

특집논문-06-11-4-11

# 실감 콘텐츠 생성을 위한 분해법 기반 다수 카메라 시스템 자동 보정 알고리즘

김기영<sup>a)</sup>, 우운택<sup>a)†</sup>

## Self-calibration of a Multi-camera System using Factorization Techniques for Realistic Contents Generation

Kiyoung Kim<sup>a)</sup> and Woontack Woo<sup>a)†</sup>

### 요 약

본 논문에서는 실감 콘텐츠 생성에 쓰이는 움직이는 다수 카메라 시스템을 위한 분해법 기반 자동 보정 알고리즘을 제안한다. 대부분의 기존 자동 보정 연구는 스테레오 카메라 혹은 고정된 다수 카메라 시스템에 한정되어왔다. 보다 넓은 영역의 3차원 복원이나, 다수 카메라를 사용하는 일반적인 응용으로의 확장을 위해서는 새로운 알고리즘에 대한 연구가 필요하다. 본 논문에서는 일반적인 다수 카메라 시스템 (동적인 시스템 포함) 구조에 대한 강건한 자동 보정 방법을 제안하고, 카메라 구조가 평면 구성을 갖는 다수 카메라 시스템 보정 알고리즘을 제안한다. 우선, 기하학적 제한 요소를 바탕으로 분해법을 사용하기 위한 수식을 유도하고, 유도된 정리를 바탕으로 실질적인 자동 보정 사용법을 설명한다. 특히, 평면을 이루는 카메라 시스템에서의 분해법 적용이 어파인 복원과 같다는 성질을 증명하고, 어파인 분해법을 활용한 자동 보정 알고리즘을 제시한다. 실험에서는 제안된 알고리즘을 시뮬레이션 데이터와 실제 영상에 적용한 실험 결과를 보인다. 제안된 알고리즘은 다수 카메라 시스템을 활용한 영상 기반 3차원 실감 콘텐츠 복원이나, 하드웨어의 지원과 함께 실시간 다수 카메라 증강 현실 시스템에 활용될 수 있다.

### Abstract

In this paper, we propose a self-calibration of a multi-camera system using factorization techniques for realistic contents generation. The traditional self-calibration algorithms for multi-camera systems have been focused on stereo-(rig) camera systems or multiple camera systems with a fixed configuration. Thus, it is required to exploit them in 3D reconstruction with a mobile multi-camera system and another general applications. For those reasons, we suggest the robust algorithm for general structured multi-camera systems including the algorithm for a plane-structured multi-camera system. In our paper, we explain the theoretical background and practical usages based on a projective factorization and the proposed affine factorization. We show experimental results with simulated data and real images as well. The proposed algorithm can be used for a 3D reconstruction and a mobile Augmented Reality.

Keywords : 카메라 자동 보정, 다수 카메라 시스템, 분해법, 실감 콘텐츠 생성

a) 광주과학기술원 정보통신공학과 U-VR 연구실  
GIST U-VR Lab.

† 교신저자 : 우운택 (wwoo@gist.ac.kr)

\* 본 연구는 21세기 프론티어 연구개발사업의 일환으로 추진되고 있는 정보통신부의 유비쿼터스 컴퓨팅 및 네트워크 원천 기반 기술 개발 사업 과 한국전자통신연구원 광통신연구센터의 광가입자망 (FTTH) 서비스개발 실험사업의 지원에 의한 것임

## I. 서 론

영상 기반 3차원 복원 (3D reconstruction)은 카메라로 촬영된 2차원 영상들을 이용하여, 영상안의 객체나 배경

을 3차원 디지털화하는 기술이다<sup>[1]</sup>. 복원된 모델은 기존 CG로 제작되는 콘텐츠 비용과 질적인 면에서 경쟁력이 높다. 그래서 영상 기반 복원은 최근 논의가 활발한 3차원 TV에 활용될 콘텐츠에 요구되는 실감성과 제작의 편의성을 제공하는 효과적인 생성 방법 중 하나로 여겨지고 있다. 실제 가상 스튜디오에서는 다수 카메라를 사용하여, Actor를 실시간으로 복원하여 다른 응용에 사용하기도 하며<sup>[2]</sup>, 단일 카메라를 활용하여 문화재를 복원하는데 이 기술을 이용하고 있다<sup>[3]</sup>. 카메라 자동 보정 (Self-calibration)은 이러한 복원 응용을 뒷받침하는 중요한 기술 중 하나이다.

카메라 자동 보정은 격자 패턴과 같은 보정 객체 없이 영상에서의 정보만을 이용하여 카메라의 내부 (Intrinsic) 및 외부 (Extrinsic) 파라미터를 구하는 과정이다<sup>[1]</sup>. 내부 파라미터는 카메라의 초점 거리등을, 외부 파라미터는 회전과 이동 행렬로 나타내어지는 카메라의 자세를 각각 의미한다. 대부분의 기존 연구는 단일 카메라를 지원하는 알고리즘에 대해 이루어졌다<sup>[3,9]</sup>. 다수 카메라용은 기존의 스테레오 시스템이나<sup>[10-17]</sup>, 카메라의 위치가 고정된 형태의 다수 카메라만을 위한 알고리즘들이 연구되었다<sup>[18][19]</sup>. 이러한 알고리즘들은 다수 카메라 시스템이 움직이는 응용에서는 적합하지 않다. 다수 카메라의 구조를 활용한 새로운 알고리즘의 연구가 필요하다.

본 논문에서는 실감 콘텐츠 생성에 쓰이는 움직이는 다수 카메라 시스템을 위한 분해법 기반 자동 보정 알고리즘을 제안한다. 알고리즘에서는 i) 일반적인 다수 카메라 시스템 구조에 대한 강건한 방법과, 특히 ii) 다수 카메라 구조가 평면을 이루는 특별한 경우를 다룬다. 모든 카메라에서 공통적으로 보이는 특징점이 있다고 가정한다. 이 특징점들로 측정 행렬을 구성하고, 분해법을 수행하여 사영 복원 (Projective reconstruction)을 수행한다. 평면 다수 카메라의 경우에는 어파인 복원 (Affine reconstruction)을 수행한다. 최종적으로 유클리드 복원 (Euclidean reconstruction)을 통해 카메라 파라미터를 얻는다.

제안하는 다수 보정 알고리즘은 다수 카메라 시스템의 구조에서 발생하는 기하학적 제한요소를 최대한 활용한다. 대표적으로 i) 다수 카메라가 움직일 경우 투영 행렬 (Projection matrix)이 불변하다는 성질, ii) 다수 카메라 움직임의 경우 각 위치에서 사영 복원된 점들 간의 호모그래피와 무한의 평면 (Plane at infinity)간의 관계를 이용한다. 평면상에 카메라가 위치하는 특수한 경우에는 어파인 분해법이 가능함을 기하학 제한 요소를 활용하여 보인다. 제안된 자동 보정은 분해법 기반으로 이루어지므로, 보정 오차가 카메라 전체에 고루 분산되는 장점과 선형 연산으로 비선형 최적화에 필요한 비교적 정확한 초기치를 제공한다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 다룰 카메라 모델 및 배경에 대해 설명한다. 3장에서는 본 논문에서 제한하는 자동 보정 알고리즘의 전체적인 흐름을 설명한다. 그리고 4장에서는 제안된 알고리즘을 보다 세부적으로 설명한다. 5장에서는 제안된 알고리즘을 활용한 실험 결과를 보이며, 6장에서는 이에 대한 결론 및 향후 연구에 대해 논한다.

## II. 배경

### 1. 단일 카메라 모델

본 논문에서는 특별한 언급 없이 다수 카메라 시스템에 장착된 카메라들이 일반적인 바늘구멍 (Pin-hole) 카메라 모델을 따르는 것으로 가정한다. 카메라 모델은 사영 (Projective), 어파인 (Affine), 그리고 유클리드 (Euclidean) 공간에서 각각 정의된다<sup>[1]</sup>. 이를 수식으로 나타내면 다음과 같다.

$$x \cong PX \cong P^A X^A \cong P^E X^E \quad (1)$$

위 식에서,  $X=[X,Y,Z,1]^T$ 는 3차원 공간의 호모지니어스

(Homogeneous)점 좌표이고,  $x=[x,y,1]^T$ 는  $X$ 가 영상 평면에 투영된 2차원 점으로 임의의 크기 인자 (up to scale)에 의존하는 값을 갖는다. 투영 관계는 투영 행렬  $P$ 의 자유도 (Degree of Freedom)에 따라 어파인 공간의  $P_A$ , 유클리드 공간의  $P_E$ 로 변환된다. 수식 아래 또는 위 첨자 'A'는 어파인 그리고 'E'는 유클리드를 뜻하고, 첨자가 없을 경우는 사영을 뜻한다.

이 중에서 유클리드 공간의 3차원 점의 투영 관계에 대한 자세한 카메라 모델의 파라미터는 식 (2)과 같다.

$$x \cong K(Rt)X^E, K = \begin{pmatrix} f & s & C_x \\ 0 & f & C_y \\ 0 & 0 & 1 \end{pmatrix} \quad (2)$$

여기서  $K$ 는 카메라 행렬 (Camera matrix)이고,  $R$ 과  $t$ 는 회전 행렬과 이동 행렬을 각각 의미한다. 카메라 행렬은 초점 거리 ( $f$ ), 초점 거리의 가로 세로 비 (Aspect ratio,  $\tau$ ), 뒤틀림 인자 (Skew parameter,  $s$ ), 그리고 주점 (Principal point,  $(C_x, C_y)$ )으로 이루어진다. 일반적으로  $s$ 는 최근 기계적 오차가 거의 없기 때문에 0이라 가정한다.

### 2. 다수 카메라 모델

그림 1과 같이 임의의  $M$ 개의 카메라로 이루어진 다수 카메라 시스템이 있다고 가정하자. 각각의 카메라는 식 (1)과 (2)에서 정의한 카메라 모델을 따른다. 이를 일반화하여, 공통된 3차원 특징점들  $X$ 를 묶으면 다수 카메라 시스템의 카메라 모델은 다음 식과 같이 표현된다.

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{pmatrix} \cong \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_M \end{pmatrix} (X) \quad (3)$$

위 식에서,  $P_i$ 는  $i$ 번째 위치에서의 카메라의 투영 행렬을

의미한다. 그리고  $x_i$ 는  $X$ 가  $P_i$ 에 의해 영상 평면에 투영된 점들의 집합이다.

다수 카메라 시스템이  $N$ 번 움직일 경우, 식 (3)은 다음 식과 같이 확장된다.

$$\begin{pmatrix} x_1^1 & x_1^2 & \dots & x_1^N \\ x_2^1 & x_2^2 & \dots & x_2^N \\ \vdots & \vdots & \ddots & \vdots \\ x_M^1 & x_M^2 & \dots & x_M^N \end{pmatrix} \cong \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_M \end{pmatrix} (X^1 X^2 \dots X^N) \quad (4)$$

$X_i$ 는  $i$ 번째 위치에서 보이는 3차원 점들을 의미한다. 다수 카메라 시스템은 시스템 자체가 움직일 경우에 내부 카메라들 간의 상대적인 기하학적 위치 관계는 불변하므로, 투영행렬은 움직임과 상관없이  $M$ 개만을 유지하게 된다<sup>[13]</sup>.

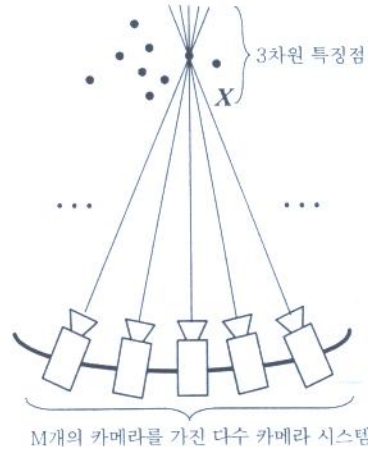


그림 1. 다수 카메라 시스템에서의 3차원 점들의 투영 관계  
Fig 1. Projection of 3D points in a multi-camera system

### 3. 자동 보정 및 문제 정의

단일 카메라의 자동 보정은 식 (1)에 정의된 투영 공간에서의 카메라 모델을 유클리드 공간의 카메라 모델로 변환시키는  $4 \times 4$  행렬  $T$ 를 구하는 과정이다<sup>[1]</sup>. 즉,  $T$ 는  $P_E = PT$

그리고  $X_e=T^1X$ 을 만족한다.  $T$ 는 다음과 같이 정의된다.

$$T = \begin{pmatrix} K & 0 \\ -p^TK1 \end{pmatrix} \tag{5}$$

위 식에서,  $p$ 는  $3 \times 1$  벡터로 무한의 평면 (Plane at infinity)을 의미한다. 다수 카메라 시스템의 자동 보정도 단일 카메라의 경우와 마찬가지로, 다수 카메라 시스템 모델 식 (3)을 유클리드 공간에서의 카메라 모델로 변환하는 행렬  $T$ 를 구하는 과정이다.

### III. 자동 보정 전체 과정

본 장에서는 다수 카메라 시스템의 자동 보정의 실질적인 세부 단계 흐름을 설명한다. 그림 2는 본 논문에서 제안하는 알고리즘의 전체 흐름도이다.

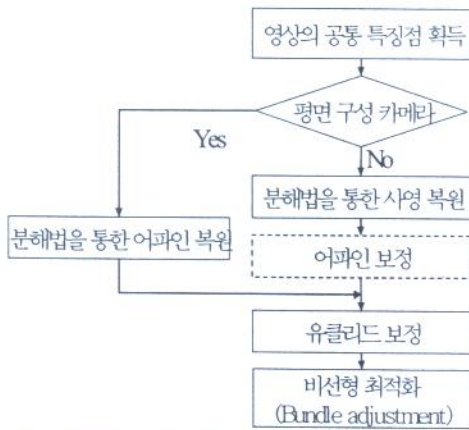


그림 2 . 다수 카메라 시스템 자동 보정 알고리즘 흐름도  
 Fig 2. The algorithm flow of the proposed self-calibration for a general multi-camera system

첫 번째 과정은 모든 위치에서 획득한 다수 카메라 시스템에서 획득한 모든 영상들에서 동일한 특징점을 찾는 것이다. 두 번째 과정은 다수 카메라의 구조에 의해, 일반 구

성용 알고리즘과 평면 구성용 알고리즘을 선택한다. 일반 구성용 알고리즘에서는 분해법을<sup>[23]</sup> 사용하여 모든 카메라의 위치에 대한 투영 행렬과 사영 복원된 점들을 구한다. 사영 복원으로부터 어파인 복원과 유클리드 복원으로의 변환은 단일 카메라의 경우와 같이 크게 두 가지 방법이 존재한다<sup>[4][8]</sup>. 유클리드 복원에서는 카메라의 내부 파라미터 및 외부 파라미터를 계산한다. 마지막 과정으로, 선형 연산만으로 전 과정에서 구해진 모든 카메라 내부 파라미터와 외부 파라미터에 비선형 최적화 알고리즘인 L-M(Levenberg- Marquardt) 알고리즘에 기반하여 Bundle adjustment를 수행하여 오차를 줄인다<sup>[21]</sup>.

### IV. 제안 알고리즘

#### 1. 일반적인 구성을 지닌 다수 카메라 시스템의 자동 보정

본 장에서는 일반적인 구성을 지닌 다수 카메라 시스템 보정 알고리즘 두 가지를 소개한다. 1.2절과 1.3절은 단계적인 (Stratified approach) 방법을 설명하고, 1.4절에서는 직접적인 방법 (Direct approach)을 설명한다.

##### 1.1 분해법 활용 사영 복원

일반적인 구성을 지닌 다수 카메라 시스템이 움직이는 경우에는 시스템에 부착된 다수 카메라들의 상대적인 기하학적 위치관계는 변하지 않는다. 그러므로 각 카메라들 간의 Fundamental 행렬이 움직임에 상관없이 유지되며, 이는 식 (4)에서 표현되듯이 투영 행렬이 다수 카메라의 움직임에 상관없이 유지됨을 의미한다.

사영 공간에서의 분해법은 영상에서 추출된 특징점들만으로 이루어진 임의의 행렬을 투영 행렬  $P$ 와 복원된 점  $X$ 로 분해하는 방법이다. 기본적으로 단일 카메라에서 수행하는 사영 공간에서의 분해법과<sup>[23]</sup> 동일하게 수행된다. 식 (4)의 좌측 행렬만 살펴보면, 다음과 같다.

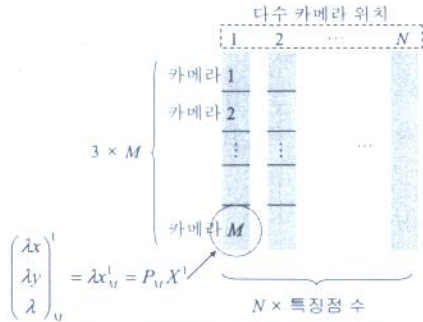


그림 3. 분해법 사용을 위한  $(3M) \times (N \times \text{특징점수})$  크기의 측정 행렬, 식 (4)의 좌측식을 의미함

Fig 3. The  $(3M) \times (N \times \# \text{ of points})$  measurement matrix for projective factorization algorithm, left matrix of the equation (4)

그림 3에서,  $\lambda$ 는 사영 깊이를 나타내는 임의의 크기 인 자이며, Fundamental 행렬을 통해 구해진다<sup>[23]</sup>. 요구되는 Fundamental 행렬은 카메라의 움직임에 불변하다는 성질을 이용하면, 보다 강건하게 계산된다. 다음 수식은 다수 카메라의 첫 번째 카메라와 두 번째 카메라간의 에피폴라 제한요소 (Epipolar constraint)를 표현한 것이다.

$$\begin{pmatrix} x_2^1 \\ x_2^2 \\ \vdots \\ x_2^N \end{pmatrix}^T F_{1,2} \begin{pmatrix} x_1^1 \\ x_1^2 \\ \vdots \\ x_1^N \end{pmatrix} = 0 \quad (6)$$

위 식에서 보는바와 같이,  $x$ 에 속한 공통점이  $K$ 개일 경우, Fundamental 행렬을 구하기 위한  $K \times M$ 개의 2차원 특징점 자원을 갖게 된다.

1.2 어파인 보정 (Affine calibration)

1.1절의 사영 복원을 통해, 각 위치에서의 사영 복원된 점  $X$ 와 다수 카메라 시스템을 이루는 카메라들의 사영 공간에서의 투영 행렬을 획득하였다. 어파인 보정은 이를 기반으로 사영 복원을 어파인 복원으로 변환하는 행렬을 구하는 과정으로 이는 식 (5)에 정의된 무한의 평면 인자 ( $p$ )를 구하는 것과 같다. [12][13]와 같이 스테레오 카메라 시

스템에서는 각각의 위치에서 복원한 점들간의  $4 \times 4$  호모그래피  $H$  전치행렬의 고유벡터 (고유치=1일 경우)가  $p$ 와 같다는 것을 보였다. 본 논문에서는  $H$ 를 다음과 같이 표기한다.

$$H_{i,j} = T D_{i,j} T^{-1} = \begin{pmatrix} K 0 \\ 0 1 \end{pmatrix} \begin{pmatrix} R_{i,j} t_{i,j} \\ 0 \quad 1 \end{pmatrix} \begin{pmatrix} K^{-1} & 0 \\ -p^T K^{-1} & 1 \end{pmatrix} \quad (7)$$

식 (7)로 부터  $p$ 는 다음 선형 연산을 구성하여 SVD (Singular Value Decomposition)를 이용하여 계산한다.

$$\begin{pmatrix} H_{1,2}^T - I_{4 \times 4} \\ \vdots \\ H_{1,N}^T - I_{4 \times 4} \end{pmatrix} \begin{pmatrix} p_{3 \times 1} \\ 1 \end{pmatrix} = 0 \quad (8)$$

위 식에서,  $H_{i,j}$ 는  $i$ 번째와  $j$ 번째에서 사영 복원된 점들 ( $X_i, X_j$ )간의 정규화된  $4 \times 4$  호모그래피를 의미하며,  $X_j = H_{i,j} X_i$ , 계산하는 방법에 대한 연구는 [20]에 제시되어 있다.

1.3 메트릭 보정 (Metric calibration)

유클리드 보정은 식 (5)에 표현된 자동 보정의 나머지 파라미터인 카메라 내부 파라미터를 구하는 과정이다. 다수 카메라들의 강제 변환은  $H_{ij} = T D_{ij} T^{-1}$  관계를 제공한다<sup>[13]</sup>. 이 관계를 1.2절에서 구해진  $p$ 를 제거하여, 전개하면 다음과 같이  $H_{ij}$ 는  $H_{ij}^A = T_A D_{ij} T_A^{-1}$ 로 변환된다.  $T_A$ 는 어파인 복원을 유클리드 복원으로 향상시키는 변환 행렬이다.

$$\begin{aligned} H_{i,j}^A &= \begin{pmatrix} K 0 \\ 0 1 \end{pmatrix} \begin{pmatrix} R_{i,j} t_{i,j} \\ 0 \quad 1 \end{pmatrix} \begin{pmatrix} K^{-1} & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} K R_{i,j} K^{-1} & K t_{i,j} \\ 0 & 1 \end{pmatrix} \end{aligned} \quad (9)$$

식 (9)에서 드러나듯이  $H_{ij}^A$ 는  $K R K^{-1}$ 라는 무한 호모그래피 (Infinite Homography)를 포함하고 있다<sup>[13]</sup>. 무한 호모그래피에서  $K$ 를 추출하는데,  $R R^T = I$ 라는 성질을 이용한다<sup>[1]</sup>.

1.4 직접적인 방법 (Direct approach)

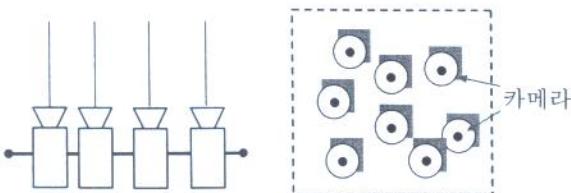
1.1절에서 사영 복원을 수행한 후, 1.2절과 1.3절의 과정을 통한 방법은 단계적인 방법이다. 단일 카메라의 경우에도 투영 행렬로부터 DIAQ (Dual Image of Absolute Quadric) 을 활용한 직접적인 방법이 존재한다<sup>[8]</sup>. 다수 카메라 시스템에도 이 알고리즘의 활용이 가능하다. 이를 위해서는 호모그래피를 이용한 투영 행렬 생성이 요구된다. 즉, 1.1절의 사영 복원 후 획득한  $i$ 번째 위치의  $k$ 번째 투영 행렬을  $P_k^i$ 라고 하면, 다수 카메라 시스템의  $j$ 번째 위치에서의 내부  $k$ 번째 카메라의 투영 행렬은 다음과 같다.

$$P_k^j = P_k^i \times H_{i,j} \tag{10}$$

그러므로 총  $M$ 개의 카메라가  $N$ 번 움직일 경우, 총  $M \times (N-1)$ 개의 투영 행렬을 생성하게 된다. 생성된 투영 행렬을 [8]에 의해 제안된 방법에서 사용하면 어파인 복원을 거치지 않고 카메라 파라미터를 얻는다.

2. 평면 구성을 지닌 다수 카메라 시스템의 자동 보정

본 논문에서 다루는 평면 구성은 다수 카메라를 이루는 카메라들의 광축이 평행하고, 동일 평면상에 카메라가 위치하여 상대적인 위치 관계가  $[a, \beta, 0]^T$ 으로 표현되는 것을 의미한다. 그림 4는 평면 구성 카메라의 예를 보인 것이다.



1차원 배열 타입 (위) 2차원 배열 타입 (정면)

그림 4. 평면 구성을 지닌 다수 카메라 시스템의 예 : 왼쪽은 1차원 배열로 이루어진 시스템, 오른쪽은 평면상에 평행하게 위치한 시스템

Fig 4. Examples of a multi-camera system with a planar configuration, left: 1D array type, right : 2D array type

2.1 어파인 투영 행렬 (Affine projection matrix)

가정한 평면 구성에서는 카메라들의 외부 파라미터의 회전 행렬  $R = I$ 가 되고,  $t = [a, \beta, 0]^T$ 라는 형식의 값을 가지게 된다. 이를 반영하여 식 (2)를 다시 정리하면 다음과 같다.

$$P^E X^E = K(Rt) X^E = \begin{pmatrix} f & s & C_x & \alpha \\ 0 & f & C_y & \beta \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{11}$$

어파인 분해법을 얻기 위해서는 알려진 정보를 활용해야 한다. 즉, 카메라 모델 중에서 주점 ( $C_x, C_y$ )이 알려져 있으며,  $s = 0$ 이라고 가정하자. 이는 기존의 여러 보정 알고리즘에서 취한 일반적인 가정이다. 그리고 3차원 점인  $X^E = [X/Z, Y/Z, 1, 1/Z]^T$ 라고 치환한다. 크기 인자가 정해지지 않았으므로 치환에 따른 오류는 없게 된다. 이를 모두 적용한 투영 행렬은 다음과 같다.

$$P^E X^E = \begin{pmatrix} f & 0 & 0 & \alpha \\ 0 & f & 0 & \beta \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X/Z \\ Y/Z \\ 1 \\ 1/Z \end{pmatrix} \tag{12}$$

위 식의 형태를 관찰하면,  $P^E$ 의 3번째 열과 행이  $X^E$ 와 곱해지는 경우 의미가 없음이 드러나게 된다. 우선, 의미 없는 행과 열을 제거한다. 그 후,  $P^E$ 를  $P^A = P^E \times T^{-1}$ 로 변환한다. 최종적으로 식 (12)으로부터 (13)와 같은 어파인 투영 관계를 얻는다.

$$w \cong QS \cong \begin{pmatrix} q_1 & 0 & q_3 \\ 0 & q_2 & q_4 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \tag{13}$$

위 식에서,  $w = [x, y]^T$  영상의 점들이고,  $[s_1, s_2, s_3] = [X/Z, Y/Z, 1/Z]$ 이다.  $Q$ 의 내부 요소는 스케일이 결정되지 않은 어파인 투영 행렬의 파라미터이다.

2.2 어파인 분해법 (Affine factorization)

2.1절에서 증명했듯이, 본 논문에서 가정한 평면 구성의 다수 카메라 시스템에서는 어파인 투영 행렬과 어파인 복원점으로 투영이 이루어졌다. 이를 다수 카메라 시스템 전체로 확장한다. 이는 수식 (13)의 확장으로 표현된다.

$$\begin{pmatrix} w_1^1 & w_1^2 & \dots & w_1^N \\ w_2^1 & w_2^2 & \dots & w_2^N \\ \vdots & \vdots & \ddots & \vdots \\ w_M^1 & w_M^2 & \dots & w_M^N \end{pmatrix} \cong \begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_M \end{pmatrix} (S^1 S^2 \dots S^N) \quad (14)$$

w가 2×(특징점 수) 크기이므로, 위 식의 좌측 행렬은 2M×(N×(특징점 수)) 크기가 된다. 특징점을 이용하여 이 행렬을 채우고, SVD를 수행한 후, 행렬의 계급(Rank)이 3인 제한 인자를 이용하여, 사영 공간에서의 분해법과 같은 과정을 수행하여 Q와 S 행렬로 분해한다.

분해법 수행 후 얻어진 Q와 S는 정확하게 수식 (13)의 형태를 따르지 않는다. 구해진 Q와 S는 임의의 3×3행렬 U가 곱해진 형태이다. 그러므로 임의의 행렬을 밝혀진 제한 요소를 활용하여 구한 후, 이를 구해진 Q와 S에 곱해 올바른 Q와 S를 구한다. 제한 요소는 식 (13)의 Q의 형태에서 얻는다.

$$\begin{aligned} Q_{1 \times 2} &= Q_{2 \times 1} = 0 \\ Q_{1 \times 2} - Q_{2 \times 1} &= 0 \text{ if } \tau = 1 \end{aligned} \quad (15)$$

Q×U는 위 조건을 동일하게 만족해야 한다. 구해진 U<sup>1</sup>을 구해진 S에 곱함으로써, 어파인 공간에서의 복원된 점들 또한 역치환을 통해 획득할 수 있다.

2.3 메트릭 보정 (Metric calibration)

마지막 과정으로 2.2절에서 추출한 어파인 공간에서의 투영행렬과 복원점을 이용하여 본 절에서는 카메라 파라미터 K를 계산한다. 1.3절에서 설명한 메트릭 보정과 동일한 방식을 사용한다. 이를 위해 사영 공간에서 복원 후 계산되었던 4×4 호모그래피처럼, 어파인 공간에서도 4×4 어파인 호

모그래피 (H<sub>A</sub>)를 계산한다. 식 (9)에 정의된 형태를 띤다.

H<sub>A</sub>는 어파인 공간의 속성 중 중심점 불변을 이용하여 구한다. 강건한 X<sub>A</sub><sup>i</sup>와 X<sub>A</sub><sup>j</sup>간의 변환 관계를 구하는 세부 과정은 다음과 같이 정리할 수 있다.

1. 복원점들 중 적정한 (40이상) 수의 점쌍들 추출
2. 각각의 중심 좌표 <X<sub>A</sub><sup>i</sup>>, <X<sub>A</sub><sup>j</sup>>를 구함
3. 좌표 변환 수행 : X<sub>A</sub><sup>i</sup> - <X<sub>A</sub><sup>i</sup>>, X<sub>A</sub><sup>j</sup> - <X<sub>A</sub><sup>j</sup>>
4. 최소 제곱법 (Least Square Method)를 이용한 H<sub>3×3</sub><sup>A</sup> 구하기  
(RANSAC<sup>[1]</sup> 방식을 이용 이상점 (outlier) 제거)
4. H<sub>A</sub>의 나머지 부분 구하기  
<X<sub>A</sub><sup>j</sup>> - H<sub>3×3</sub><sup>A</sup> <X<sub>A</sub><sup>i</sup>>
5. 구해진 H<sub>A</sub>를 활용하여, 모든 점들에 대한 오차값 계산 :  
|| X<sub>A</sub><sup>j</sup> - H<sub>A</sub>X<sub>A</sub><sup>i</sup> ||
6. 오차값이 어느 정도의 임계치를 만족할 때까지 1-5의 과정을 반복함

3. 비선형 최적화 (Bundle adjustment)

제안한 알고리즘은 선형 연산만으로 카메라 파라미터를 획득한다. 선형 연산의 단점은 잡음에 민감하기 때문에, 결과의 정확성을 보장하기 위해서는 잡음을 최소화하는 비선형 최적화 작업이 동반되어야만 한다. 다수 카메라 시스템에서는 다음 유클리드 복원 오차 함수를 최적화한다.

$$\frac{1}{N \times M} \sum_{j=1}^N \sum_{i=1}^M \| x_i^j - P_i^{j,E} X^{j,E} \|^2 \quad (16)$$

위 식은 M개의 내부 카메라를 지닌 다수 카메라가 N번 움직였을 경우, 하나의 카메라에서의 복원된 점을 복원된 카메라 투영행렬로 재투영 (Reprojection)했을 경우 관찰된 영상점과 추정된 투영된 영상점간의 평균 거리 오차이다.

다수 카메라 시스템의 경우, 실제 L-M 구현 시 최적화 과정에서 단일 카메라의 경우보다 규모가 상당히 큰 행렬이 생성되므로<sup>[21]</sup>, 일반적인 행렬 연산보다는 희소 (Sparse) 행렬 연산으로 구현해야 컴퓨팅 자원의 효율성을 높이는 방법이다.

### V. 실험 및 결과

제안된 논문의 강건성과 활용성을 보이기 위해 본 장에서는 다양한 자동 보정 실험 결과를 보인다. 본 논문에서 제안된 모든 알고리즘들은 공개 소스인 OpenCV<sup>[24]</sup>를 이용하여 Windows 환경에서 구현되었고, 카메라 보정의 결과로 획득하는 카메라의 위치와 복원된 점들의 가상 렌더링은 OpenGL을 이용하여 렌더링하였다.

#### 1. 시뮬레이션 데이터를 이용한 실험 결과

시뮬레이션 데이터 생성에는 총 6대의 가상 카메라를 가진 다수 카메라 시스템이 사용되었다. 다수 카메라 시스템은 평면 구성의 카메라 구조와 일반적인 구조를 갖춘 두 가지 경우를 테스트 하였다. 총 6대의 가상 카메라는 600.0 (픽셀)에서 650.0 (픽셀) 사이의 서로 다른 초점 거리를 지니며, 영상 평면은 640 픽셀 × 480 픽셀로 설정되었다. 실험에서는 주점 (320 픽셀 × 240 픽셀로 가정)이 알려져 있고, 뒤틀림 인자 s는 0이라고 가정하였다. 가상의 3차원 객체는 91개의 특징

점들을 지닌 삼면 직각 모형을 사용하였다. 그림 5는 실험에서 시뮬레이션한 다수 카메라 시스템을 보여준다. 영상 특징점 획득을 위해 다수 카메라는 총 6번 이동하였다.



그림 5. 시뮬레이션에 사용된 가상의 6대 카메라로 이루어진 다수 카메라 시스템과 91개의 특징점을 지닌 직각인 가상 패턴

Fig 5. The simulated multiple camera system with 6 cameras configuration and 3D orthogonal pattern with 91 feature points

첫 번째 실험에서는 알고리즘의 강건성을 평가하기 위해, 임의의 평균 0.0, 표준 편차 (Standard Deviation)가 0.0에서 0.30까지인 가우시안 (Gaussian) 잡음을 수집된 영상 평면 특징점에 부가한 후 알고리즘을 수행하였다. 모든 실험 결과는 30번씩 반복하여 수행된 값의 평균치이다. 그림 6에서 'DIAQ'는 3장에서 제안한 방법 중 직접 방법으로 보정한 결과를 의미하고, 'Stratified'는 단계적인 방법을

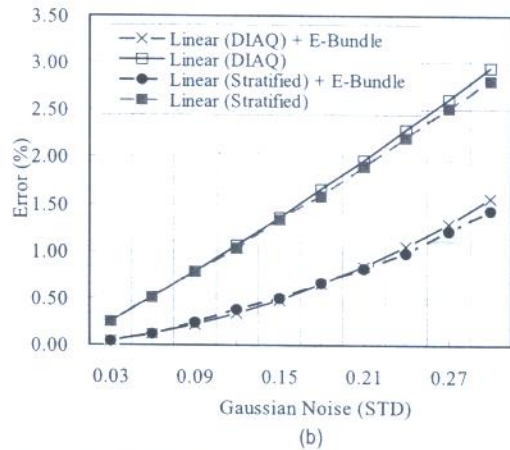
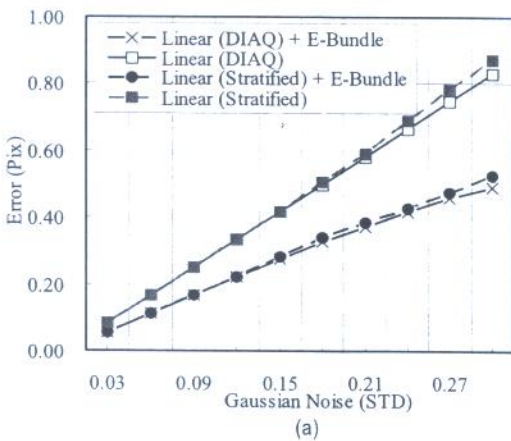


그림 6. 가우시안 잡음에 따른 다수 카메라 자동 보정 알고리즘 복원 오차값, a) 유클리드 복원 후 재투영 (Reprojection, 수식 (16)) 오차 b) 카메라 초점거리 오차 비율 평균:  $\frac{\|추정값 - 참값\|}{참값} \times 100$

Fig 6. Experimental results according to Gaussian noise (standard deviation), a) Reprojection errors of Euclidean reconstruction, b) Focal length errors =  $\frac{\|Estimated\ value - Ground\ truth\|}{Ground\ truth} \times 100$

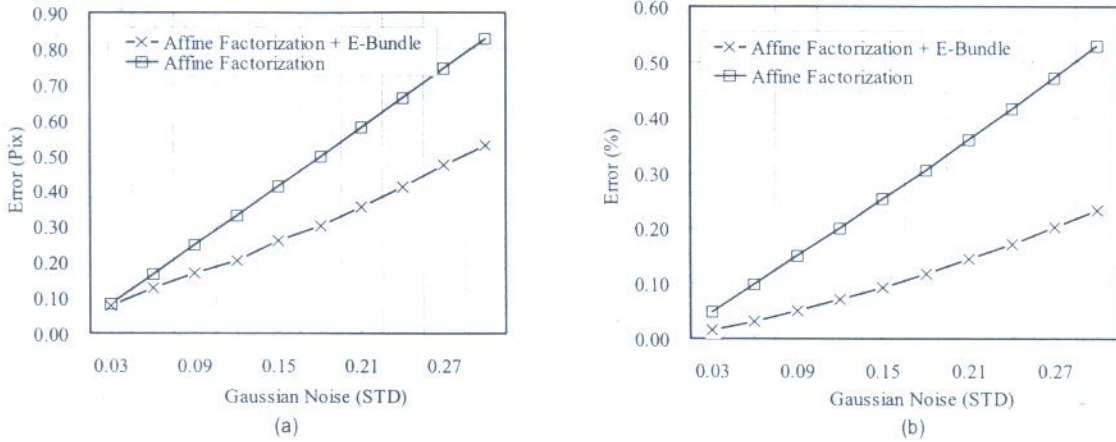


그림 7. 가우시안 잡음에 따른 평면 구조를 갖는 다수 카메라 자동 보정 알고리즘 (어파인 복원 방법 이용) 복원 오차값 a) 재투영 오차 b) 초점 거리 오차  
 Fig 7. Experimental results using affine factorization techniques according to Gaussian noise (standard deviation), a) Reprojection errors b) Focal length errors

각각 나타낸다. 'E-Bundle'은 비선형 최적화를 L-M 방법으로 유클리드 복원 후 수행했다는 것을 의미한다.

실험에서 비선형 최적화를 수행하면 선형 결과로 얻은 오차를 약 1/2정도로 낮출 수 있음이 관찰되었다. 이 결과는 두 가지 알고리즘에서 공통적으로 관찰되었으며, 잡음이 증가함에 따라 그 효과를 뚜렷하게 관찰할 수 있었다. 3장에서 제안된 'DIAQ' 직접방법과 'Stratified' 단계적방법은 재현된 시뮬레이션 모션에서 비슷한 결과를 보여주고 있다.

동일 조건아래, 제안한 평면 구조에 특화된 어파인 분해법 기반 알고리즘을 수행하였다. 그림 7에서 'Affine Factorization'은 선형 연산으로 구한 오차를 나타내고, 'E-Bundle'은 L-M 최적화 수행 후 오차를 의미한다. 'DIAQ'나 'Stratified'와 경우와 비슷한 결과를 보이지만, 약간의 재투영 오차의 개선을 가져온 것을 알 수 있으며, 특히 초점 거리의 오차는 상대적으로 급감함을 관찰할 수 있다.

2. 실제 영상을 이용한 실험 결과

제안된 알고리즘을 테스트하기 위해 PointGrey 사의 Digiclops를 사용하였다<sup>[25]</sup>. Digiclops는 다수 카메라 시스템

으로 장착된 카메라 3 대가 모두 평면 구성을 지니고 있다. 각 카메라의 영상 평면 해상도는 640 픽셀 × 480 픽셀이다.

실험은 두 종류의 영상을 이용하였고, 총 8번 카메라를 움직였다. 첫 번째 영상은 실제 144개의 특징점을 지닌 직각 모양의 보정 패턴이 포함된 영상을 이용하여 실험을 수행하였고, 두 번째는 일반적인 영상을 이용하여 알고리즘을 수행하였다. 그림 8은 실험에 사용된 두 종류의 영상이다.

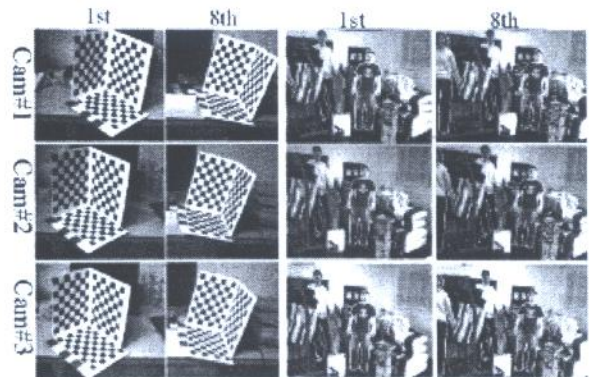


그림 8. 실험에 사용된 실제 영상, 좌) 패턴 영상 포함한 영상들 중 1번째와 8번째 영상 프레임, 우) 일반적인 실내 환경의 영상들  
 Fig 8. Real images, left) 1st and 8th image sets including the orthogonal pattern, right) 1st and 8th image sets including general scene

실험 결과는 표 1과 같다. 실험 결과에서는 패턴과 실제 영상 모두에서 어파인 분해법 알고리즘이 향상된 복원 오차를 보이는 것으로 나왔다. L-M 최적화 후 수행 결과는 초기 선형 연산으로 구했던 복원 오차를 2/3 정도로 줄여 줌을 확인할 수 있었다. 초점 거리 결과는 모두 기준 카메라에 해당하는 값으로, 840.2923 픽셀이 평면 패턴을 이용한 [7]의 알고리즘으로 구해졌다. 이 값을 근거로 살펴보면, 어파인 분해법 기반이 평면 구조 카메라 시스템에는 적합하다는 사실을 확인할 수 있었다. 그림 9는 패턴 영상에서 나온 결과를 토대로, 다수 카메라들의 위치들을 복원

표 1 실제 영상을 이용한 자동 보정 알고리즘 실험 결과  
Table 1. Experimental results of the proposed self-calibration algorithm using real images

방법 측정값		직접적 방법 (DIAQ)	단계적 방법 (Stratified)	어파인 분해법
복원 오차 (픽셀)	패턴 영상	0.3336 (0.2180)	0.3037 (0.2070)	0.2976 (0.1970)
	실제 영상	0.6764 (0.4590)	0.6975 (0.4610)	0.5976 (0.3913)
초점 거리 (픽셀)	패턴 영상	897.3947 (897.4235)	915.1437 (914.9432)	826.8544 (826.3968)
	실제 영상	866.7356 (866.8210)	930.2311 (931.0051)	845.8394 (838.9965)

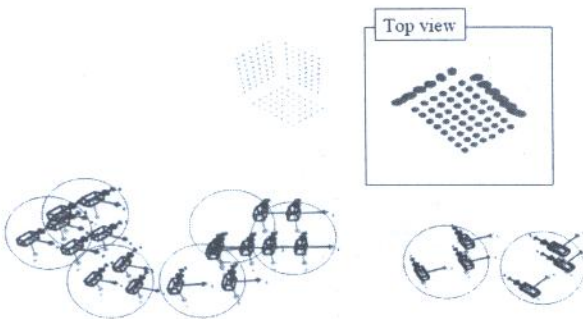


그림 9. 8번의 움직임을 통해 획득한 패턴 영상을 이용하여 복원된 패턴상의 모습과 카메라 위치들, 3대의 카메라로 이루어진 다수 카메라 시스템의 8번 움직임 보여줌

Fig 9. The 3D pattern and real camera poses reconstructed by using 8 movements of the multiple camera system having 3 cameras

한 결과이다. 3대의 카메라로 구성된 다수 카메라 시스템 이므로, 각 원들의 카메라들이 시스템을 이루는 카메라들의 묶음이다. 그림 9에서 보듯 직각 패턴이 실제 직각 패턴에 근접한 형태를 보여준다.

실제 영상에서 최소 다수 카메라인 스테레오 카메라의 자동 보정을 수행하였다. 그림 10에서와 같은 실내 공간에서 스테레오 카메라를 왼쪽에서 오른쪽으로 회전시키면서 5번 영상을 획득하였다. 그림 10은 획득된 영상 중 하나로, 원영상에 추출된 특징점의 위치, 그리고 기준영상에서의 특징점과 동일한 점들의 관계를 표시한 결과를 보여준다. 추출된 특징점을 이용하여 자동 보정을 수행한 후 얻어지는 특징점들의 3차원 복원점들을 보였으며, 카메라의 위치를 동시 렌더링하였다 (카메라는 회전만 가했기 때문에, 그림에서처럼 축만 뒤틀려 보이고, 전체적으로 겹쳐 보임). 3차원 복원점들을 위에서 바라볼 경우, 실제 환경에서의 직각인 모습과 비슷함을 알 수 있으며, 장면의 앞쪽에 위치하는 박스의 구조 또한 바르게 복원되었음을 알 수 있다. 정밀한 3차원 복원은 본 논문의 결과물을 활용하면 가능할 것이라 기대한다.

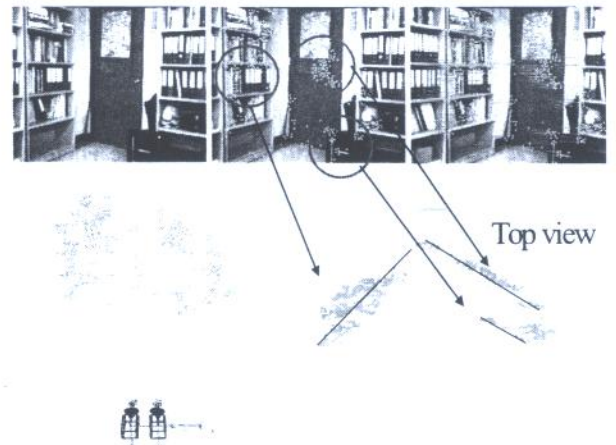


그림 10. 스테레오 카메라를 5번 움직여 획득한 실제 영상에 제안된 알고리즘 적용 후 획득한 특징점 (초기 장면 구조)의 3차원 복원점과 카메라 위치

Fig 10. Initial scene structure and poses of cameras after applying the proposed method with stereo camera : 10 images (5 movements) are used in this experiment

## VI. 결 론

본 논문에서는 움직이는 다수 카메라 시스템에 적합한 새로운 분해법 기반 알고리즘을 제안하였다. 특히, 일반적인 구조를 갖는 다수 카메라 시스템 알고리즘과 평면 구조를 갖는 다수 카메라 시스템에 최적화된 알고리즘을 제시하였다. 또한, 다수 카메라 대신 움직이는 객체 (보정 직각 패턴)의 추적이 가능할 경우, 고정된 다수 카메라 시스템에도 알고리즘의 수정 없이 사용가능 하다. 장착되는 카메라의 개수 (최소 2대 이상)에 영향을 받지 않는다는 장점이 있다. 이러한 유연함은 실제 다수 카메라 시스템의 구축 및 활용에 큰 도움을 줄 것이라 생각한다.

실험에서는 알고리즘의 활용성과 강건성을 점검하였다. 시뮬레이션으로 생성된 직각 패턴과 가상 평면 구성의 카메라를 이용하여, 가우시안 잡음에 따른 보정 결과를 보였으며, 비선형 최적화 전과 후의 결과를 통해 최적화 과정이 1/2 정도의 결과를 향상시킴을 보였다. 또한, 제안된 어과인 복원 방법을 사용하여 동일 조건아래 수행한 보정 결과를 보였다. 시뮬레이션 과정을 기반으로, 실제 3대의 카메라로 구성된 다수 카메라를 통해 획득된 영상을 이용한 실험을 수행하였으며, 초점 거리, 투영 오차와 복원된 카메라 위치를 정리하였다. 향후에는 제안된 알고리즘을 이용, 다수 카메라 시스템을 이용한 조밀한 3차원 복원에 대한 연구와 증강 현실에의 응용에 대한 연구를 진행할 예정이다.

## 참 고 문 헌

- [1] R. Hartley and A. Zisserman, "Multiple view geometry", 2nd Edition, Cambridge University Press, 2003.
- [2] F. Jean-Sebastien, Lapierre, Marc and E. Boyer, "Visual Shapes of Silhouette Sets", 3DPVT, Chapel Hill (USA), 2006.
- [3] M. Pollefeys, R. Koch, M. Vergauwen, L. Van Gool. "Automated reconstruction of 3D scenes from sequences of images", ISPRS Journal Of Photogrammetry And Remote Sensing (55)4 pp. 251-267, 2000.
- [4] M. Pollefeys and L. Van Gool. "A stratified approach to metric self-calibration", CVPR, pp.407-412, Puerto Rico, 1997.
- [5] M. Pollefeys, L. Van Gool, "Stratified Self-Calibration with the Modulus Constraint", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 21, No.8, pp.707-724, 1999.
- [6] P. Sturm, "Critical Motion Sequences for Monocular Self-Calibration and Uncalibrated Euclidean Reconstruction", CVPR, pp.1100, 1997.
- [7] Z.Zhang, "A Flexible New Technique for Camera Calibration", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol."22", pp.1330-1334, 2000.
- [8] B. Triggs, "Autocalibration and the Absolute Quadric", CVPR, pp.609-714, 1997.
- [9] B. Triggs, "Autocalibration from Planar Scenes", ECCV, vol. 1, pp.89--105, June, 1998.
- [10] A. Zisserman, P. Beardsley, and I. Reid, "Metric calibration of a stereo rig", in Proc. IEEE Wksp. Representation Visual Scenes, pp. 93 - 100, June 1995.
- [11] F. Devernay and O. Faugeras, "From projective to Euclidean reconstruction", CVPR, pp.264-269, San Francisco, CA, 1996.
- [12] G. Csurka, D. Demirdjian, A. Ruf and R. Horaud, "Closed-form solutions for the Euclidean calibration of a stereo rig", ECCV, pp.426-442, 1998.
- [13] R. Horaud, G. Csurka and D. Demirdjian, "Stereo Calibration from Rigid Motions", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol."22", pp.1446-1452, 2000.
- [14] D. Demirdjian, A. Zisserman and R. Horaud, "Stereo Autocalibration from One Plane", ECCV, pp.625-639, 2000.
- [15] J. Knight and I. Reid, "Self-Calibration of a Stereo Rig in a Planar Scene by Data Combination", ICPR, vol."1", pp.1411, 2000.
- [16] F. Dornaika, "Self-calibration of a stereo rig using monocular epipolar geometry", ICCV, vol.2, pp.467-472, 2001.
- [17] F. Dornaika and C. Chung, "Stereo geometry from 3D egomotion streams", IEEE Transactions on Systems, Man and Cybernetics, Part B, Volume 33, Issue 2, pp.308 - 323, April 2003.
- [18] T. Svoboda and D. Martinec and T. Pajdla, "A Convenient Multi-Camera Self-Calibration for Virtual Environments", PRESENCE: Teleoperators and Virtual Environments, pp 407-422, 14(4), August 2005.
- [19] A. Zaharescu, R. Horaud, R. Ronfard, L. Lefort, "Multiple Camera Calibration using Robust Perspective Factorization", 3DPVT, 2006.
- [20] G. Csurka, D. Demirdjian and R. Horaud, "Finding the Collineation between Two Projective Reconstructions", CVIU, vol.75, pp.260-268, 1999.
- [21] B. Triggs, P. McLauchlan, R. Hartley and A. FitzgibbonR, "Bundle Adjustment - A Modern Synthesis", Vision Algorithms: Theory and Practice, LNCS, pp.298-375, 2000.
- [22] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: a factorization method", IJCV, vol.9(2), pp.137-154, 1992.
- [23] P. Sturm, and B. Triggs, "A Factorization Based Algorithm for

multi-Image Projective Structure and Motion", ECCV, pp.709-720, 1996.

[24] <http://sourceforge.net/projects/opencvlibrary/>,

[25] <http://www.ptgrey.com/>

---

## 저 자 소 개



김 기 영

- 2003년 : 중앙대학교 컴퓨터공학과 (학사)
- 2004년 : 광주과학기술원 정보통신공학과 (석사)
- 2004년 ~ 현재 : 광주과학기술원 정보통신공학과 박사 과정
- 주관심분야 : 3D 컴퓨터비전, 증강 현실 (Augmented Reality), HCI



우 운 택

- 1989년 : 경북대학교 전자공학과 (학사)
- 1991년 : 포항공과대학교 전기전자공학과 (석사)
- 1998년 : University of Southern California, Electrical Engineering-System (박사)
- 1999년 ~ 2001년 ATR. 초빙 연구원
- 2001년 ~ 현재 광주과학기술원 정보기전공학부 부교수
- 2005년 ~ 현재 광주과학기술원 문화기술연구센터장
- 주관심분야 : 3D 컴퓨터비전, 증강현실, HCI, 유비쿼터스 컴퓨팅, 컨텍스트 인식, 문화기술 등