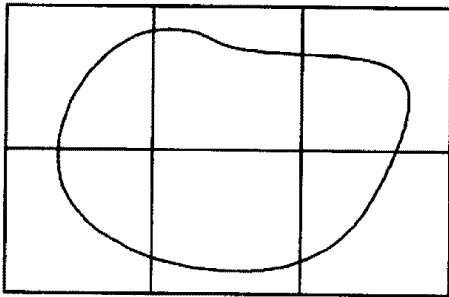


그림 1: 일정한 크기의 블록 형태의 영역 분할

그림2는 하나의 영역에 대하여 근사화한 affine 변환을 도시한 것이고, 식(4)은 근사화된 변환을 이용한 역방향 워핑의 역변환 식이다.

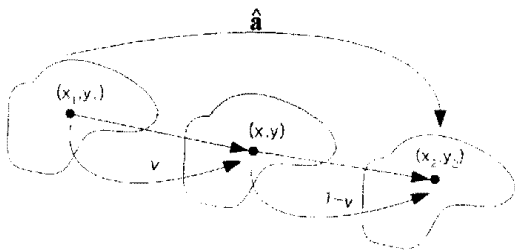


그림 2: 영역별 역방향 워핑

$$\begin{cases} x_2 = \hat{a}_1 x_1 + \hat{a}_2 y_1 + \hat{a}_3 \\ y_2 = \hat{a}_4 x_1 + \hat{a}_5 y_1 + \hat{a}_6 \end{cases} \quad (3)$$

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} v \cdot (\hat{a}_1 - 1) + 1 & v \cdot \hat{a}_2 \\ v \cdot \hat{a}_4 & v \cdot (\hat{a}_5 - 1) + 1 \end{pmatrix}^{-1} \begin{pmatrix} x - v \cdot \hat{a}_3 \\ y - v \cdot \hat{a}_6 \end{pmatrix} \quad (4)$$

식(4)를 이용하여 각 영역 별로 역방향 워핑을 하여 시점 합성된 영상을 얻는다.

3. 모의 실험

그림3, 그림4, 그림5는 모의 실험에 사용된 256 × 256 크기의 영상들이다. 각각 140°, 150°, 160° 각도에서 바라본 영상들이며 140°와 160°를 가지고 150° 영상을 합성한다.



그림 3: 테스트 영상 #1

그림6과 그림7은 각각 이동 블록 기법으로 직접 계산된 변이장과 이로부터 추정된 8 파라미터로 대체되어 새롭게 얻어진 변이장을 보여준다.

그림8은 지금까지의 결과를 가지고 150°에서 바라본 영상을 시점 합성한 결과이다. 실제 영상과 비교해 보았을 때 상당히 유사한 합성 영상을 얻을 수 있음을 알 수 있다.



그림 4: 테스트 영상 #2

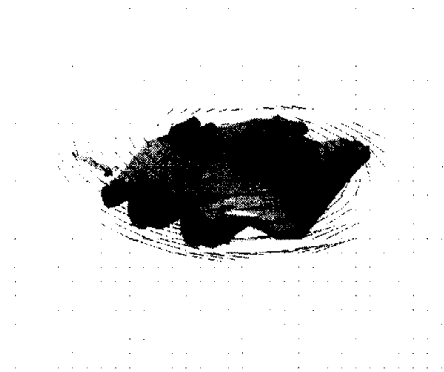


그림 7: 8 파라미터로 수정된 변이장



그림 5: 테스트 영상 #3

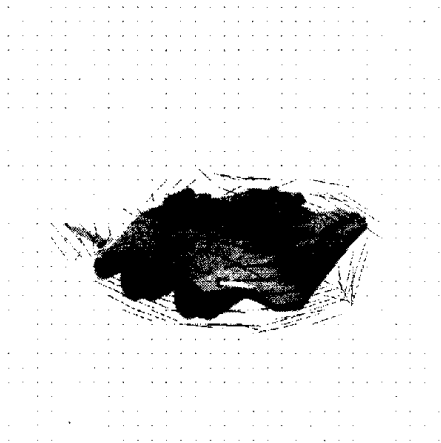


그림 6: 이동 블럭 기법으로 계산된 변이장



그림 8: 시점 합성된 결과 영상

4. 결론

본 논문에서는 간단하면서도 실제적인 적용이 쉬운 시점 합성 기법을 제안하였다. 화소간의 정확한 대응 관계와 depth 정보가 없어도, 3차원 움직임 파라미터 추정과 역방향 워핑을 이용하여 실제와 가까운 시점 합성 영상을 얻을 수 있었다.

또한, 제안한 기법은 변이 지도등을 저장할 필요가 없이 기존 영상들과 8개로 이루어진 3차원 움직임 파라미터들만 양자화하여 저장하므로 물체의 3차원 데이터베이스 구축에 적용이 용이하다.

참고 문헌

- [1] Andrew Lippman, "Movie-maps : An application of the Optical Videodisc to Computer Graphics," Proc. SIGGRAPH 1980. In *Computer Graphics Proceedings*, Annual Conference Series, 1980, ACM SIGGRAPH, pp.32-43.
- [2] Shenchang Eric Chen and Lance Williams(Apple Computer, Inc.), "View Interpolation for Image Synthesis," Proc. SIGGRAPH 1993. In *Computer Graphics Proceedings*, Annual Conference Series, 1993, ACM SIGGRAPH, pp.279-288.
- [3] Leonard McMillan and Gary Bishop, "Head-tracked Stereoscopic Display using Image Warping," *Proc. of SPIE*, Vol. 2409, pp. 21-30, San Jose, CA, Feb. 1995.
- [4] Shenchang Eric Chen, "QuickTime VR - An Image-based Approach to Virtual Environment Navigation," Proc. SIGGRAPH 1995. In *Computer Graphics Proceedings*, Annual Conference Series, 1995, ACM SIGGRAPH, pp.29-38.
- [5] William R. Mark, Meonard McMillan, and Gary Bishop, "Post-Rendering 3D Warping," Proc. Symposium on Interactive 3D Graphics 1997. pp.7-16.
- [6] Marc Levoy and Pat Hanrahan, "Light Field Rendering." Proc. SIGGRAPH 1996. In *Computer Graphics Proceedings*, Annual Conference Series, 1993, ACM SIGGRAPH, pp.31-42.
- [7] Microsoft Research, "The Lumigraph," Proc. SIGGRAPH 1996. In *Computer Graphics Proceedings*, Annual Conference Series, 1996, ACM SIGGRAPH, pp.43-54.
- [8] Peter-Pike Sloan and Michael F. Cohen, "Time Critical Lumigraph Rendering," Proc. Symposium on Interactive 3D Graphics 1997, pp.17-23.
- [9] Stéphane Laveau and Olivier Faugeras, "3D Scene Representation as a Collection of Images and Fundamental Matrices," Technical Report N2205, INRIA, Feb. 1994.