

논문 2005-42TE-1-10

넓은 가상환경 구축을 위한 다수의 TIP (Tour into the Picture) 영상 합성

(Multiple TIP Images Blending for Wide Virtual Environment)

노 창 현*, 이 완 복*, 류 대 현**, 강 정 진***

(Chang Hyun Roh, Wan Bok Lee, Daehyun Ryu, and Jung Jin Kang)

요 약

영상 기반 렌더링은 장면을 3차원 기하 모델로 구축할 필요없이 실사 영상만을 이용하여 빠른 속도로 사실감 높은 결과 영상을 렌더링 할 수 있는 기법이다. 특히, '그림 속으로의 여행(Tour into the Picture: TIP)' 기법은 3차원 배경을 매우 간단히 구축할 수 있어 널리 이용되고 있다. 그러나 기존 TIP 기법은 영상 촬영 지점을 많이 벗어나면 원거리 지형 정보 부족으로 인해 시점 이동 범위에 한계가 있다. 이점을 보완하기 위해 본 연구에서는 다수의 TIP 영상을 합성하여 넓은 가상환경을 구축하는 방법을 개발하였다. 제안된 방법은 먼저, 다수의 시점에서 각각 TIP 모델을 구축하고, TIP 모델을 구성하는 전경 물체와 배경 물체를 각각 보간하여 렌더링 함으로써 부드러운 네비게이션 영상을 얻었다. 본 연구는 게임, 차량용 네비게이션 등 다양한 산업 분야에 응용될 수 있다.

Abstract

Image-based rendering is an approach to generate realistic images in real-time without modeling explicit 3D geometry. Especially, owing to its simplicity, TIP(Tour Into the Picture) is preferred to constructing a 3D background scene. Because existing TIP methods have a limitation in that they lack geometrical information, we can not expect a accurate scene if the viewpoint is far from the origin of the TIP. In this paper, we propose the method of constructing a virtual environment of a wide area by blending multiple TIP images. Firstly, we construct multiple TIP models of the virtual environment. Then we interpolate foreground and background objects respectively, to generate a smooth navigation image. The method proposed here can be applied to various industry applications, such as computer game, 3D car navigation, and so on.

Keywords : TIP, virtual environment, Image based rendering, Image blending

I. 서 론

3차원 컴퓨터 그래픽에서 일반적으로 사용되는 렌더링 기법은 크게 2가지로 나눌 수 있다. 시간이 많이 걸

리더라도 마치 사진과 같은 정확한 영상을 얻어내는 포토 리얼리스틱 렌더링(photo-realistic rendering)과 전경(scene)을 조금 단순화하여 부정확한 영상을 얻어내더라도 실시간으로 빠른 렌더링이 가능하게 하는 실시간 렌더링(real time rendering)이 그것이다^[1].

실시간 렌더링 분야는 그래픽 하드웨어 가격이 고가였기 때문에 상대적으로 관심을 받지 못하고 있었다. 그러나 최근 OpenGL이나 Direct3D 등의 3D 그래픽 라이브러리를 가속하는 하드웨어가 저렴한 가격에 널리 보급되면서 실시간 렌더링 분야가 각광 받고 있다.

* 정회원, 중부대학교
(Joongbu University)

** 정회원, 한세대학교
(Hansei University)

*** 정회원, 동서울대학
(Dong Seoul College)

접수일자: 2004년10월16일, 수정완료일: 2004년11월30일

특히 영상기반 렌더링 분야가 많은 각광을 받고 있는데 이 기법은 환경 또는 물체를 담은 참조 영상들로부터 얻은 정보를 바탕으로 임의의 시점에서 그 환경이나 물체를 볼 때의 영상을 얻는다. 기존의 렌더링 기법과는 달리 표현하고자 하는 장면의 3D 모델링 과정이 필요 없으며, 출력 영상의 생성 속도가 대상물의 기하학적 복잡도와는 상관없이 참조 영상의 해상도에 의해 결정되므로 영상의 실시간 렌더링이 가능하다. 또한, 참조 영상과 같은 수준의 사실적인 영상을 생성할 수 있다. 그러므로 현실감 있는 3D 영상을 얻을 수 있게 된다.

Chen 등은 여러장의 파노라마 영상을 이용하여 가상 환경을 네비게이션(navigation)하는 Quicktime VR 기법을 제안하였다^[2]. McMillan 등은 플레놉틱 함수(pleoptic function)를 최초로 정의하여 영상기반렌더링의 이론적 기반을 세웠다^[3]. Gortler 등^[4]과 Levoy 등^[5]은 복잡한 플레놉틱 함수를 4차원 함수로 간략화하였고 털과 같은 물체의 사실적인 렌더링이 가능함을 보였다.

Horry등은 다수의 참조 영상을 입력으로 하는 기존의 영상 기반 렌더링 기법들과 달리, 비교적 간단하면서도 실용적인 접근 방법으로서, 주어진 하나의 영상 속을 여행할 수 있는 그림 속으로의 여행(Tour into the Picture:TIP)" 기법을 제안하였다^[6]. Horry의 TIP 기법은 참조 영상이 하나의 소실점을 갖는 경우로 국한되어 있어서 이후 소실선을 이용하여 가상 배경을 구성하도록 Kang 등에 의해 확장되었다^[7].

그러나 기존 TIP 기법으로 구축된 가상환경을 네비게이션 할 때 영상을 취득한 지점을 많이 벗어나면 원거리 지형 정보 부족으로 인해 시점 이동 범위에 한계가 있다. 이점을 보완하기 위해 본 연구에서는 여러 시점에 촬영된 다수의 TIP 영상을 합성하여 넓은 가상환경을 구축하는 방법을 개발하였다.

제안된 방법은 먼저, 다수의 시점에서 촬영된 영상을 Kang 등이 제안한 모델을 이용하여 각각 TIP 모델을 구축하고, TIP 모델을 구성하는 전경 물체와 배경 물체를 각각 보간하여 렌더링 함으로써 부드러운 네비게이션 영상을 얻는다.

II. 그림속으로의 여행(Tour into the Picture: TIP)

1. 소실점을 이용한 TIP

Horry^[6]등은 한 장의 입력 영상을 이용하여 단순화된 3D 모델을 구성하여 시점의 변화에 따른 새로운 영상

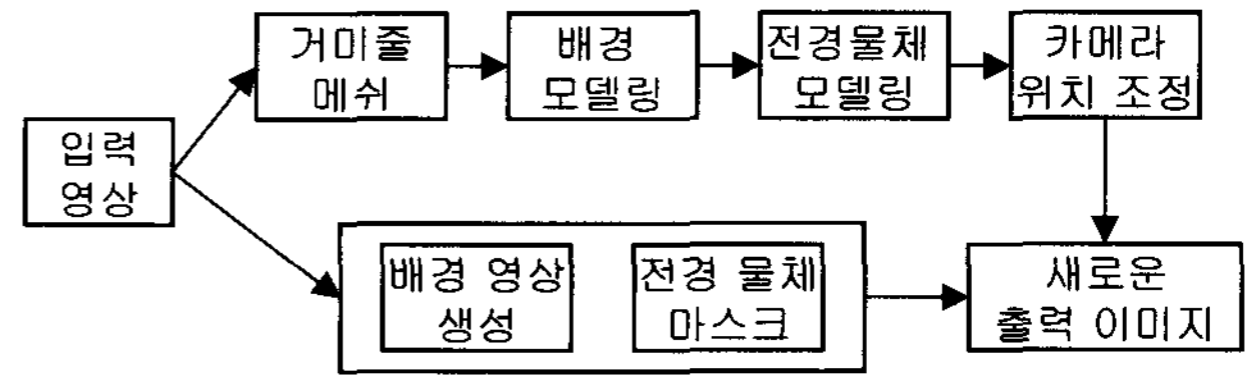


그림 1. 소실점을 이용한 TIP 환경 구축
Fig. 1. TIP environment construction using a vanishing point.

을 생성할 수 있는 방법을 제안하였다. 이 방법은 여러장의 입력 이미지들을 필요로 하는 일반적인 컴퓨터 비전 기법들 대신, 2D 이미지의 소실점 위치로부터 배경의 3D 정보를 복원하여, 한 장의 이미지만으로 그림이나 사진 속으로의 전진이나 비행 같은 시점 이동의 애니메이션을 구현한 것이다.

그림 1은 Horry^[6] 등이 제안한 방법의 TIP 환경 구축 과정이다. 한 장의 입력 영상에서 전경 물체를 추출하고, 최초의 입력 영상을 참조하여 전경 물체 추출로 인해 생긴 여백을 채워넣어 배경 영상을 얻는다. 때어낸 전경 물체의 영상을 저장하고 전경 물체 마스크(mask)를 구성한다. 그리고 입력 영상에서 소실점을 찾아 거미줄 메쉬(spiderly mesh)를 그린다. 거미줄 메쉬에서 바닥, 천정, 벽 등을 구분하여 배경을 모델링 한 후 완성된 배경 모델속에 전경 물체를 배치시켜 전경 물체를 모델링 한다. 이렇게 완성된 가상 공간을 카메라의 위치 변화에 따라 새롭게 렌더링된 결과를 생성할 수 있게 된다.

2. 소실선을 이용한 TIP

Horry^[6]등의 연구는 무엇보다도 한 장의 입력 영상만으로 애니메이션을 구현하는 간단한 방법을 제시했다는 데 큰 의의가 있다. 그러나 참조 영상이 하나의 소실점을 갖는 경우로 국한되어 있어서 소실점이 여러 개이거나 소실점이 분명하지 않거나 없는 경우에 적용하는데 어려움이 있다. 이러한 단점을 해결하기 위해 Kang 등은 소실선을 이용한 TIP모델을 제안하였다^[7].

그림 2는 소실선을 이용한 배경 모델을 설명한 것이다. 그림에서 배경 평면을 무한대에 위치한 평면으로 가정하면 카메라 영상에 나타난 소실선을 기준으로 상부에 위치한 영상은 배경 평면에 대응하고 하부에 위치한 영상은 바닥 평면에 대응하게 된다. 그러므로 한 장의 영상이 주어졌을 때 이를 카메라 영상 평면으로 하여 사영기하학을 적용하면 배경 평면과 바닥 평면으로

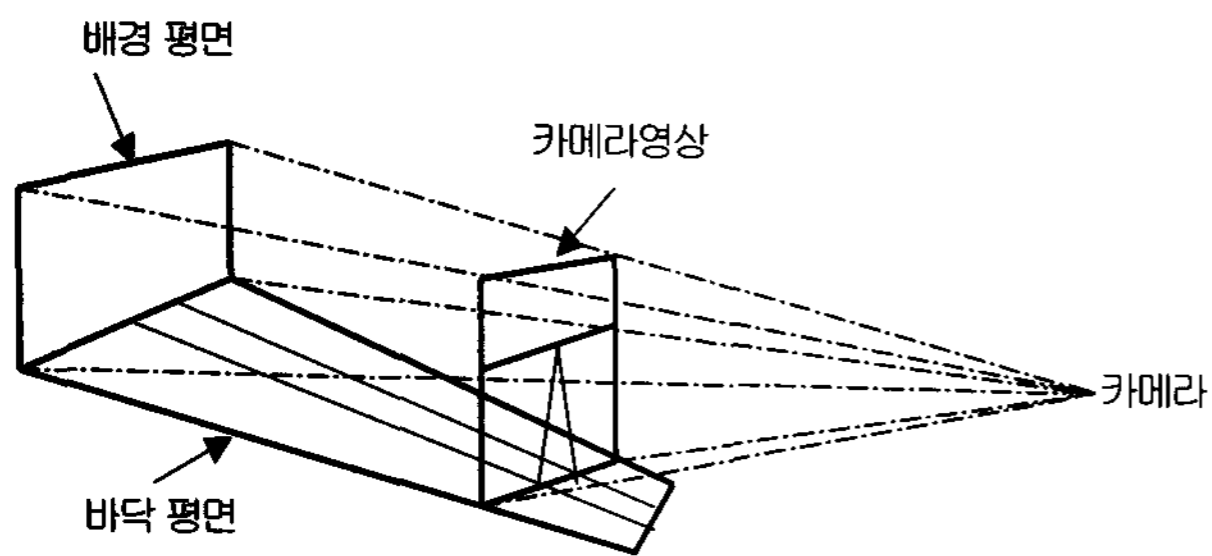


그림 2. 소실선을 이용한 배경 모델
Fig. 2. A background model using a vanishing line.

구성된 배경 모델이 완성된다. Horry가 제안했던 모델은 배경을 이루는 평면이 최고 5개로 구성되었으나 이 모델은 바닥과 배경 평면 2개로 간단히 구성된다.

배경 모델과 마찬가지로 카메라 영상 평면에 존재하는 전경 물체들도 사영기하학 이론을 적용하여 바닥 평면에 대응하는 위치를 찾을 수 있으므로 전경 물체 모델링 또한 가능하다.

파노라마 영상에서 카메라로부터 보여지는 배경은 카메라 위치를 원점으로 하는 구에 매핑(mapping)될 수 있기 때문에 이들은 이 방법을 파노라마 영상에도 적용하였다.

소실선을 이용한 TIP 기법에서 재연된 환경 모델은 간단한 원리에 의해 구성되면서도 매우 사실적인 장면을 생성할 수 있고 소실점의 유무 및 개수에 제한받지 않는 장점이 있다.

III. 다수의 TIP 영상 합성

1. 다수의 TIP 영상을 통한 가상환경 구축 방법

기존 TIP 모델은 한 장의 영상을 취득한 후 이를 기반으로 가상환경을 구축하고 이 환경을 자유롭게 네비게이션하는 모델이다. 그런데 영상이 취득한 시점 부근에서는 화질이 선명하나 이로부터 멀어지면 원거리지형에 대한 정보 부족으로 영상의 품질이 매우 낮다. 그러므로 3차원 관광 사이트, 카네비게이션 등과 같이 넓은 가상환경을 네비게이션해야 하는 분야에 응용되기에는 부적합하다.

이러한 문제를 해결하기 위해 본 연구에서는 그림 3과 같은 다수의 TIP 영상을 합성하여 넓은 영역을 자유롭게 네비게이션 할 수 있는 방법을 제안한다.

그림 3에서 보는 바와 같이 도로 또는 공원 등과 같은 넓은 영역에서 골고루 여러 지점을 선택하고 각 지

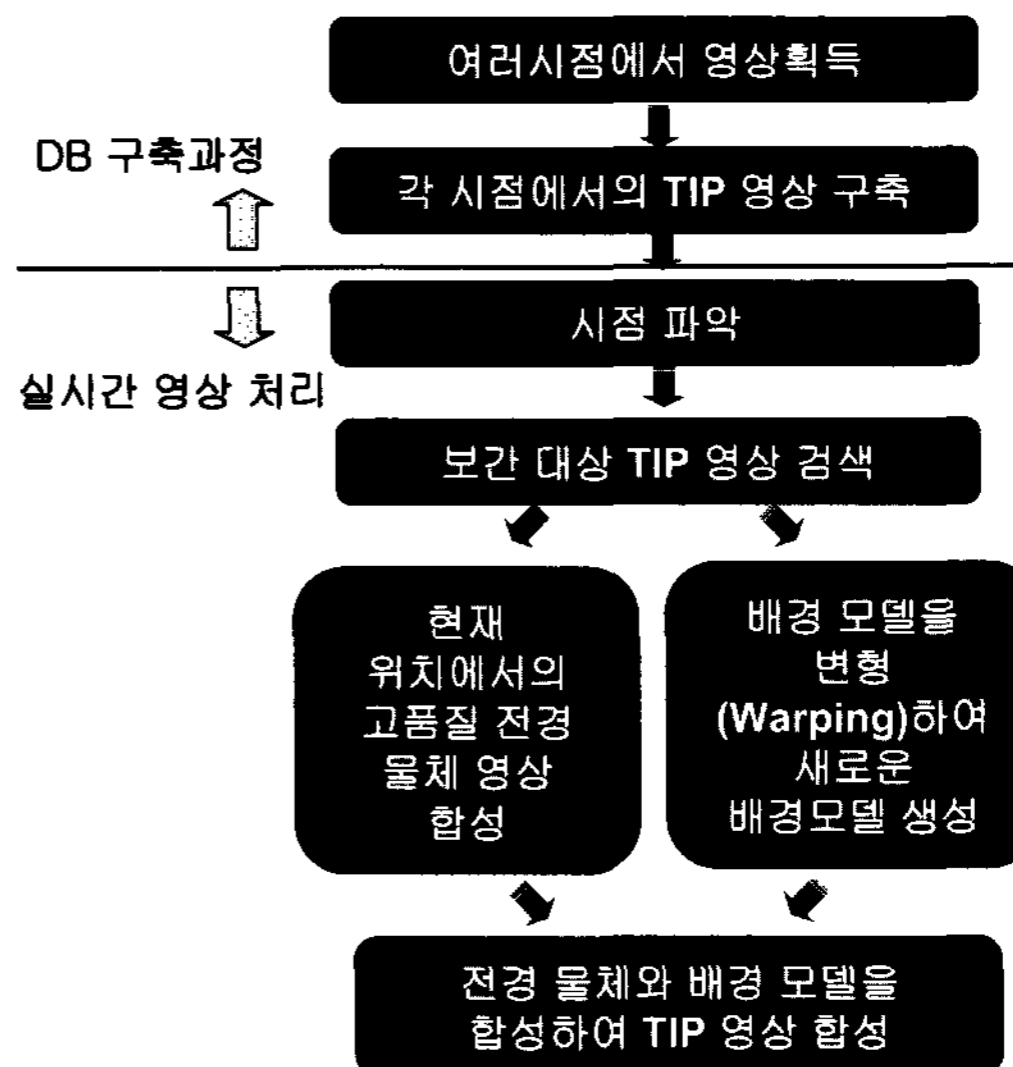


그림 3. 다수의 TIP 영상 합성을 통한 가상환경 구축 과정

Fig. 3. Procedure of virtual environment construction using multiple TIP images blending.

점마다 파노라마 영상을 취득한다. 그리고 Kang 등의 연구에서 수행된 것처럼 각 시점에서 전경물체와 배경영상을 분리하여 각 지점별로 TIP 환경을 구축한다. 넓은 영역에서 여러 시점의 영상을 취득함으로써 동일 전경물체가 각 TIP 환경에 나타나므로 이들 전경물체간의 대응관계를 설정해 주어야 한다.

이제 가상환경을 실시간 네비게이션하고자 할때 먼저 현재 시점을 파악하여 TIP 영상 들 중 현재 위치에 가까운 보간 대상 영상을 파악한다. 이후 보간 대상 영상간의 전경물체를 합성하여 새로운 전경물체를 생성하고 배경모델을 변형(Warping)하여 새로운 배경모델을 생성하여 이를 합성하여 새로운 영상을 실시간 출력하여 준다. 이런 방법을 통해 넓은 영역의 가상환경을 자유롭게 실시간 이동하더라도 고 품질의 영상을 보장할 수 있게 된다.

다수의 TIP 영상 합성을 통한 가상환경 구축에 필요한 주요 과정을 아래 각 절에서 자세히 설명한다.

2. 영상 획득 위치 추정 및 TIP 영상 검색

현재 위치에서의 결과 영상을 얻기 위해서, 본 연구에서는 가까운 위치에 있는 세 개의 TIP 환경 모델을 이용하였다. 가장 가깝게 놓여있는 세 개의 TIP 환경 모델을 검색하는 것은 n차원 공간에서 가장 가까운 점 k개를 검색하는 k-최근점 문제(k-nearest point

problem)의 특수한 경우이며, 본 연구에서는 쿼드트리(quadtree)를 이용하여 공간을 셀(cell)들로 분할하고, 현재 위치 주변의 셀만을 검색함으로써 빠른 시간에 주변 TIP 환경 모델들을 찾을 수 있었다.

3. 전경 물체간의 대응관계 설정 및 보간

현재 위치에서 전경물체의 형태를 얻기 위하여, 검색된 주변 TIP 환경 모델에 포함된 전경물체들을 다음과 같이 보간한다. 우선 사용자는 세 TIP 환경 모델에 포함된 전경 물체들 중 같은 물체들이 어떤것인지를 지정한다. 같은 물체가 세 환경 모델 모두에 포함되어 있지 않은 경우에는 우선 크기가 0인 가상의 전경 물체를 생성한다. 같은 전경 물체들이 지정되면, 전경 물체를 이루는 정점들의 대응 관계를 설정한다. 서로 다른 TIP 환경 모델에 포함된 전경 물체들이 원과같은 위상을 가지고 있지만, 같은 개수와 형태의 정점을 가지지는 않는다고 가정한다.

본 연구에서는 기존의 영상 모핑(image morphing) 기법을 이용하여 정점 사이의 대응 관계를 설정하고 새로운 위치에서의 전경 물체 형태를 생성하였다. 같은 물체가 세 환경 모델에 모두 포함되지 않아 가상물체가 생성된 경우에는 컴퓨터 비전 분야의 운동 추정(motion estimation) 기법을 이용하여 가상 물체의 형태를 추정하였다.

가상 물체의 형태가 얻어지면, 정점과 픽셀 간의 거리를 이용한 가중합과 영상변형(image warping)을 통하여 새로운 위치에서의 텍스처를 얻어낸다. 세 환경 모델로부터 텍스처 영상이 얻어지면 현재 위치와 환경 모델의 중심 간의 위치의 역에 비례하는 가중치를 설정하여 이들을 픽셀 단위로 합성(blending)함으로써 최종 텍스처 영상을 얻는다.

4. 배경 모델의 변형 및 합성

전경 모델과는 달리, 배경 모델은 TIP 환경 모델의 위치에 관계없이 일정한 형태를 취하고 있다. 따라서 형태의 변형보다도 텍스처 영상 간의 변형 및 합성이 중요한 문제가 된다. 우선 사영기하학을 통하여 배경 모델들을 이루는 픽셀들 간의 대응 관계를 찾는다.

이를 기반으로, 각 배경 모델로부터 새로운 위치에서의 텍스처 영상을 영상 변형을 통하여 생성한다. 각 배경 모델로부터 생성된 배경 영상들을 현재 위치와 환경 모델의 중심 간의 위치의 역에 비례하는 가중치를 설정

하여 이들을 픽셀 단위로 합성(blending)함으로써 최종 텍스처 영상을 얻는다

5. 전경 모델 배치 및 새로운 시점에서의 TIP 영상 합성

최종적으로, 보간에 사용된 전경 물체들의 영상획득 시의 3차원 좌표를 기반으로 하여, 보간된 전경 물체의 3차원 좌표를 사영 기하학을 이용하여 추정하고, 이들을 3차원 배경 모델 위에 배치한다. 전경 물체의 보간을 통하여 얻어진 텍스처 영상은 알파 텍스처(alpha texture)로서 배경 모델과 전경 모델의 부드러운 융합을 위하여 사용되었다.

IV. 실험결과

본 실험은 파노라마에 기반한 TIP 모델을 사용하였으며, Kaidan 사의 360 OneShot VR 시스템을 이용하여 파노라마 영상을 촬영하였다. 촬영 간격은 약 50미터 당 하나였으며 영상 렌더링 시스템은 HP사의 iPaq 3850기종으로, 64MB의 ROM과 64MB의 RAM을 내장하고 있으며 Microsoft Windows CE에 기반하고 있고, 320 x 240의 해상도에서 약 65000 가지의 색깔을 표현할 수 있었다. 개발 환경은 PC 상에서 Microsoft Embedded Toolkit을 이용하여 구축되었다.

그림 4는 실험에 사용된 입력 파노라마 영상을 획득한 위치이며 대전지방기상청 정문에서 한국과학기술원 정문까지 총 23 지점에서 촬영되었다. 각 위치는 붉은색 점으로 표시되었다. 그림 5은 대전 기상청에서 과기원 방향으로 약 50M 간격으로 촬영된 파노라마 영상의

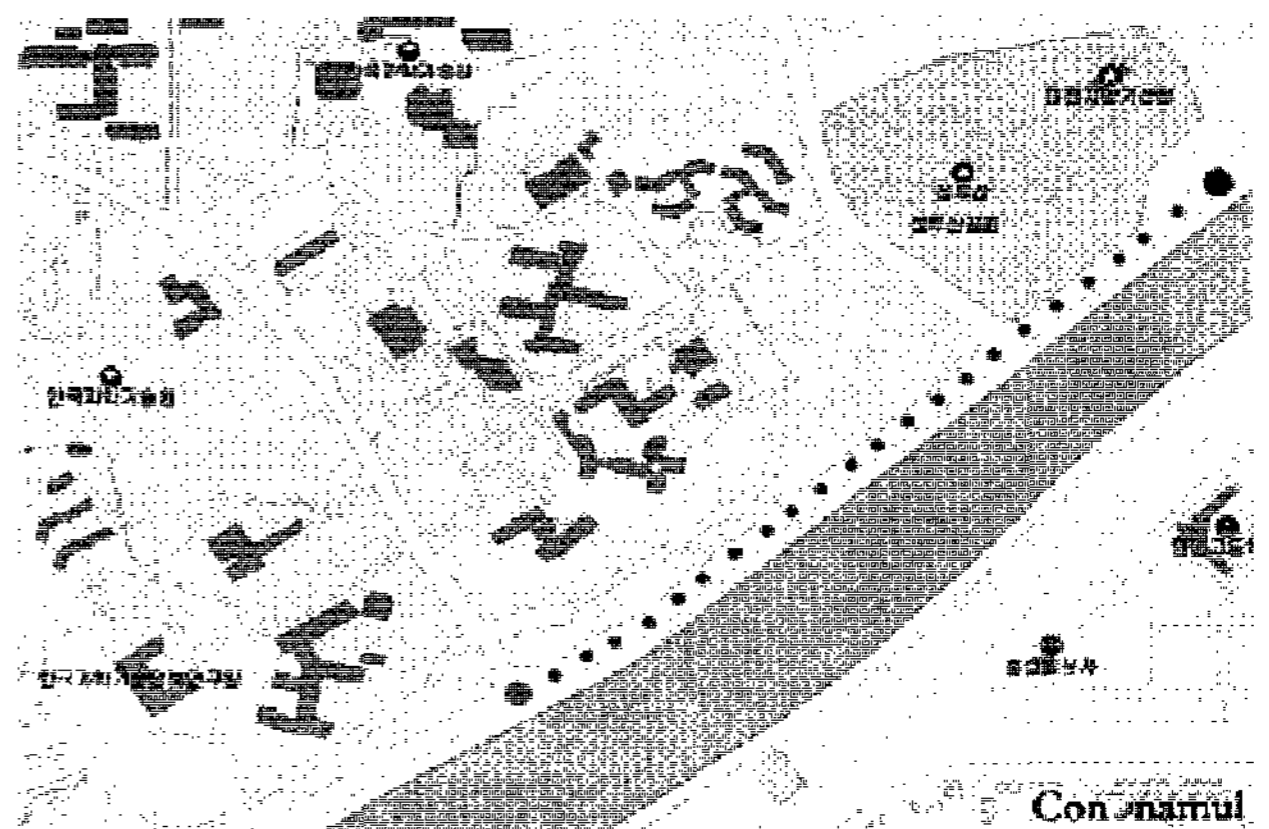


그림 4. 파노라마 영상 취득 위치 (붉은색 점)

Fig. 4. The 23 locations (red dots) for panoramic image acquisition.

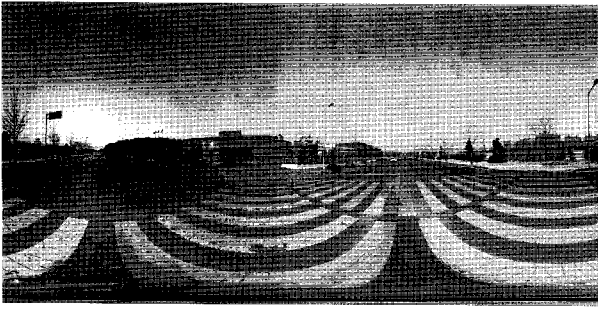


그림 5a. 기상청 앞

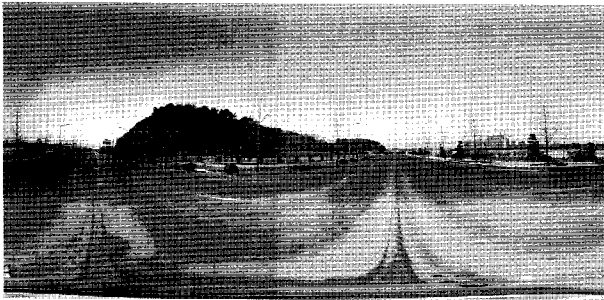


그림 5b. 기상청에서 과기원 방향 50m 지점

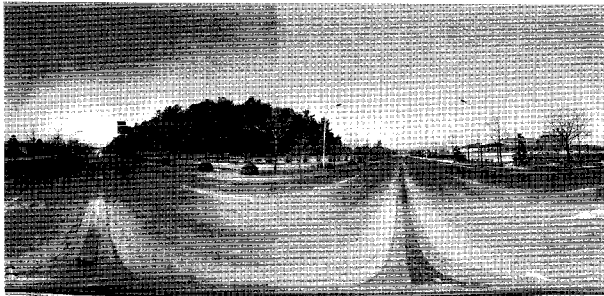


그림 5c. 기상청에서 과기원 방향 100m 지점

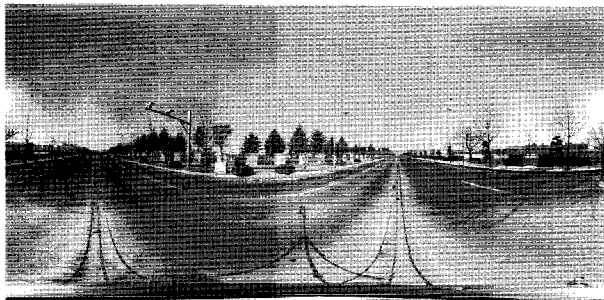


그림 5d. 기상청에서 과기원 방향 150m 지점



그림 5e. 기상청에서 과기원 방향 200m 지점

그림 5. 입력파노라마 영상 샘플
Fig. 5. Input panoramic image samples.

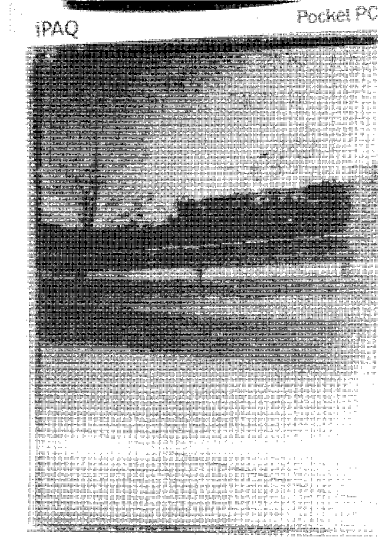


그림 6a. 기상청 앞에서 우측을 바라봄

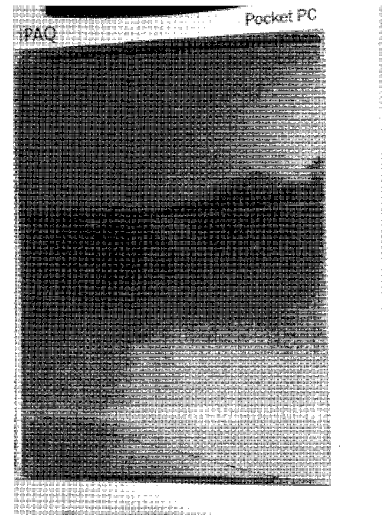


그림 6b. 기상청에서 과기원으로 전진

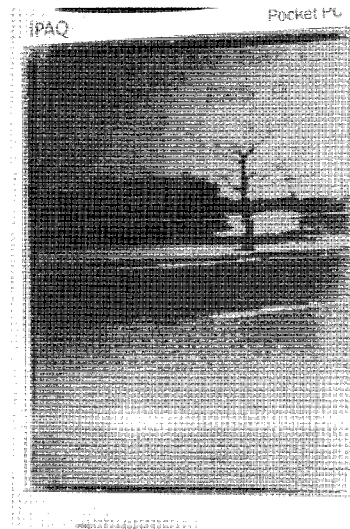


그림 6c. 기상청과 과기원 중간에서 우측 45도

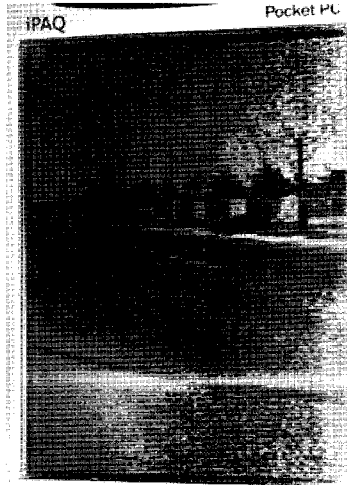


그림 6d. 과기원 옆길

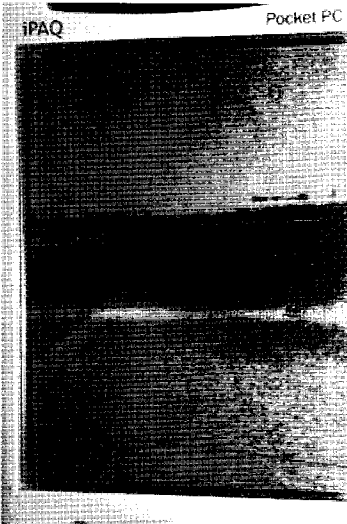


그림 6e. 도로 바닥을 응시

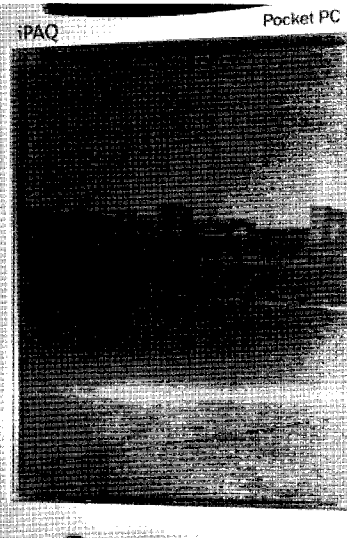


그림 6f. 과기원 정문앞에서 우측 정문을 봄
 그림 6. 가상환경 네비게이션 결과 (PDA)
 Fig. 6. Navigation images for the virtual environment (PDA).

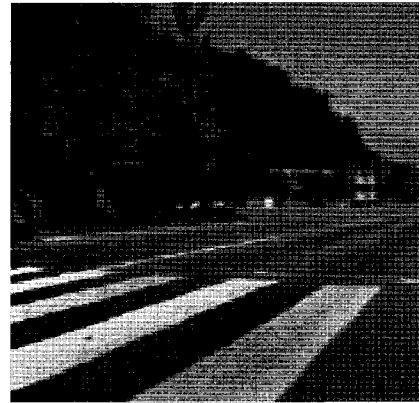


그림 7a. 기상청 앞 횡단보도 좌측 30도

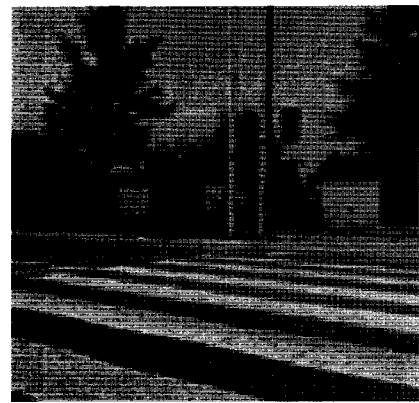


그림 7b. 기상청 앞 횡단보도 우측 30도

그림 7. 가상환경 네비게이션 결과 (PC)
 Fig. 7. Navigation images for the virtual environment (PC).

샘플이며, 그림 6는 대전지방기상청 정문에서 한국과학기술원 까지 이동하며 얻은 PDA 상의 렌더링 결과이다. 그리고 그림 7a), 7b)는 그림 6과 같은 방법으로 PC에서 얻은 결과 영상이다. 각 파노라마 영상의 획득 방향은 길 진행방향과 일치한다.

실험 결과, 320 x 240 크기의 최종 영상을 평균적으로 초당 25 프레임의 속도로 얻을 수 있었다. 영상 생성에 사용되는 주위 TIP 모델이 변경되어 별도의 하드디스크 액세스가 발생할 경우에는 약초당 15프레임의 속도를 얻을 수 있었다.

V. 결론 및 향후 연구

본 연구에서는 기존의 TIP 기법을 보완하여 원거리 지형에서의 시점 이동 범위를 넓혔다. 넓은 면적의 가상 환경을 구축하기 위하여 다수의 시점에서 파노라마 영상을 촬영하였고, 이를 이용하여 다수의 TIP 모델을

구성하였다.

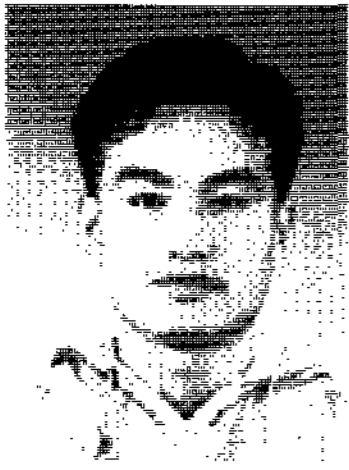
임의의 시점에서의 렌더링 결과를 얻기 위하여 주변의 가장 가까운 세 개의 TIP 모델을 보간하였다. 전경 물체의 경우 물체를 이루는 각각의 정점 위치 및 색상 정보를 보간하여 새로운 전경 물체의 모양과 색상을 얻었으며, 배경 모델의 경우 텍스처 영상을 영상 변형을 통하여 생성함으로써 새로운 배경영상을 얻을 수 있었다.

실제 실험을 통하여 일반 PC에서도 만족할 만한 속도를 얻을 수 있음을 보였고, 부드러운 내비게이션이 가능함을 입증하였다.

참 고 문 헌

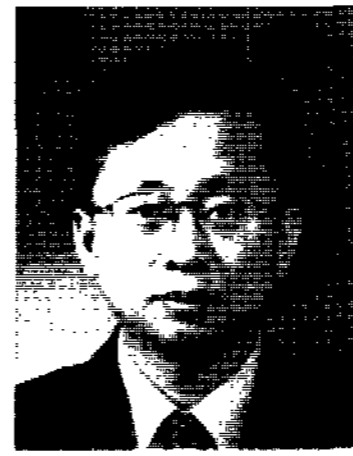
- [1] 임인성, 이중연, "영상기반 렌더링 기법을 이용한 실시간 그림자 생성," 컴퓨터그래픽스학회논문지, 제7권, 1호, 27-35쪽, 2001년 1월
- [2] Chen S. E, "QuickTime VR - An Image-Based Approach to Virtual Environment Navigation", Computer Graphics (SIGGRAPH 95), 29-38, 1995.
- [3] [McMillan95] McMillan L. and Bishop G., "Plenoptic Modeling: An Image-Based Rendering System", Computer Graphics (SIGGRAPH95), 39-46, 1995.
- [4] Gortler S., Grzeszczuk R., Szeliski R., and Cohen M., "The Lumigraph", Computer Graphics (SIGGRAPH 96), 43-54, 1996.
- [5] Levoy M. and Hanrahan P., "Light Field Rendering", Computer Graphics (SIGGRAPH 96), 31-42, 1996.
- [6] Y. Horry, K. Anjyo, and K. Arai. Tour into the picture: using a spidery mesh interface to make animation from single image. In Proc. SIGGRAPH, pages 225--232, 1997.
- [7] Hyung Woo Kang, Soon Hyung Pyo, Ken-ichi Anjyo, Sung Yong Shin: Tour Into the Picture using a Vanishing Line and its Extension to Panoramic Images. Computer Graphics Forum 20(3): (2001).

저 자 소 개



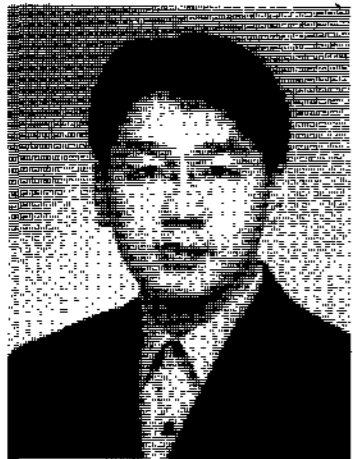
노 창 현(정회원)
 1993년 한국과학기술원 원자력 공학과 학사 졸업
 1995년 한국과학기술원 원자력 공학과 석사
 2001년 한국과학기술원 원자력 공학과 박사

2000년 8월~2003년 6월 (주)에스포라 연구소장
 2002년 3월~현재 한국게임학회 이사
 2002년 3월~현재 중부대학교 게임학과 교수
 <주관심분야 : 컴퓨터게임, VR, Interactive Media>



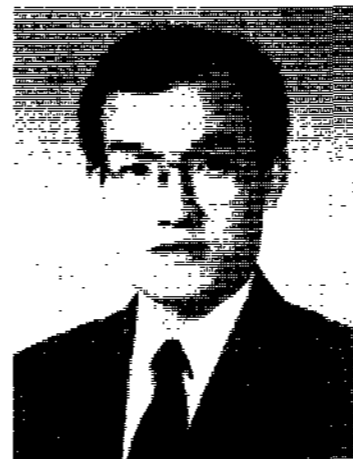
류 대 현(정회원)
 1983년 2월 부산대학교 전기기계 공학과 졸업
 1985년 2월 부산대학교 전자공학과 석사
 1997년 2월 부산대학교 전자공학과 박사

1987년 2월~1998년 2월 한국전자통신연구원 선임연구원
 1998년 3월~현재 한세대학교 IT학부 부교수
 <주관심분야: 정보보호, 영상처리, 통신공학 >



이 완 복(정회원)
 1993년 한국과학기술원 전기 및 전자공학과 학사 졸업
 1995년 한국과학기술원 전기 및 전자공학과 석사 졸업
 2004년 한국과학기술원 전자전산학과 박사 졸업

2000년 (주)시큐어넥서스 개발팀장
 2004년~현재 중부대학교 정보보호학과 전임강사
 <주관심분야 : 정보보호, 컴퓨터 네트워크, 이산사건 시스템, 시뮬레이션>



강 정 진(정회원)
 1991년 2월 건국대학교 대학원 전자공학과(공학박사)
 1991년 3월~2004년 현재 동서울대학 정보통신과 교수 (학과장)

<주관심분야 : RFID 및 Ubiquitous, 이동통신, RF회로, 전파전파 및 Antenna>