

## 비교정 영상에서의 반자동 3차원 건물 모델링

장경호<sup>†</sup>, 장재석<sup>‡</sup>, 이석준<sup>\*\*\*</sup>, 정순기<sup>\*\*\*\*</sup>

### 요 약

본 논문에서는 실외에서 촬영된 비교정 영상으로부터 3차원 건물 구조를 복원하는 반자동화 방법론을 제안한다. 사용자는 관심있는 건물을 임의의 위치에서 촬영한다. 본 논문에서 제안하는 시스템은 먼저, 입력 영상에 대하여 SIFT 알고리즘을 이용하여 특징점과 대응점을 추출한다. 두 번째로, 각 영상에 존재하는 선과 소실점을 추정하고, 추정된 소실점으로 추출한 선들을 그룹화 한다. 다음으로, 각 영상 간의 관계를 대응점의 개수로 정의한 인접 그래프를 이용하여 입력 영상에 대한 순서를 정의하여, 각 영상을 촬영한 카메라의 위치 정보를 보정한다. 최종적으로 추정한 카메라의 정보와 각 영상에서 그룹화된 선을 이용하여 건물의 대략적인 3차원 구조를 복원한다. 하프 에지(half-edge) 자료 구조와 오일러 연산자(Euler operator)를 이용한 상세 모델링을 수행함으로써 완성된 건물 구조를 복원할 수 있다. 본 논문에서는 자동으로 추출된 기하학적 정보를 이용하여 최소한의 사용자 입력으로 건물을 복원 할 수 있도록 하였다.

## Semi-automatic 3D Building Reconstruction from Uncalibrated Images

Kyung Ho Jang<sup>†</sup>, Jae Seok Jang<sup>‡</sup>, Seok Jun Lee<sup>\*\*\*</sup>, Soon Ki Jung<sup>\*\*\*\*</sup>

### ABSTRACT

In this paper, we propose a semi-automatic 3D building reconstruction method using uncalibrated images which includes the facade of target building. First, we extract feature points in all images and find corresponding points between each pair of images. Second, we extract lines on each image and estimate the vanishing points. Extracted lines are grouped with respect to their corresponding vanishing points. The adjacency graph is used to organize the image sequence based on the number of corresponding points between image pairs and camera calibration is performed. The initial solid model can be generated by some user interactions using grouped lines and camera pose information. From initial solid model, a detailed building model is reconstructed by a combination of predefined basic Euler operators on half-edge data structure. Automatically computed geometric information is visualized to help user's interaction during the detail modeling process. The proposed system allow the user to get a 3D building model with less user interaction by augmenting various automatically generated geometric information.

**Key words:** Camera Calibration(카메라 보정), 3D Vision(3D 비전), Building Reconstruction(건물 복원)

※ 교신저자(Corresponding Author) : 정순기, 주소 : 대구광역시 북구 산격동 1370번지(702-701), 전화 : 053)950-5555, FAX : 053)957-4846, E-mail : skjung@knu.ac.kr  
접수일 : 2008년 8월 25일, 수정일 : 2009년 4월 12일  
완료일 : 2009년 6월 11일

<sup>†</sup> 준회원, 경북대학교 전자전기컴퓨터학부 박사과정  
(E-mail : khjang@vr.knu.ac.kr)

<sup>‡</sup> 준회원, 경북대학교 전자전기컴퓨터학부 박사과정  
(E-mail : jsjang@vr.knu.ac.kr)

<sup>\*\*\*</sup> 준회원, 경북대학교 전자전기컴퓨터학부 박사과정  
(E-mail : sukjuni@vr.knu.ac.kr)

<sup>\*\*\*\*</sup> 종신회원, 경북대학교 컴퓨터공학과 부교수

※ 본 연구는 국토해양부가 출연하고 한국건설교통기술평생  
가원에서 관리하며, 지능형국토정보기술혁신사업단에서  
위탁시행한 건설기술연구개발사업의 지원으로 진행되었습니다.  
또한, 본 논문은 Brain Korea 21 정보기술 인력양성  
사업 과제에 의하여 지원되었습니다.

## 1. 서 론

촬영된 영상으로부터 3차원 건물 구조를 복원하는 기술은 컴퓨터 비전에서 중요한 문제이다. 최근에 문화재 복원과 같은 정밀한 복원 작업에 주로 사용되고 있는 레이저 스캐너와 같은 특수 장비[1]는 대상물의 정확한 기하학적 구조를 복원할 수 있으며, 영상과 레이저 스캐너를 동시에 사용하는 혼합 접근법[2]은 더욱더 정확하게 3차원 건물 모델을 자동으로 생성하기에 적합하게 개발되고 있다. 그러나 대부분의 상업적인 장비들은 정확한 정보를 얻기 위해 고가의 장비를 사용해야 하며, 여러 시점에서 얻은 깊이 정보들을 정합하기 위해서는 보다 강인한 알고리즘이 요구된다[1,2].

이에 비해 영상만을 이용하는 컴퓨터 비전 기술 기반의 건물 모델링 방법은 비용적인 면에서 효율적이며 입력 영상으로부터 고해상도의 건물 모델과 텍스쳐 정보를 동시에 얻을 수 있다는 장점을 가진다. 그러나 영상 기반의 건물 모델링 기법을 이용하기 위해서는 정확한 카메라의 위치 정보 보정(camera calibration) 과정이 요구된다.

Debevec[3]은 주어진 입력 영상에서 사용자의 입력에 의해 주어진 직선과 그들 간의 대응점의 관계로부터 입방체(cube), 원뿔(cone)과 같은 그래픽스 기본 객체와 직선 간의 오차를 최소화하여 건물을 모델링하는 Facade 시스템을 제안하였다. 하지만, 카메라의 위치 정보 없이 사용자의 입력에 의해서만 얻어진 직선 선분들을 이용하여 카메라와 모델을 동시에 추정하기 때문에 모델링 시 시간이 길다는 단점이 있다.

Cipolla[4]는 건물을 모델링하기 위해 영상으로부터 추출된 소실점의 정보를 이용하여 건물을 모델링하였다. 이를 위해서는 영상으로부터 최소 세 방향에 대한 소실점의 정보를 정확하게 추출할 수 있어야 한다. 하지만, 대부분의 촬영된 영상에서는 3개의 소실점이 존재하지 않는 경우가 많다는 점이 문제이다.

Photo Modeler[5]에서는 카메라의 위치 보정이 완료된 후, 건물의 실제적인 외부 모습을 복원하기 위해서는, 각 영상에 대한 대응점과 선에 대한 모든 필요한 정보를 입력해 주고, 세밀한 기하학적 모양을 사용자가 일일이 지정해 주어야 건물의 세밀한 부분까지 모델링 할 수 있었다. 이와 같은 방법은 영상에

서 나타나는 특정 정보에 대한 관리를 사용자에게 전적으로 의존하기 때문에, 건물의 창문이나 지붕과 같은 세밀한 부분까지 사용자가 조정해 주어야 하므로 완성된 건물을 모델링하기 위해서 필요한 시간이 비교적 길다는 것이 단점이다.

건물의 3차원 구조를 자동으로 복원하기 위해서는 대부분의 연구자들은 SFM(structure from motion) 알고리즘과 자기 보정(self-calibration) 알고리즘을 사용하였다. 이는 순서가 주어진 입력 영상에서 추출된 대응점의 관계를 이용하여 SFM 알고리즘으로 투영(projective) 공간에서 카메라의 위치 정보를 보정한 후, 자기 보정 알고리즘을 사용하여 메트릭(metric) 공간에서의 카메라의 위치 정보를 보정하는 것이다[3,4]. 만약 카메라의 내부 매개 행렬 정보를 알고 있다면 메트릭 공간상에서 SFM 알고리즘만을 적용함으로써 카메라 위치 정보를 보정할 수 있다[6]. SFM 알고리즘으로 카메라 위치 보정이 완료되면, 대부분의 자동화 기법들은 스테레오(stereo) 기반의 알고리즘을 적용하여 건물의 3차원 구조를 점대점의 관계로부터 구하고, 다수 영상들 간의 관계로부터 구해진 점대점의 관계를 재정의함으로써 조밀한 건물의 3차원 구조를 복원하는 과정을 수행한다. 그러나 영상의 겹쳐짐에 의해 보이지 않는 부분에 대한 깊이 정보를 추출하지 못하는 문제점이 생기며, 점대점의 관계로부터 얻어진 깊이 정보로부터 모델을 생성하기 위해서는 이를 메쉬화하는 단계가 요구된다.

앞서 살펴본 바와 같이 자동으로 건물의 3차원 구조를 복원하고자 할 때는, 깊이 정보를 추출하기 못하거나 조밀한 건물의 3차원 구조를 복원하기 위한 메쉬화 단계가 필요하며, 수동으로 건물을 복원하고자 할 때는, 사용자에 전적으로 의존해야 하기 때문에 복원하는데 긴 시간이 요구된다는 단점들이 있다.

본 논문에서는 이와 같은 단점들을 보완하여 건물을 모델링하기 위한 반자동화 시스템을 제안한다. 또한, 본 논문에서 제안하는 시스템의 입력은 사용자가 복원하고자 하는 건물에 대하여 임의의 위치에서 촬영한 영상으로 순서가 정해져 있지 않다. 카메라의 위치 정보는 자동화가 가능한 전통적인 SFM 기반의 알고리즘을 사용한다. 그러나, 카메라 위치 보정을 위해서는 영상의 순서를 정의할 필요가 있다. 이를 위해 영상을 자동적으로 정렬하기 위한 인접 그래프

개념을 도입하였다.

카메라 보정 후 모델링 과정은 초기 입방 모델 생성과 상세 모델링 과정으로 구분된다. 이를 위해 소실점의 정보와 소실점에 의해 그룹화된 직선 성분을 사용한다. 소실점을 이용한 기준의 방법과 달리 제안하는 시스템은 소실점을 카메라 보정을 위해 사용하기보다는 초기 입방 모델을 생성하기 위해 필요한 직선 성분을 정의하기 위한 매체로서 사용한다. 그러므로 3개의 직교 소실점이 존재하지 않아도 된다. 또한, 다양한 전물의 상세 모델링을 위해 하프 에지(half-edge) 자료 구조를 사용하고, 새로운 오일러 연산자(Euler operator) 집합을 정의한 후 적용하여 사용자의 최소한의 입력으로 건물의 상세 모델링을 가능하게 하였다.

본 논문에서 제안하는 방법은 그림 1의 처리흐름을 가진다.

제안하는 시스템은 크게 특징 정보 추출, 인접 그래프 생성, 카메라 위치 정보 보정, 건물 모델링 단계로 나누어진다. 특징 정보 추출 단계에서는 카메라 위치 정보를 보정하기 위한 특징점 및 대응점 추출, 건물의 평면을 결정하기 위한 직선 추출 및 직선 그룹화 과정을 수행한다. 인접 그래프 생성 단계에서는 카메라 위치 정보 보정을 위한 영상의 순서를 부여하며, 카메라 위치 정보 보정 단계에서는 정렬된 영상과 영상에 존재하는 대응점을 이용하여 카메라의 위치 정보를 추정한다. 건물 모델링 단계에서는 카메라의 위치 정보와 그룹화된 직선 성분을 이용하여 사용자의 입력에 의한 모델을 생성한다.

본 논문의 구성은 다음과 같다. 2절에서는 영상의 특징 정보를 추출하기 위한 기반 기술과 인접 그래프 생성 알고리즘에 대해 먼저 정리한다. 3절에서는 본 논문에서 적용한 카메라 위치 정보 보정 기술을 설명

하고, 4절에서는 건물 모델링에 필요한 모델링 정의 과정과 그 적용 예제를 기술한다. 5절에서는 본 논문의 실험 결과를 나타내고, 6절에서 본 논문의 전체적인 내용을 정리한다.

## 2. 특징 정보 추출

제안하는 시스템에서 건물을 모델링하기 위한 특징 정보는 영상으로부터 추출된 소실점(vanishing point), 직선 성분들(lines) 그리고 특징점(feature point)이다. 본 장에서는 이러한 정보를 영상에서 추출하기 위한 방법들을 알아보고 본 논문에 적용한 방법에 관하여 설명한다.

### 2.1 직선 검출

호프 변환(Hough transform)[7]은 영상에 존재하는 직선 성분을 추출하기 위한 표준 도구로써 널리 사용되고 있다. 많은 연구자들은 직선 추출의 정확성을 향상시키기 위해 많은 노력을 해왔으며 그 결과 극 좌표계(polar coordinates)에서 수행되어지는 많은 알고리즘이 존재한다[8,9]. 하지만 그 결과는 영상에 존재하는 직선 성분과 항상 정확하게 대응되지는 않는다. 전통적인 호프 변환과 달리 역호프 변환(inverse hough transform)[10]은 극 좌표공간이 아닌 영상 공간에서 누적(voting) 알고리즘을 적용하여 보다 정확한 직선 성분을 추출할 수 있으며, 그 결과는 항상 영상에 존재하는 직선 성분과 정확하게 대응된다. 본 논문에서는 직선 성분 추출을 위해 영상 공간에서 직선을 추출하는 역호프 변환을 사용하였다[11].

### 2.2 소실점 검출

소실점을 추출하기 위해서는 같은 소실점에 맷히게 되는 직선 성분을 찾아 그룹화하는 처리과정이 요구된다. 소실점의 추출을 위해 가우시안 구(gaussian sphere)를 사용하였으며, 영상위의 직선 성분들은 그림 2와 같이 가우시안 구상에서 대원(great circle)으로 표현된다. 실제로 이러한 그룹화 작업은 비선형(non-linear) 계산이 필요하므로 계산하기 까다로운 문제이다. 이를 해결하기 위해 본 논문에서는 비선형적인 문제를 선형적(linear)인 문제

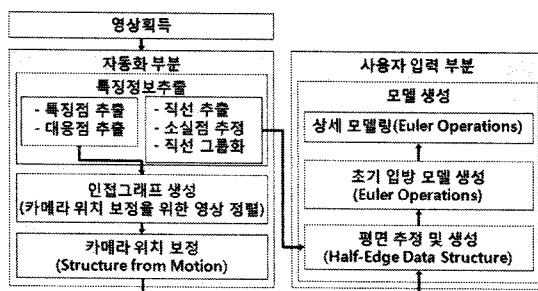


그림 1. 시스템 처리흐름도

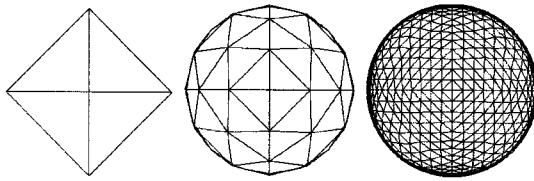


그림 2. 소실점 추출과 직선 그룹화를 위한 가우시안 반구

로 변환하기 위해 일정한 간격으로 나누어진 가우시안 반구(gaussian half sphere)를 사용하였다[12]. 초기에는 그림 3에서 보는 바와 같이 사변형의 피라미드를 생성하여 충분한 해상도를 가질 때까지 피라미드를 쪼개나간다.

주어진 직선성분  $L_{ij}$ 의 시작점  $(x_{1i}, y_{1i})$ 와 끝점  $(x_{2i}, y_{2i})$ 가 주어진다면, 수식 (1)로부터 그림 2와 같이 가우시안 구상에서 대원(great circle)의 법선 벡터  $\bar{n}$ 을 구해낼 수 있다.

$$\bar{n} = \bar{x}_{1i} / \| \bar{x}_{1i} \| \times \bar{x}_{2i} / \| \bar{x}_{2i} \| \quad (1)$$

수식 (1)에서  $\bar{x}_{1i} = (x_{1i} - x_c, y_{1i} - y_c, f)$ ,  $\bar{x}_{2i} = (x_{2i} - x_c, y_{2i} - y_c, f)$ 의 관계를 가지며,  $x_c, y_c, f$ 는 각각 영상의 중점과 렌즈의 초점 거리이며, 소실점 추출 시 카메라의 중점은 영상의 중점으로 카메라 렌즈의 초점 거리는 영상의 높이 정보와 넓이 정보 중 큰 값을 가지는 값으로 사용하였다. 소실점 추출 시 영상의 원점과 렌즈의 초점 거리는 소실점의 위치 결정에 영향을 미치지 않는다[13].

대원의 법선 벡터가 구해진 후에 각 셀에 해당되는 삼각형과 대원이 교차(intersect)되는 부분이 발생하게 될 것이다. 이처럼 대원과 삼각 셀이 교차되는 셀에 대해서는 누적 값을 증가시킨다. 이러한 방

식으로 모든 직선 성분에 대하여 누적(voting)을 수행하였다면, 가장 큰 누적 값을 가지는 셀을 선택함으로써 임의의 소실점을 지나는 직선 성분과 소실점을 손쉽게 계산해 낼 수 있게 된다. 본 논문에서는 보다 정확한 소실점을 추출하기 위해 추가적으로 EM(expectation maximization) 알고리즘을 사용하여 최적의 값을 구해내었다[14]. 소실점의 정보는 카메라 보정을 위해 사용되기 보다는 초기 입방 모델 생성 시 직선 성분을 구하기 위한 용도로서 사용되어진다.

### 2.3 인접 그래프 생성

카메라 위치 정보 추정을 위해서는 무작위로 촬영된 영상을 순서화 필요가 있다. 이를 위해 본 절에서는 인접 그래프를 소개한다. 인접 그래프내의 영상 간 대응 관계를 두 영상간의 대응점의 개수로 정의하였다. 두 영상간의 대응점들은 두 영상에서 판측되는 동일한 점들로, 대응점의 개수가 많으면 두 영상에서 판측되는 동일한 점이 많다는 것이며, 대응점의 개수가 많으면 두 영상의 유사도가 높은 것으로 판단할 수 있다. 또한, 대응점의 개수가 많으면 후에 추정할 카메라의 위치 정보에 대한 오차를 크게 줄일 수 있기 때문에 두 영상간의 관계는 대응점의 개수로 정의하였다. 대응점의 개수로 영상간의 관계를 정의한 인접 그래프는 영상 간의 관계를 효과적으로 표현 할 수 있으며, 영상의 순서를 정의할 수 있기 때문에, 카메라 위치 정보 추정을 쉽게 할 수 있도록 한다.

#### 2.3.1 특징점 추출과 매칭

SFM 기반의 카메라 보정 기술을 적용하기 위해 영상과 영상사이에 대응되는 특징점이 필요하다. 이를 위해 본 논문에서는 SIFT(Scale Invariant Feature Transform)[15] 알고리즘을 이용하여 특징점을 추출하고, 추출된 특징점을 이용하여 영상간의 대응 관계를 정의하였다. SIFT 알고리즘에 의해 구해진 대응 관계는 80%정도의 정확도를 유지하고 있지만, 여전히 오류나 부 대응(mis-match)하는 경우가 동반되므로 잘못된 대응결과도 수반된다. 더구나 본 논문에서와 같이 순서가 없는 영상을 정렬하기 위해서 영상사이 대응점의 개수를 사용하므로 보다 정확한 대응 관계를 정의할 필요가 있다. 이를 위해 본 논문에서는 RANSAC(RANdom SAmples Consensus)[16] 알고리즘을 사용하여 대응점의 개수로 대응 관계를 정의하였다.

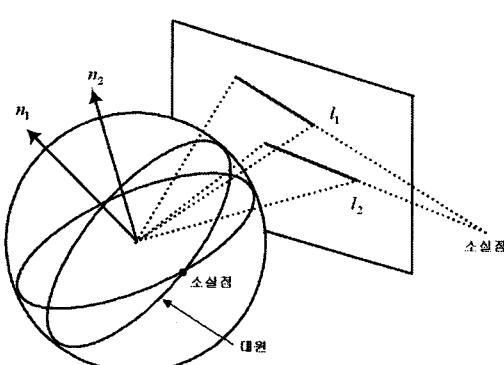


그림 3. 소실점, 직선과 가우시안 구의 관계

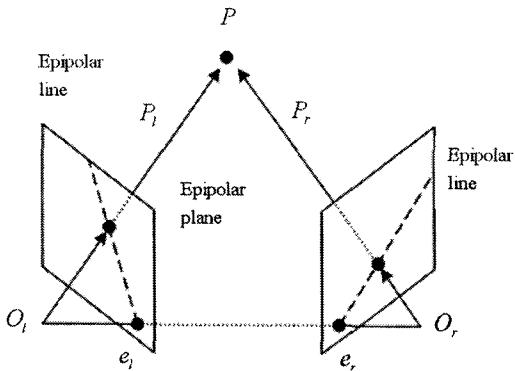


그림 4. 에피풀라 기하학[17]

리즘과 에피풀라 기하학(epipolar geometry)[17]을 이용하여 불필요하고 잘못된 대응 관계를 제거하였다. 그림 4는 에피풀라 기하학의 기본적인 정의를 나타내고 있다.

에피풀라 기하학은 두 영상 사이의 대응점과 각각의 에피풀을 지나는 에피풀라 직선과 그에 의해 정의되어진 에피풀라 평면에 의해 정의되어진다. 이러한 에피풀라 기하학은 기본 행렬(fundamental matrix)  $F$ 에 의해 정의되며,  $F$ 는 수식 (2)와 같이 구해진다. 에피풀  $e_l, e_r$ 은 구해진  $F$ 에 SVD(Singular Value Decomposition)을 적용함으로써 구할 수 있으며 에피풀라 직선은 각각의 대응점에  $F$ 를 곱함으로써 구할 수 있다.

$$p_r^T F p_l = 0, F = \begin{pmatrix} f_1 f_2 f_3 \\ f_4 f_5 f_6 \\ f_7 f_8 f_9 \end{pmatrix} \quad (2)$$

수식 (2)에서  $p_l = (x_l, y_l, 1)$ 과  $p_r = (x_r, y_r, 1)$ 은 대응되는 2차원 점으로 그림 4에서 보는 바와 같이 3차원 상의 점  $P$ 가 영상 평면에 투영된 점이다. 만약 두 개의 대응점이 정확하게 에피풀라 기하학을 만족한다면 에피풀라 직선과 대응되는 점들 간의 수직거리는 0에 가까운 값을 가진다. 그러므로 오차 함수는 수식 (3)과 같이 정의할 수 있다[17].

$$err_i = \left( \frac{1}{e_{rx}^2 + e_{ry}^2} + \frac{1}{e_{lx}^2 + e_{ly}^2} \right) p_r^T F p_l \quad (3)$$

수식 (3)에서  $(e_{rx}, e_{ry})$ 와  $(e_{lx}, e_{ly})$ 는 에피풀라 직선의 매개변수로서  $p_l F$ 과  $p_r F^T$ 에 의해 유도되어진다. 필요한 요소를 제거하기 위해 오차 함수를 수식 (3)과 같이 정의한 후, RANSAC[16] 알고리즘을 사용하였

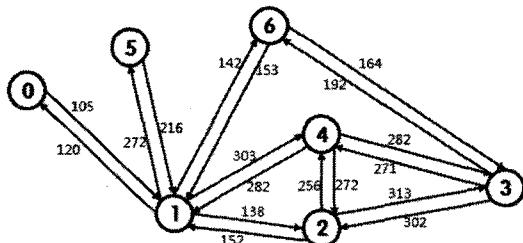
다. RANSAC 알고리즘은 전체의 대응 관계를 이용하기보다는 대응 관계에서 m개의 표본(에피풀라 기하학의 경우 샘플의 크기는 7 또는 8)을 생성하고 각각의 표본에서 구해진 에피풀라 기하학으로 모든 대응 관계에 대해 오차 함수를 적용하여 가장 많은 유효 대응점을 구하는 방법으로 잘못된 대응 관계를 효과적으로 제거하는 방법으로 널리 사용되고 있다. 보통 유효 대응점을 오차함수의 값이 0.5 또는 1화소 차이 내에 있으면 유효한 대응점으로 인식한다. 이를 통해 구해진 유효한 대응 관계는 카메라 보정을 위해 사용되어지며, 두 영상간의 유효한 대응점의 수는 영상의 순서를 결정함에 있어 인접 그래프의 에지의 가중치로 사용되어진다. 만약 유효한 대응점의 수가 주어진 값  $\epsilon_{corr}$ 보다 작다면 두 영상은 관계가 없는 영상으로 인식되어 대응점의 관계가 성립되지 않는다.

### 2.3.2 인접 그래프 생성

인접 그래프는 다수의 영상들의 관계를 효과적으로 표현하여 SFM 알고리즘 적용 시 영상을 손쉽게 다루기 위한 도구로서 사용된다. 입력되는 영상은 각각의 고유 번호가 부여되며, 인접 그래프는 영상의 소실점, 특징점, 대응점을 포함하는 정점(vertex)과 각 정점 간의 관계를 정의한 에지(edge)로 표현된다. 두 정점 사이의 에지는 2.3.1절에서 설명한 에피풀라 기하학으로 정의된 대응점의 개수가  $\epsilon_{corr}$ 보다 많으면 에지가 생성되며, 에지의 가중치는 에피풀라 기하학으로 정의된 대응점의 개수이다. 초기 인접 그래프  $G_I(V, E)$ 는 정점의 집합  $V(G_I)$ 와 에지의 집합  $E(G_I)$ 로 표현된다.

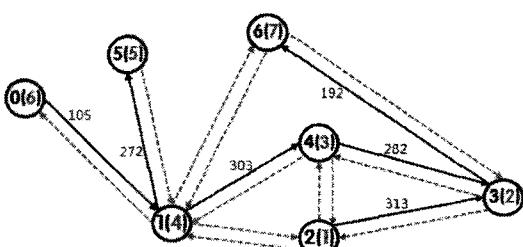
$$\begin{aligned} V(G_I) &= \{I_i | i = 0, 1, \dots, n-1, \text{입력 영상}\} \\ E(G_I) &= \{(I_i, I_j) | i \neq j, \\ &\quad \text{[에피풀라 기하학에 의한 영상간 대응관계]}\} \end{aligned} \quad (4)$$

정점 내의 하나의 대응점은 다수의 영상과 대응 관계를 표현하고 있다. 각 영상간의 관계를 나타내는 에지는 SFM 에지와 묶음(bundle) 에지로 정의된다. SFM 에지는 카메라의 위치 정보를 보정하기 위한 경로를 나타내는 에지이며, 묶음 에지는 묶음 정렬을 위한 경로를 나타내는 에지이다. 에피풀라 기하학으로 대응 관계가 성립된 정점들은 묶음 에지를 가진다. 그림 5는 초기 인접 그래프  $G_I$ 를 나타내며, 정점 내의 번호는 영상에 부여된 고유 번호이다. 정점간의 관계는 묶음 에지로 정의되며, 에지는 방향성을 가지

그림 5. 초기 인접그래프  $G_I$ 

고, 에지의 가중치는 에피폴라 기하학으로 정의된 대응점의 개수이다.

카메라 위치 정보 보정을 위한 최종 인접 그래프  $G_S$ 는 초기 인접 그래프의 하위 그래프(sub-graph)로 표현되며, 이 때 각 영상간의 관계를 나타내는 에지는 SFM 에지이다. 에피폴라 기하학으로 초기 인접 그래프를 생성한 후, 카메라 보정을 수행하기 위해서는 카메라 보정을 위한 최종적인 경로인 신장 트리(spanning tree)를 생성해야 한다. 이를 위해 본 논문에서는 [18]의 탐욕적 방법(Greedy algorithm)을 이용하였다. 또한, 탐욕적 방법 중 크루스칼의 알고리즘(Kruskal's algorithm)을 사용하여 신장 트리를 생성하였다. 그림 6은 카메라 위치 정보 보정을 위한 경로를 설정한 후의 인접 그래프  $G_S$ 를 보여주고 있다. 그림 6에서 점은색으로 표현된 에지가 SFM 에지이며, SFM 에지로 이루어진 정점들과 에지의 집합이 최종 인접 그래프  $G_S$ 로 정의된다. 그림 6에서 정점내의 고유 번호 외의 번호는 카메라 위치 정보 보정을 위한 영상의 입력 순서가 된다. 카메라 위치 정보 보정을 위한 영상의 입력 순서는 에피폴라 기하학으로 정의된 대응점의 개수가 가장 많은 영상을 초기 입력으로 정하고, SFM 에지를 따라 각 정점을 읊어우선 탐색(depth first search)하면서 순서를 부여한다.

그림 6. 최종 인접그래프  $G_S$ 

### 3. 카메라 보정

#### 3.1 카메라 위치 보정

인접 그래프 생성 시 각 영상들은 SIFT와 에피폴라 기하학을 통해 대응 관계가 정립되어 있다. 그러므로 본 절에서는 다수 영상간의 관계를 이용하여 카메라의 위치 정보와 3차원 상의 점을 복원하는 방법에 대하여 설명하도록 한다.

##### 3.1.1 투영 공간(projective space) 구조 복원

영상에 대응하는 3차원 점을 추정해 내기 위해서는 에피폴라 기하학을 이용한 삼각법(triangulation)을 사용된다. 먼저 SFM 에지로 연결되어 있는 임의의 두 장의 영상 중 하나의 영상에 대한 카메라 좌표계를 기준 좌표계로 설정하여 각 영상에 대한 카메라 행렬을 다음과 같이 계산한다[17].

$$P_0 = [I|0], P_1 = [F|e] \quad (5)$$

수식 (5)에서  $F$ 는 초기에 선택한 두 장의 영상에 대한 근본행렬이며,  $e$ 는 에피폴(epipole)을 나타낸다. 초기 3차원 점들은 수식 (6)을 사용하여 계산한다. 이 때, 해를 구하기 위하여 SVD(Singular Value Decomposition)을 사용하였다[17].

$$\begin{pmatrix} P_{02}x_{0k} - P_{00} \\ P_{02}y_{0k} - P_{01} \\ P_{12}x_{1k} - P_{10} \\ P_{12}y_{1k} - P_{11} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix} = 0 \quad (6)$$

수식 (6)에서  $P_{ij}$ 는  $i$ 번째 영상의 카메라 행렬의  $j$ 번째 열(row) 벡터이며,  $(x_{ik}, y_{ik})$ 와  $(x_{jk}, y_{jk})$ 는  $i$ 번째 영상과  $j$ 번째 영상간의 대응점 중  $k$ 번째 대응점의 영상 좌표이고,  $(X, Y, Z, W)^T$ 는 추정하고자 하는 3차원 좌표이다.

초기 두 개의 카메라와 3차원 점이 설정되었다면 다음에 입력으로 들어오는 영상은 2차원 점과 3차원 점간의 대응 관계를 이용하여 수식 (7)을 사용하여 직접적으로 카메라 행렬을 구할 수 있다. 본 논문에서는 해를 구하기 위하여 SVD(Singular Value Decomposition)을 사용하였다[17].

$$\begin{pmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -X_i x_k - Y_i y_k - Z_i z_k - x_k \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -X_i y_k - Y_i y_k - Z_i y_k - y_k \end{pmatrix} \bar{p} = 0 \quad (7)$$

수식 (7)에서  $\bar{p} = (p_0, p_1, \dots, p_{11})^T$ 이며,  $(X_i, Y_i, Z_i)$ 는

3차원 점,  $(x_k, y_k)$ 는 대응되는 2차원 점이다. 이 경우에는 잘못된 대응점의 관계를 없애기 위해 RANSAC 알고리즘[15]을 적용하여 최적의 조건에 만족하는 대응점만 남겨둔다.

카메라 행렬을 복원한 후, 기존의 3차원 점들은 대응되는 모든 점들과 카메라 행렬 정보를 이용하여 재계산되어지며, 3차원 점으로 복원되지 않은 2차원 대응점들에 대하여 삼각법을 이용하여 3차원 점을 계산하여 추가한다.

### 3.1.2 메트릭(metric) 구조 복원

투영 공간상의 구조 복원이 수행되었다면 자기 보정(self-calibration) 알고리즘[17]을 사용하여 투영 공간 구조를 메트릭 구조로 복원하여야 한다. 이를 위해서는 수식 (8)과 같이 프로젝트 구조를 메트릭 구조로 변환하기 위한 변환행렬  $T$ 를 구하면 된다.

$$P_{euc} \simeq P_{proj} T, M_{euc} = T^{-1} M_{proj} \quad (8)$$

수식 (8)에서  $P_{proj}$ 과  $M_{proj}$ 은 투영 공간상에서의 카메라 행렬과 3차원 점이며,  $P_{euc}$ 과  $M_{euc}$ 는 메트릭 구조에서의 카메라 행렬과 3차원점이다. 여기서  $T$ 는 수식 (9)와 같이 표현할 수 있다.

$$T = \begin{pmatrix} A & 0 \\ r^T & 1 \end{pmatrix}, A = \begin{pmatrix} f_x & s & u \\ 0 & f_y & v \\ 0 & 0 & 1 \end{pmatrix}, r^T = (a \ b \ c) \quad (9)$$

수식 (9)에서  $(a, b, c)^T$ 는 투영 공간을 아핀(affine) 공간으로 변환하기 위해 필요한 무한대의 평면이며,  $A$ 는 카메라의 내부 매개 변수 행렬을 의미한다. 그러므로 수식 (8)은 수식 (10)으로 표현이 가능하다.

$$P_{euc} = P_{proj} T = P_{proj} \begin{pmatrix} A \\ r^T \end{pmatrix} = AR \quad (10)$$

수식 (10)에서  $R$ 은 회전 행렬이며,  $RR^T = I$ 므로 다음의 수식 (11)을 유도할 수 있다.

$$\begin{aligned} P_{proj} \begin{pmatrix} A \\ r^T \end{pmatrix} P_{proj}^T &= P_{proj} \begin{pmatrix} AA^T Ar \\ r^T A^T rr^T \end{pmatrix} P_{proj}^T = AA^T \\ P_{proj} QP_{proj}^T &= C \end{aligned} \quad (11)$$

수식 (11)에서  $Q$ 를 코드릭(quadric),  $C$ 를 코닉(conic)이라고 정의한다. 그러므로 투영 공간 구조를 메트릭 구조로 변환하기 위한 변환 행렬  $T$ 를 구하는 문제는 코드릭  $Q$ 를 구하는 문제로 볼 수 있다.

그러나 자기 보정 알고리즘을 사용할 때는, 카메라에 대한 사전 정보가 없다면 수행하기 힘들다. 그러므로 투영 공간상에서 복원된 카메라 행렬로부터 적접적으로 코드릭  $Q$ 를 구하기 위해서는 카메라에 대한 사전 지식이 필요하다[17]. 본 논문에서는 모든 카메라의 내부 매개 변수가 동일하다는 즉, 상수라는 가정 하에서 카메라의 위치 정보를 보정하였다. 이를 위해 수식 (12)를 사용하였다.

$$\sum_{i=0}^n \| \|AA^T\|_F - \|P_{proj} QP_{proj}^T\|_F \|_F \quad (12)$$

수식 (12)에서  $\|\cdot\|_F$ 는 Forenium 거리이다.

최적화 수행 후 구해진 코드릭  $Q$ 로부터 변환 행렬  $T$ 를 구하기 위해서  $Q$ 를 수식 (13)으로 분해함으로써 구해 낼 수 있다.

$$\begin{aligned} Q &= \begin{bmatrix} f_x^2 + s^2 + u^2 & sf_y + uv & u & af_x + sb + uc \\ sf_y + uv & f_y^2 + v^2 & v & bf_y + vc \\ v & v & 1 & c \\ af_x + sb + uc & bf_y + vc & c & a^2 + b^2 + c^2 \end{bmatrix} \\ &= \begin{bmatrix} q_0 & q_2 & q_3 & q_5 \\ q_2 & q_1 & q_4 & q_6 \\ q_3 & q_4 & 1 & q_7 \\ q_5 & q_6 & q_7 & q_8 \end{bmatrix} = \begin{bmatrix} AA^T Ar \\ r^T A^T rr^T \end{bmatrix} \quad (13) \\ u &= q_3, v = q_4, f_y = \sqrt{q_1 - v^2}, \\ s &= (q_2 - uv)/f_y, f_x = \sqrt{q_0 - u^2 - s^2} \\ c &= q_7, b = (q_6 - vc)/f_y, \\ a &= (q_5 - sb - uc)/f_x \end{aligned}$$

### 3.2 묶음 정렬(bundle adjustment)

앞 단계까지는 영상과 영상의 관계만 정의하였고, 이러한 과정은 단지 한 쌍의 영상에 대해서만 계산이 되었을 뿐이다. 따라서 계산이 진행되어 영상이 계속 추가 될수록 영상 정렬에 대한 오차가 계속적으로 커지는 지터(jitter)가 발생할 수 있다. 이와 같이 카메라 보정 시 발생되는 오차 전파(error propagation)를 최소화하기 위해서는 카메라 보정 마지막 단계에서 모든 카메라와 모든 영상들의 대응 관계가 정의된 묶음 에지를 이용한 묶음 정렬[20]을 수행한다. 카메라 보정 단계에서 발생되는 문제는 실제로는 정확한 대응점이지만 카메라 보정 과정을 수행할 때 잡음으로 인식되어 대응 관계가 끊어지는 현상이 자주 발생하게 된다. 이러한 경우 묶음 정렬을 수행시키게 되면, 동일한 3차원점이 두 개 또는 여러 개의 3차원 점으로 인식되어 묶음 정렬의 효율성을 저하시키는 원인이 된다. 그러므로 앞서 정의된 영상간의 대응

관계를 재정의 할 필요가 있다. 이를 위한 영상 재정렬 과정은 그림 7과 같이 나타낼 수 있다.

그림 7에서 점선은 뮤음 에지를 나타내며, 실선은 SFM 에지를 나타낸다. 그림 7(a)는 SFM 에지로 연결되어 있었으나, 카메라 위치 보정 과정에서 실제로는 에러가 아니지만, 에러로 판단되어 대응 관계가 끊어진 경우이며, 그림 7(b)는 뮤음 에지로는 대응 관계가 정의되어 있으나, SFM 에지로는 대응 관계가 정의되어 있지 않아 대응 관계가 없는 것으로 판단되어 지는 경우이다. 그림 7과 같이 카메라 위치 보정 과정에서 대응 관계가 끊어진 경우와 SFM 에지로 대응 관계가 정의되지 않은 대응점들의 경우에도 3차원 점을 동일하게 유지할 필요가 있다. 이를 위해 본 논문에서는 다음의 2가지의 조건을 만족할 경우에 대응점들이 같은 3차원 점을 가진다고 판단하여, 끊어진 에지를 연결하거나 동일한 3차원 점을 가지도록 설정한다.

**조건1.** 끊어지거나 뮤음 에지로 연결된 점의 3차원 점을 각각의 영상에 투영하여 그 오차가  $0 \leq \epsilon_{pr} < 1.96\sigma + \bar{m}$ 을 만족해야 한다.

**조건2.** 연결이 되어있는 모든 2차원 점과 카메라

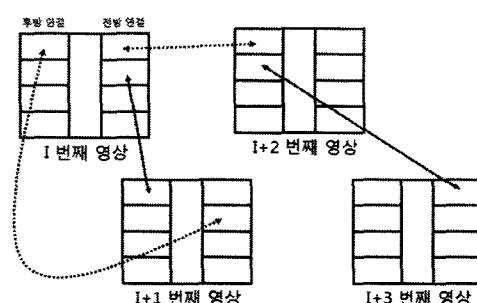
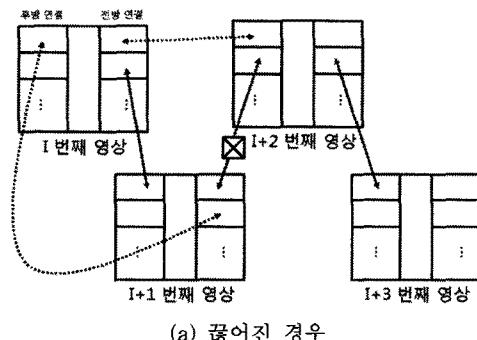


그림 7. 뮤음 정렬을 위한 3차원 점의 재 설정

행렬을 이용하여 구해진 3차원 점을 계산하여 이를 각 영상에 재 투영시킨 점과 실제 영상 상에 존재하는 점과의 오차가  $0 \leq \epsilon_{pr} < 1.96\sigma + \bar{m}$ 을 만족해야 한다.

여기서  $\epsilon_{pr}$ 은 생성된 3차원 점을 영상에 재 투영한 점으로부터 실제 영상 상에 존재하는 2차원 점까지의 거리이고,  $\bar{m}$ 은 각 영상에서 생성된 3차원 점을 2차원 상에 투영하여 얻어진 점과 실제 영상 상에 존재하는 점과의 오차 평균값이며, 이는 3개의 카메라에 나타나는 3차원 점을 기준으로 계산한다.  $\sigma$ 는 분산 값으로써 오차 범위가 신뢰구간 95% 범위 내에 있다면 두 개의 점이 같은 3차원 점을 보고 있다고 판단한다. 만약 위 두 조건을 만족한다면 모든 카메라의 대응점과 카메라를 이용하여 3차원 점을 다시 계산하여 모든 점들이 동일한 3차원 점을 가지도록 설정한다.

대응 관계를 다시 정의한 후,  $\hat{P}_k$ 와  $\hat{M}_i$ ,  $m_{ki}$ 를 각각 카메라 행렬과 3차원 점, 영상상의 2차원 점으로 둔다면 뮤음 정렬은 수식 (14)와 같이 표현할 수 있다.

$$\min_{\hat{P}_k, \hat{M}_i} \quad (14)$$

수식 (14)에서  $D(m_{ki}, \hat{P}_k \hat{M}_i)$ 는 유클리디안(euclidean) 거리이다[20].

최종적으로 추정된 카메라의 위치 정보는 다음 장에서 소개할 모델 생성 과정에서 영상에 존재하는 그룹화된 선들과 소실점 그리고, 사용자의 간단한 입력과 함께 3차원 공간에서 복원될 건물의 위치를 결정하는데 사용된다.

## 4. 모델 생성

본 논문에서 초기에 구해진 건물의 형태는 평면을 기준으로 생성한 간단한 뼈대 구조(wire-frame structure)이다. 그런 다음 하프에지(half-edge) 자료 구조와 오일러 연산자(Euler operator)를 적용한 모델링 양식을 이용하여 상세 모델을 생성하게 된다.

### 4.1 건물의 입방체(Solid) 구조 복원

컴퓨터 비전 기술만으로는 영상을 통해 건물을 복원하기 위한 모든 정보를 추출하기 어렵다. 따라서, Li는 단일 영상으로부터 건물을 복원하기 위하여 사용자의 전반적인 저작 과정을 통하여 건물의 외곽

형태를 이루는 구조를 복원하였고[21], Devebec은 앞선 연구에 비해 입력의 횟수가 줄어들기는 했지만 [3], 여전히 사용자의 입력을 동반하는 직선과 직선 간의 관계로부터 건물을 복원하는 방법을 사용하였다. 하지만, 일련의 앞선 연구들에서는 최종적으로 완성된 건물을 복원하기 위하여 여전히 사용자의 입력이 많이 필요하고, 특히 사용자의 기하학적 통찰력에 의존하는 작업이 상당수 차지하게 된다. 본 논문에서는 전반부의 자동화 단계에서 추출한 건물 모델링에 필요한 정보를 이용하여 대상 건물의 기하학적 연관 관계나 배치를 구성하도록 함으로써 사용자의 기하학적 통찰에 의한 의존도를 줄이고, 건물의 모델링하는 단계에서의 사용자 입력을 간소화함으로써 사용자의 노고를 줄일 수 있도록 하였다. 여기서 사용되는 주요한 정보는 건물 영상으로부터 추출되는 선과 소실점인데, 특히 소실점과 소실점을 추출할 때 그룹화된 직선 성분들을 사용한다. 사용자는 추출된 소실점과 소실점으로 그룹화된 직선 성분들을 이용하여 간단한 조작을 함으로써 기본적인 입방 모델을 생성할 수 있다. 그렇게 하기 위하여 앞서 자동으로 그룹화된 직선 성분들은 해당되는 소실점의 방향으로 확장시키고, 이렇게 소실점 방향으로 확장된 다수의 선들은 특정 위치에서 교차하게 되며, 이 교차점을 연결하게 되면 2차원 상에서도 평면을 결정할 수 있게 된다. 만약, 이 과정을 거쳤는데도 불구하고 평면을 결정하기에 충분한 직선 성분이 추출되지 않는다면 본 논문에서는 사용자의 입력에 의해 임의의 점을 소실점 방향으로 확장시켜 새로운 직선 성분을 추가하여 평면을 결정하도록 구현하였다. 이러한 과정을 거쳐 모든 영상에 대해 평면을 결정한 후에는 2차원 평면상의 점 사이에 대응 관계를 이용하여 3차원 상의 평면을 결정하게 된다. 물론 카메라의 위치 정보와 2차원 특징점들의 대응 관계로부터 얻어진 3차원 점은 평면상의 점들에 대한 관계를 결정함에 있어 유용한 정보로 사용되어 쉽게 대응 관계를 결정할 수 있다.

## 4.2 건물의 상세 모델링

### 4.2.1 하프 에지 자료 구조(half-edge data structure)와 오일러 연산자(Euler operator)

본 논문에서는 건물의 기하 정보를 조작하기 위

하여 하프 에지(half-edge) 자료 구조를 사용하였다. 하프 에지는 경계 표현(boundary representation)방법으로 입방 모델을 표현하는 전형적인 자료 구조로써 입체를 구성하는 에지는 두 개의 서로 반대되는 방향을 가지는 하프 에지로 구성되며, 각 하프 에지는 하나의 면(face)만을 가진다. 각 하프 에지는 그림 8과 같이 다음 하프 에지(next half-edge)와 이전 하프 에지(previous half-edge), 반대 방향 하프 에지(opposite half-edge), 생성 면(incident face)을 가진다.

하프 에지 자료 구조를 사용하여 실제로 3차원의 다면체를 생성할 때는 생성된 다면체의 기하학적 유효성을 보장하기 위해서 수식 (15)의 Euler-Poincaré 공식을 만족하여야 한다.

$$F + V - E - L = 2(S - H). \quad (15)$$

수식 (15)에서  $F$ 는 면(face)의 수,  $V$ 는 꼭지점(vertex)의 수,  $E$ 는 모서리(edge)의 수,  $L$ 은 루프(loop)의 수 그리고  $H$ 는 구멍(hole)의 수,  $S$ 는 셀(shell)의 수이다.

만약 다면체를 변경하고자 하는 경우 하프 에지는 Euler-Poincaré 공식을 만족하는 오일러 연산자(euler operator)를 사용한다[21]. 오일러 연산자는 크게 생성(make) 연산자와 삭제(kill)연산자로 나누어진다. 그림 9는 기본적인 생성 연산자의 예를 보여주고 있다. 삭제 연산자는 생성 연산자의 반대 순서로 동작한다.

그림 9에서 mev(make edge vertex), msfv(make shell face vertex), mfe(make face edge)와 같은 전체적인 오일러 연산자는 [22]에 정의되어져 있다.

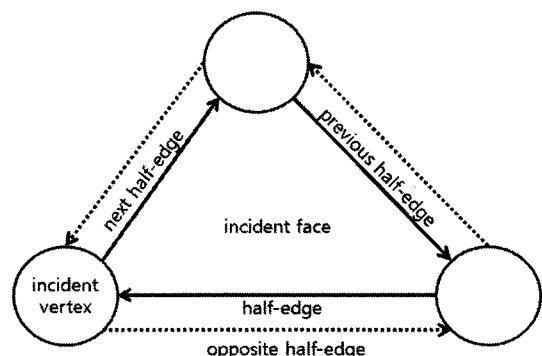


그림 8. 하프에지(half-edge) 자료구조[22]

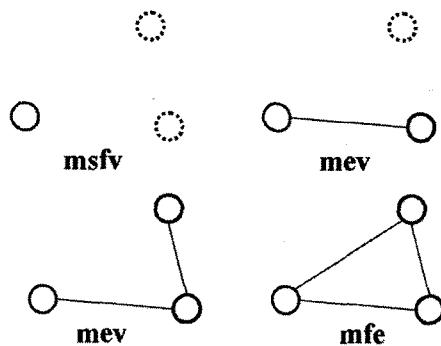


그림 9. 오일러 연산자의 예

#### 4.2.2 상세 모델 조작(Detailed Model Manipulation)

본 논문에서는 보다 효율적인 모델링을 위해 오일러 연산자들을 이용하여 다양한 모델링이 가능하도록 사용자 모델링 방법을 정의하고 이를 사용하였다. 그림 10은 오일러 연산자의 조합을 통해 창문이나 현관의 구조를 복원하기 위한 새로운 연산자 집합을 정의한 예이다.

그림 10(b)는 3개의 mev 연산자, 한 개의 mfe 연산자, 한 개의 mhe(make hole edge) 연산자의 집합을 조합되어 하나의 창문을 만들어 내는 연산자이며, 그림 10(c)의 경우는 평면의 움직임과 평면이 가지는 에지의 수 만큼인 3개의 mev 연산자, 한 개의 mfe 연산자 조합이 사용되어졌다. 물론 기둥을 가진 현관을 효율적으로 모델링하기 위해 구멍을 생성하는 연

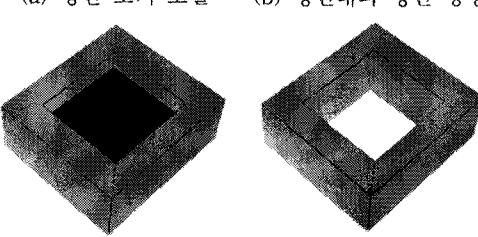
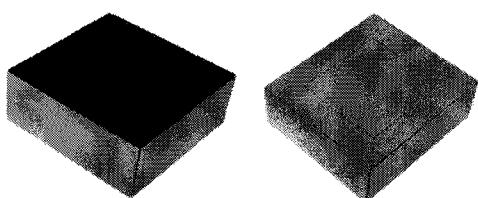


그림 10. 현관이나 창문 구조를 표현하기 위한 오일러 연산자 집합

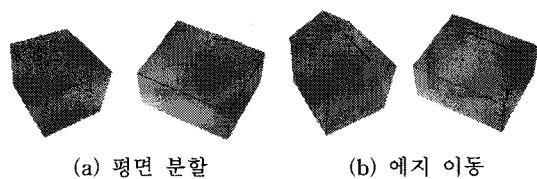


그림 11. 지붕 모델링을 위한 오일러 연산자와 그 집합

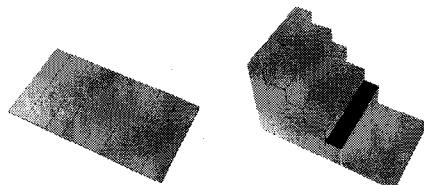


그림 12. 계단을 표현하기 위한 오일러 연산자 집합의 예

산자를 정의하여 사용하였으며 이러한 경우 그림 10(c)의 연산자 수행 후, 하나의 kfl(kill face loop)와 하나의 mhe 연산자를 적용하여 정의하였다.

그림 11은 지붕 모델링을 위한 오일러 연산자와 그 집합으로 구성된 기능의 예를 보여주고 있다.

그림 11(a)에서는 평면 분할을 수행할 수 있는 기능의 예를 보여주고 있으며 그림 11(b)는 평면 분할 후 에지를 이동시킴으로서 지붕의 모양을 모델링하고 있는 예이다. 그림 12는 건물에서 자주 나타나는 모형인 계단을 모델링하기 위한 방법으로 평면 결정 후 분할 할 평면의 개수와 높이 정보를 입력으로 하여 계단을 생성하는 예를 보여주고 있다.

## 5. 구현 및 결과

본 논문에서 제안하는 모델링 방법은 윈도우 XP, Visual Studio .Net 2005에서 구현되었다. 대화상자 기반의 윈도우로 구현되었으며, 기본적인 모든 인터페이스는 GUI 기반이며 초기 설정과 오일러 연산자를 이용한 상세 모델링을 제외하고는 사용자의 입력 없이 자동으로 수행이 된다. 앞서 살펴본 그림 10은 GUI 기반에서 수행되어지는 예이다.

그림 13은 제안하는 시스템의 메인 화면을 나타낸다. 그림 13(a)는 카메라 위치 보정을 위한 특정점 추출 및 카메라 위치 보정의 화면을 보여주며, 그림 13(b)는 건물의 초기 모델을 결정하기 위한 그룹화된 선과 소실점으로의 확장된 화면을 나타내며, 그림 13(c)는 모델 생성을 위한 메인 화면을 나타낸다.



그림 13. 시스템 메인 화면

그림 14는 2절에서 제안한 역호프 변환에 의해 추출된 직선 성분의 결과를 보여주고 있다. 그림 14에서 보는 바와 같은 역호프 변환은 나무와 같은 에지 성분들이 누적되어 발생되는 잘못된 직선 성분을 효과적으로 제거하여 보다 정확한 직선 성분을 추출할 수 있음을 알 수 있다.

그림 15(a)는 추출된 직선 성분을 가우시안 반구에 투영하여 얻어진 결과이며, 그림 15(b)는 같은 소실점을 공유하는 직선 성분의 그룹화 결과이다.

그림 16은 무작위로 촬영된 영상으로부터 대응점의 개수를 가중치로 하여 신장 트리(spanning tree) 알고리즘으로 획득한 인접 그래프이다. 인접 그래프는 무작위로 획득한 영상 집합에서 각 영상간의 연관

성을 대응점의 개수로 정의하여 연관성이 높은 영상들로 영상간의 관계를 정렬하였다. 획득한 인접 그래프는 카메라 보정을 위한 최종적인 경로로 사용된다.

그림 17은 건물의 초기 입방 모델을 얻기 위해 사용자의 선택에 의해 확장된 직선 성분과 그들 간의 교차점을 보여주고 있다. 이 단계에서 사용자는 그룹화된 직선 중에서 소실점 방향으로 확장할 직선 성분을 선택하고, 직선의 대응 관계는 카메라의 위치 정보와 에피폴라 기하학으로 구해낸다.

그림 18은 초기 입방 모델과 4.2.1 절의 기본적인 오일러 연산자와 4.2.2 절의 상세 모델링 연산자 집합을 통해 생성된 상세 모델을 보여주고 있다. 모델링을 위한 사용자의 조작은 GUI 기반의 자유도를 제공한다.



그림 14. 역호프 변환에 의해 구해진 직선 성분

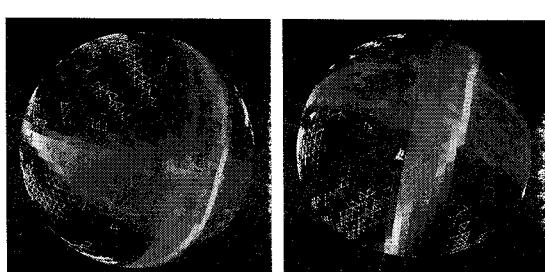


그림 15. 가우시안 반구를 이용한 소실점 추출과 그룹화 직선 성분

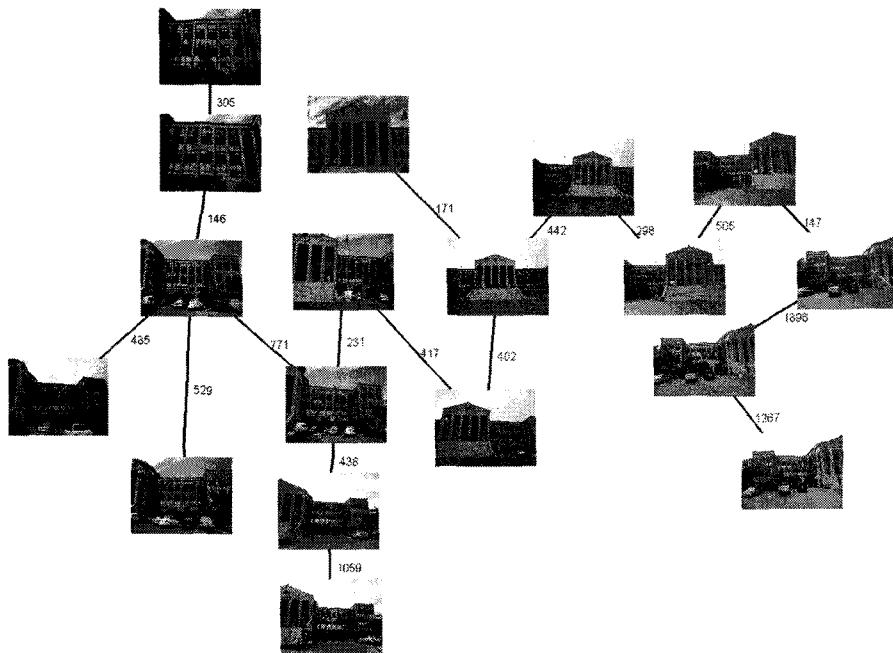


그림 16. 카메라 위치 보정을 위한 인접 그래프

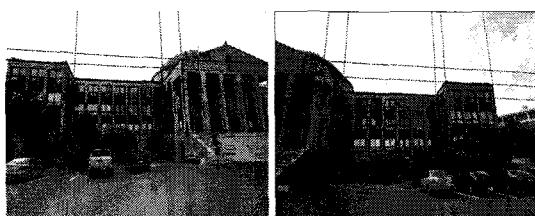


그림 17. 소실점 방향으로 직선 성분 확장

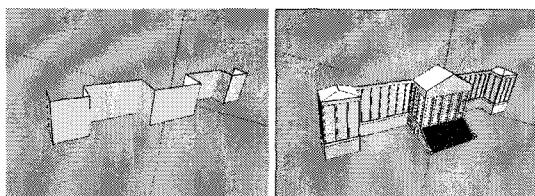


그림 18. 오일러 연산자 집합을 이용한 건물의 상세 모델링

그림 19는 초기 입방 모델에서 지붕 부분에 해당되는 구조를 복원하는 예를 보여주고 있다. 먼저 GUI 기반에서 사용자가 입력에 의해 지붕에 해당되는 평면을 생성하고 오일러 연산자를 이용하여 평면 분할 다음 선을 평면의 법선 방향으로 이동시켜 새로운 평면을 재정의함으로써 지붕의 모양을 결정한다.

그림 20은 지붕 복원 후 초기 평면 모델 좌표계의 Y축으로 확장시켜 초기 모델에서 복원되지 않는 건물의 밑 부분을 생성한 예를 보여주고 있다. 그런 다음 계단 생성을 위해 평면의 선을 선택한 후 이를 평면의 법선 방향으로 확장시켜 계단을 생성하는 예를 보여주고 있다.

그림 21은 건물의 입구 부분에 존재하는 기둥을 생성하기 위한 예를 보여주고 있다. 그림 21(a)는 계단을 복원한 확장 모델에서 구멍을 생성할 면을 선택하였고, 그림 21(b)는 선택한 면을 법선 벡터 방향으로 이동하는 예를 나타내며, 그림 21(c)는 생성된 구멍을 보여준다. 그림 21(d)는 건물의 입구 부분에 존재하는 기둥을 생성한 결과이다.

그림 22는 동일한 유형의 면을 조작하는 인터페이스를 나타내고 있다. 그림 22(a)는 면을 선택한 후, 복사 및 붙이기를 이용하여 동일한 유형의 면을 생성한 예를 보여준다. 그림 22(b)는 복사된 모든 면을 동시에 밀기 연산자를 이용하여 창문을 생성한 예를 보여준다.

그림 23은 건물의 상세 모델링으로부터 얻어진 결과를 다시 영상에 재 투영하여 텍스쳐를 계산하는 예를 보여주고 있다. 여기에서 사용자의 입력은 단지

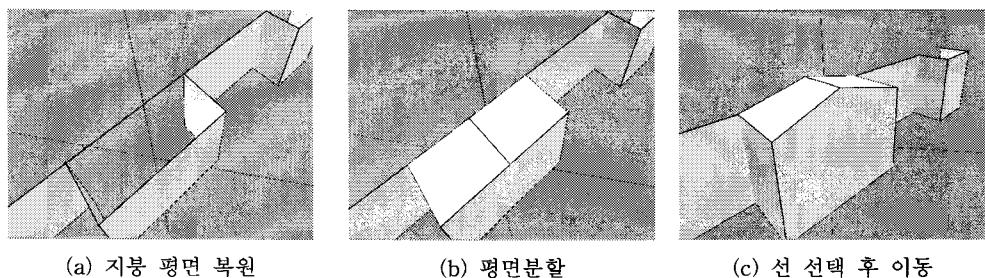


그림 19. 지붕 복원을 위한 오일러 연산자와 인터페이스

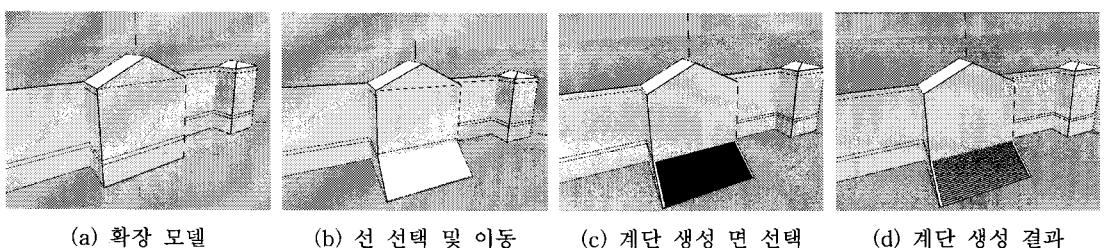


그림 20. 계단 생성을 위한 사용자 인터페이스와 결과

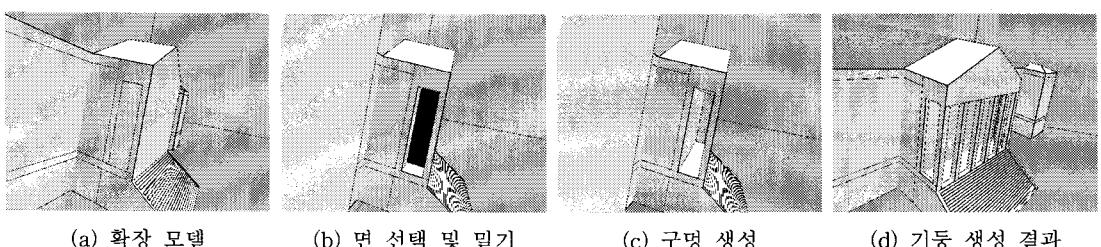


그림 21. 건물의 기둥 생성을 위한 사용자 인터페이스와 결과

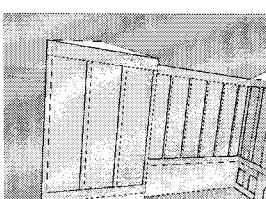
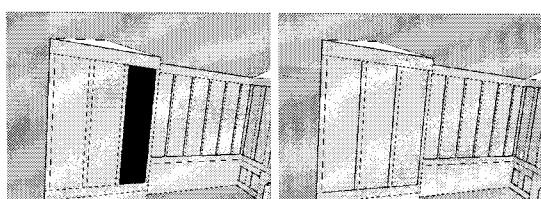


그림 22. 동일한 유형의 면의 조작 인터페이스와 결과

어느 영상에서 가장 좋은 텍스쳐를 얻을 수 있는지를 선택할 수 있도록 하였다.

텍스쳐는 그림 23과 같이 모델을 영상에 재 투영하여 선택할 수 있을 뿐 아니라, 그림 24와 같이 텍스쳐 형판(template)을 이용하여 생성할 수도 있다. 그림 24(a)는 텍스쳐 형판을 보여주며, 그림 24(b)는 형판을 이용하여 생성한 텍스쳐이다. 텍스쳐 형판과 형

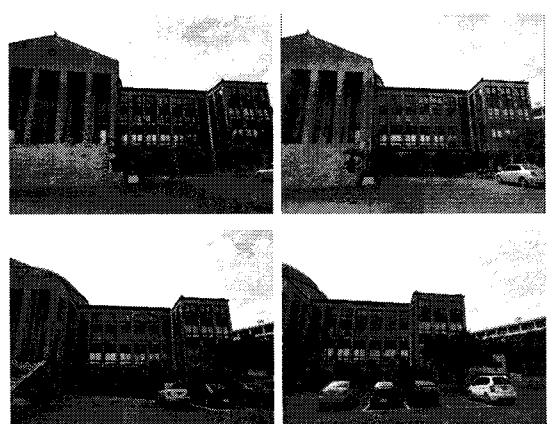
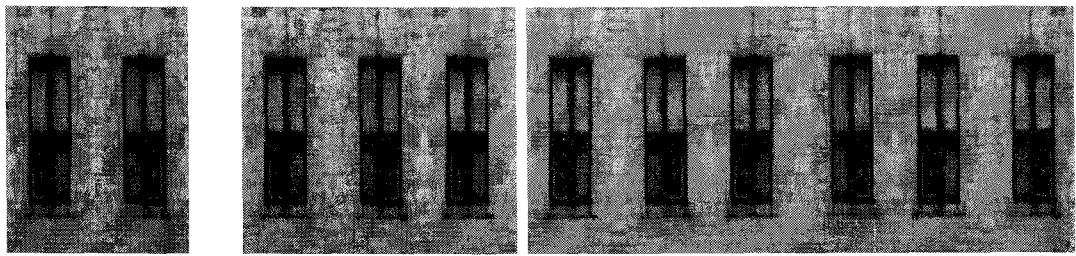


그림 23. 텍스쳐를 선택할 수 있는 영상

판을 이용하여 새롭게 생성한 텍스쳐를 이용하여 건물의 밑 부분에 텍스쳐를 적용하였고, 동일한 유형을



(a) 텍스처 형판

(b) 형판을 이용하여 생성한 텍스처

그림 24. 형판을 이용한 텍스처 생성

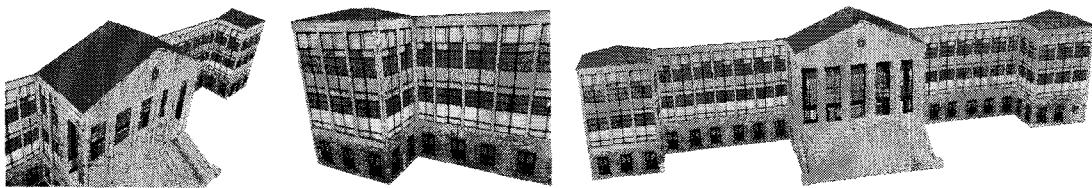


그림 25. 최종결과

가지는 면에 대해서는 동일한 텍스쳐를 적용하였다.

그림 25는 텍스쳐 결정 후 최종적인 모델링 결과를 보여주고 있다.

## 6. 결 론

본 논문은 보이지 않는 부분을 복원할 수 없는 자동화 건물 복원의 단점과 오랜 시간이 요구되는 수동화 건물 모델링의 단점을 보완하기 위한 반자동 건물 모델링 시스템을 제안하였다. 먼저, 비교정 영상은 카메라 보정을 위해 영상 정렬 기법을 이용하여 순서를 정의하고 이를 인접 그래프를 도입하여 표현하였다. 인접 그래프는 영상간의 관계, 대응점 관계 및 카메라 추정 순서를 효과적으로 표현한다. 인접 그래프 생성 후 카메라 보정 과정이 수행되며, 보정과정 중 추정된 3차원 점 중에서 동일한 3차원 점임에도 불구하고 다른 점으로 인식되는 점을 사용자가 직접 확인하고 수정함으로써, 보다 정확한 카메라 보정 과정이 수행 가능하도록 하였다. 카메라 보정 후 소실 점 추출 시 얻어진 직선 성분들을 간단한 사용자의 입력을 이용하여 각 소실점 방향으로 확장시켜 얻어진 교차점의 대응관계를 이용하여 초기 입방 모델을 생성하였다. 건물의 상세모델링을 위해 초기 입방 모델은 하프 에지 자료 구조를 이용하여 표현하고, 오

일려 연산자와 그들의 조합을 통하여 정의되어진 새로운 오일러 연산자를 이용하여 창문, 계단 및 지붕과 같은 기하 구조를 효과적으로 모델링 할 수 있었다. 그리므로 수동화와 자동화의 단점을 보완하여 효과적으로 건물을 모델링 할 수 있었다.

그러나 제안한 시스템은 영상에서 나타나는 소실점과 직선 성분을 이용하여 건물의 초기 입방 모델을 생성하는 모델링 기법을 적용하였기 때문에, 주로 고덕 양식의 건물과 같이 직선 위주의 건물에 대하여 최적화 되어 있으므로, 복잡한 곡면으로 이루어진 건물은 복원하기 어렵다. 본 연구에서는 차후에 타원이나 입방체, 원뿔과 같은 다양한 기본 객체를 지원하는 연산자를 추가하여 다양한 형태의 복합 건물 모델링이 가능하도록 할 것이다.

## 참 고 문 헌

- [1] I. Stomas and P.K. Allen, "Automatic Registration of 2D with 3D Imagery in Urban Environment," *International Conference on Computer Vision*, pp. 731-736, 2001.
- [2] S.F. El-Hakim et al., "Detailed 3D Reconstruction of Large-scale Heritage Sites Integrated Techniques," *IEEE Computer*

- Graphics and Applications*, Vol.23, No.3, pp. 21-29, 2004.
- [3] P.E. Debevec, C.J. Taylor and J. Malik, "Modeling and Rendering Architecture from Photographs: A hybrid geometry- and image-based approach," *International Conference on Computer Graphics and Interactive Techniques*, pp. 11-20, 1996.
- [4] R. Cipolla, T. Drummond and D. Robertson, "Camera Calibration from Vanishing Points in Images of Architectural Scenes," *The 10th British Machine Vision Conference*, pp. 382-391, 1999.
- [5] "Photo Modeler," Eos System, <http://www.photomodeler.com/>
- [6] T. Sato, M. Kanbara and N. Yokoya, "Outdoor Scene Reconstruction from Multiple Image Sequences Captured by a Hand-held Video Camera," *IEEE International Conference on Multisensor Fusion and Integration for Intelligent System*, pp. 13-118, 2003.
- [7] P. V. C. Hough, "Machine Analysis of Bubble Chamber Pictures," *International Conference on High Energy Accelerators and Instrumentation*, pp. 554-556, 1959.
- [8] J. Illingworth and J. Kittler, "The Adaptive Hough Transform," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.9, pp. 690-698, 1987.
- [9] P. Palmer, J. Kittler and M. Petrou, "Using Focus of Attention with the Hough Transform for Accurate Line Parameter-Estimation," *Pattern Recognition*, Vol.27, No.1, pp. 127-134, 1994.
- [10] A. Kesidis, and N. Paramarkos, "On the Inverse Hough Transform," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.21, No.12, pp. 1329-1343, 1999.
- [11] K.H. Jang and S.K. Jung, "3D City Model Generation from Ground Images," *The 24th Computer Graphics International*, pp. 630-638, 2006.
- [12] S.C. Lee, S.K. Jung and R. Nevatia, "Automatic Integration of Facade Textures into 3D Building Models with a Projective Geometry Based Line Clustering," *Computer Graphics Forum*, Vol.21, pp. 511-519, 2002.
- [13] 차영미, 적응적인 공선상의 애지점 그룹을 이용한 소설점 검출, 석사학위논문, 경북대학교, 2005.
- [14] J. Kosecka and W. Zhang, "Efficient Computation of Vanishing Points," *IEEE International Conference on Robotics and Automation*, Vol.1, pp. 223-229, 2002.
- [15] D. G. Lowe, "Distinctive Image Features from Scale-invariant Keypoints," *International Journal of Computer Vision*, Vol.60, No.2, pp. 91-110, 2004.
- [16] M. Fischler and R. Bolles, "RANdom SAMpling Consensus: a paradigm for model fitting with application to image analysis and automated cartography," *Communications of the ACM*, Vol.24, pp. 381-395, 1981.
- [17] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge Press, New York, 2003.
- [18] E. Horowitz, S. Sahni and D. Mehta, *Fundamentals of Data Structures in C++*, Computer Science Press, division of W. H. Freeman, New York, 1995.
- [19] M. Pollefeys, L.V. Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, R. Koch, "Visual Modeling with a Hand-held Camera," *International Journal of Computer Vision*, Vol.53, No.3, pp. 207-232, 2004.
- [20] B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon, "Bundle Adjustment - A Modern Synthesis," *Lecture Notes in Computer Science*, Vol.1883, pp. 298-372, 2000.
- [21] Z. Li, J. Liu and X. Tang, "Shape from Regularities for Interactive 3D Reconstruction of Piecewise Planar Objects from Single Images," *Proceedings of the 14th annual*

- ACM International Conference on Multimedia.  
pp. 85~88, 2006.
- [22] M. Mantyla, *An Introduction to Solid Modeling*, Computer Science Press, New York, 1998.

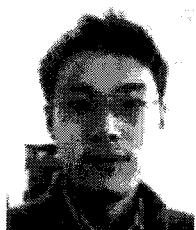


### 장 경 호

1991년 3월~1998년 2월 경성대학교 전산통계학과 이학사  
1998년 3월~2000년 2월 경북대학교 컴퓨터공학과 공학석사  
2000년 3월~2005년 2월 경북대학교 컴퓨터공학과 공학박사

2005년~현재 경북대학교 전자전기컴퓨터학부 연구교수  
관심분야 : Image-Based Modeling and Rendering, Virtual Reality, Computer Graphics, Compute Vision

### 장재석



2000년 2월~2007년 2월 경북대학교 컴퓨터공학과 공학사  
2007년 3월~2009년 2월 경북대학교 전자전기컴퓨터학부 공학석사  
2009년~현재 경북대학교 전자전기컴퓨터학부 박사과정

관심분야 : 3D Reconstruction, Computer Vision, Computer Graphics, Virtual Reality



### 이 석 준

1999년 2월~2004년 2월 대구대학교 정보통신공학부 공학사  
2004년 3월~2006년 2월 경북대학교 컴퓨터공학과 공학석사  
2006년~현재 경북대학교 컴퓨터공학과 박사과정

관심분야 : Scientific Information Visualization / Authoring System, 3D based Visual System, Wearable Computing, HCI.



### 정순기

1986년 3월~1990년 2월 경북대학교 컴퓨터공학과 공학사  
1990년 3월~1992년 2월 한국과학기술원 전산학과 이학석사

1992년 3월~1997년 2월 한국과학기술원 전산학과 공학박사

1997년~1998년 University of Maryland, Research Associate.

2001년~2002년 IRIS, University of Southern California, Research Associate

1998년~현재 경북대학교 컴퓨터공학과 교수

1999년~현재 (주)아이디스 기술고문

2008년~2009년 IRIS, University of Southern California, Visiting Faculty

2008년~현재 (주)아이디스 사외이사

관심분야 : Virtual Reality, Artificial Intelligence, Computer Vision, Image Processing, Computer Graphics.