

## 불변 특징모델을 이용한 카메라 동작인수 측정

차 정 회\*, 이 근 수\*\*

### Estimation of Camera Motion Parameter using Invariant Feature Models

Jeong-Hee Cha\*, Keun-Soo Lee\*\*

#### 요 약

본 논문에서는 카메라의 뷰포인트에 무관한 효율적인 불변특징을 기반으로 카메라의 동작인수를 산출하는 방법을 제안한다. 기존연구에서 사용된 특징정보는 카메라의 뷰포인트에 따라 변하기 때문에 정보양이 증가하여 정확한 특징추출이 어렵다. 또한 카메라 외부인수 산출을 위해 사용되는 LM(Levenberg-Marquardt) 방법은 정확하게 목표 값에 수렴하지만 작은 스텝크기로 최소화를 진행하므로 소요시간이 긴 단점이 있다. 따라서 본 논문에서는 뷰포인트에 무관한 불변특징 추출방법과 이 특징들을 이용하여 2D 호모그래피로 찾은 카메라 동작인수를 LM 방법의 초기값으로 사용, 정확성과 수렴도를 향상시키는 2단계 카메라 동작인수 산출 방법을 제안한다. 제안하는 방법은 특징 추출단계, 정합 단계, 2단계 카메라 동작인수 산출단계로 구성된다. 실험에서는 다양한 실내영상으로 제안한 방법과 기존 방법을 비교, 분석함으로써 제안한 알고리즘의 우수성을 입증하였다.

#### Abstract

In this paper, we propose a method to calculate camera motion parameter, which is based on efficient invariant features irrelevant to the camera viewpoint. As feature information in previous research is variant to camera viewpoint, information content is increased, therefore, extraction of accurate features is difficult. LM(Levenberg-Marquardt) method for camera extrinsic parameter converges on the goal value exactly, but it has also drawback to take long time because of minimization process by small step size. Therefore, in this paper, we propose the extracting method of invariant features to camera viewpoint and two-stage calculation method of camera motion parameter which enhances accuracy and convergent degree by using camera motion parameter by 2D homography to the initial value of LM method. The proposed method are composed of features extraction stage, matching stage and calculation stage of motion parameter. In the experiments, we compare and analyse the proposed method with existing methods by using various indoor images to demonstrate the superiority of the proposed algorithm.

▶ Keyword : 카메라 동작인수(camera motion parameter), LM 방법(Levenberg-Marquardt) Method, 2D 호모그래피(2D Homography), 불변특징(Invariant Feature)

• 제1저자 : 차정회  
• 접수일 : 2005.07.31, 심사완료일 : 2005.09.05  
\* 송실대학교 컴퓨터학부 박사과정, \*\* 한경대학교 컴퓨터공학과 교수

## I. 서론

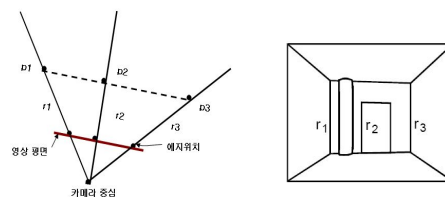
컴퓨터의 새로운 패러다임으로 등장한 유비쿼터스화는 유비쿼터스 컴퓨팅과 유비쿼터스 네트워크를 기반으로 물리 공간을 지능화함과 동시에 물리공간에 펼쳐진 각종 사물들을 네트워크로 연결시키려는 노력이다[1]. 유비쿼터스 컴퓨팅은 모든 컴퓨터가 서로 연결되고 이용자 눈에 보이지 않으며 언제 어디서나 사용 가능하고 현실세계의 사물과 환경 속으로 스며들어 일상생활에 통합되는 것을 기본 전제로 하는데, 이중 중요한 요소 기술인 위치인식 기술이 현재 선진 각국에서 활발한 연구가 진행되고 있다. 위치인식에는 센서를 이용하는 방법과 컴퓨터 비전 기술을 이용하는 방법, 서로 다른 센서 정보들을 융합(fusion)하여 사용하는 방법 등이 있다. 컴퓨터 비전기술을 이용할 경우 센서가 가지고 있는 오차의 한계점을 극복할 수 있는데, 이는 센서 정보보다 정확한 영상 정보를 이용하기 때문이며, 위치인식을 위해 거리나 각도와 같은 지리적인 특징이 필요하지 않기 때문이다. 그러나 위치인식에 사용되는 특징들이 카메라의 자세에 따라 변하므로 정보량이 증가한다는 문제점이 있다. 또한 카메라 외부인수 산출을 위해 사용되는 LM 방법은 오차를 최소화함으로써 최적화된 해를 산출하나 초기치에 따라 최소점에 근사되는 반복회수가 다르고 지역적인 최소점에 빠질 경우 수렴시간이 길어진다는 단점이 있다. 따라서 본 논문에서는 환경변화에 따라 증가하는 정보 양을 줄이기 위해 효율적인 불변특징을 이용하여 두 영상의 정확한 대응점을 확인하고 구해진 대응점들을 이용하여 카메라의 동작인수를 산출할 때 정확성과 수렴도를 향상시키는 2단계 카메라 동작인수 산출방법을 제안한다.

## II. 기존의 카메라 위치인식 방법

광학 센서를 이용하는 비전기반 위치인식의 경우 환경 표현방법은 점이나 선 같은 매우 간단한 특징의 형태로 주어지기도 하고 복잡한 패턴인 3차원 환경이나 모델과 같은 형태로 주어지기도 하는데, 여기서는 본 논문의 내용과 관련된 있는 몇 가지를 간단하게 기술하고자 한다.

### 2.1 한대의 카메라를 사용한 방법

환경지도가 있고, 지표면과 평행한 광축을 가진 한대의 카메라가 차량이나 이동 로봇위에 탑재되어 있다면, 주변의 수직 모서리가 랜드마크가 될 수 있다[2]. 이 경우 위치인식 문제는 두 가지 경우를 고려할 수 있는데, 첫 번째는 영상에 나타난 각 수직 모서리가 서로 구분이 가능하고 각 수직 모서리의 높이 방향으로 변화가 없어서 2차원 지도상에 수직 폴(vertical pole)로 표시되는 경우이다. 이러한 경우 수직모서리는 점 특징이 되고, 2차원상에서의 위치인식은 세 개의 유일한 특징들의 인식을 필요로 한다. 만일 특징들이 유일하게 인식되고 특징들의 위치가 알려져 있으면, 핀홀 카메라의 위치와 방향은 (그림 1)의 (a)와 같이 결정될 수 있다. 두 번째 경우는 (b)와 같이 각 수직 모서리의 구분이 불가능한 경우로, 이때는 환경지도상에 나타나는 수직 폴과 그에 해당하는 측정된 수직 모서리와의 가능한 모든 대응관계를 조사해야 해를 찾을 수가 있다. 즉, 모든 가능한 해를 고려하는 복잡한 절차를 거쳐야 한다. 또한 환경지도가 주어지지 않으면 위치인식을 할 수 없게 된다.



(a)특징점 구분이 가능한 경우 (b)구분이 어려운 경우  
 그림 1. 한대의 카메라를 이용한 위치인식방법  
 Fig 1. Localization using Single Camera

### 2.2 스테레오카메라를 사용한 위치인식 방법

관찰된 랜드마크와 작성된 지도사이의 대응성을 결정하고 그 대응성으로부터 센서의 2차원 위치를 추정하기 위하여 스테레오 카메라 쌍을 이용하는 방법이 있다[3]. 이 방법은 두 대의 카메라로 획득된 스테레오 영상을 이용해서 관찰된 점들의 2차원 위치들을 결정하고, 모든 가능한 지도상의 세 개의 점과 이 점들의 관측 위치사이의 대응성을 인식함으로써 위치인식을 수행한다. 이러한 관측된 자료와 저장된 지도상의 점들은 이동과 회전에 불변한 변환식을 찾아 내어 적용해야 대응점들을 찾을 수 있다.

### 2.3 카메라 보정을 통한 위치인식 방법

이 방법은 추출된 특징점 집합들로부터 카메라의 내부인자와 외부인자를 계산하는 과정 때문에 앞서 설명된 2차원 위치인식 알고리즘보다 복잡하다. 카메라 보정을 통한 위치인식이란 기준 좌표계에 대한 카메라의 3차원 위치와 방향을 설정하는 것을 말하며, 내부 인자는 초점거리, 렌즈 왜곡변수와 확대 변수(scale parameter)를 말하고, 외부변수는 카메라의 상태를 파악할 수 있는 회전변수(rotational parameter)로 구성된다. 계산된 외부변수로부터 물체나 특징점이 나타나는 세계 좌표계에 대한 카메라 좌표시스템의 3차원위치나 회전정보를 획득할 수 있다. 분석에 의하면 12개의 변수(9개의 회전 변수와 3개의 이동변수)를 구하기 위하여 제약사항에 의해 여러 단계로 해를 분해하게 되므로 유일한 해를 구하기 위해서는 적어도 6개의 특징 대응점들이 필요하다. 그러나 모든 3차원 점들이 동일한 평면상에 존재 한다면 최소 4개의 특징점에 대한 정보만으로도 구할 수 있게 된다.

## III. 특징추출 단계

(그림 2)는 본 논문에서 제안하는 전체 시스템 구성도이다. 특징 추출단계에서는 입력영상에서 두드러진 영역을 찾아낸 후 코너점을 추출하고 전체 벡터값의 평균값 이상을 지닌 강건 특징모델 후보를 형성한다. 학습과 인식단계에서는 5개의 점 집합이 동일평면상의 점인지와 이중 세 점이 동일선상에 존재하는지를 검사하고 벡터오더링 기법을 이용하여 전체 벡터값의 평균값 이상을 지닌 강건 특징모델 후

보를 형성한다. 정합 단계에서는 형성된 가정을 블록면 검사로 대응성이 확인된 두 영상의 변환행렬을 계산한다. 마지막으로 카메라 외부인수 측정단계에서는 대응성이 확인된 두 영상의 특징점 좌표를 이용하여 카메라 외부인수를 산출한다. 본 논문에서 사용한 뷰포인트에 무관한 특징은 평면상에 존재하는 5개의 점으로 계산되는 기하학의 불변 특징 벡터이다. 이장에서는 불변특징을 추출하기 위해 사용된 사영기하학의 복비와 특징모델을 정의하고 특징 추출 과정을 설명한다.

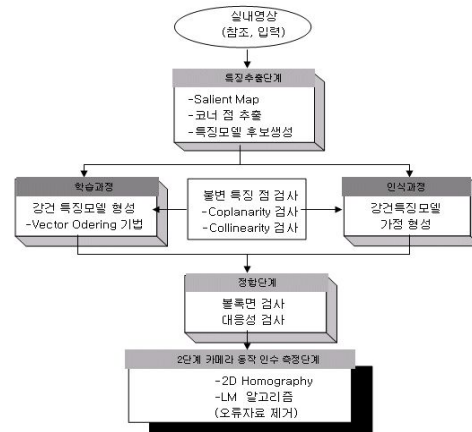


그림 2. 전체 시스템 구성도  
Fig 2. Overall System Configuration

### 3.1 사영기하학과 복비의 정의

(그림 3)에서 직선 L를 보이는 대로 평면위에 그린다는 것은, 그리려는 도형과 보고 있는 눈 O 사이에 투명한 평면  $\pi$ 를 놓아 O에서 L를 투시하여 M을 만드는 것이라 할 수 있는데, L에서 M을 구하는 것을 L을  $\pi$ 에 사영한다고 한다. 사영에 의하여 만들어진 M은 L이 가지고 있는 모든 성질을 다 가지고 있지 않으나 어떤 성질은 보존되는데, 이와 같이 사영에 의해 불변인 성질을 연구하는 것이 사영기하학이다.

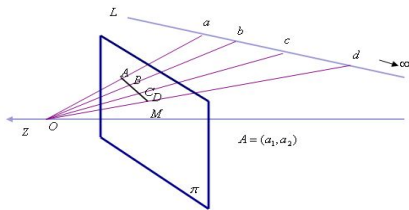


그림 3. 사영기하학의 성질  
Fig 3. Characteristic of Projective Geometry

사영직선  $M$  위의 4점  $A=[a_1, a_2], B=[b_1, b_2], C=[c_1, c_2], D=[d_1, d_2]$ 에 대해 식(3.1)를 사영기하학에서의  $a, b, c, d$ 의 복비(cross ratio)라고 한다.

$$(ab/cd) = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix}}{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} \begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix}} = \frac{(a_1c_2 - a_2c_1)(b_1d_2 - b_2d_1)}{(a_1d_2 - a_2d_1)(b_1c_2 - b_2c_1)} \dots (3.1)$$

실험에서는 특징 추출을 위해 동일 평면상의 5개 점의 복비를 사용한다. 사영기하학의 정리에 의하면 2차원 공간에서 동차좌표계내의 5개의 점  $c_i, i=1,2,3,4,5$ 가 있을 때, 이 점들이 동일 평면상에 존재하고, 이 중 3개의 점이 동일선상에 없으면, 식(3.2)와 같은 두개의 독립적인 사영 불변(projective invariant)값인 복비  $\lambda_1, \lambda_2$ 가 존재한다[4].

$$\lambda_1 = \frac{det(m_{431})det(m_{521})}{det(m_{421})det(m_{531})} \dots (3.2)$$

$$\lambda_2 = \frac{det(m_{432})det(m_{512})}{det(m_{412})det(m_{532})}$$

위 식에서  $det(m_{431})$ 은 점  $c_1, c_3, c_4$ 의  $x, y$  좌표 값을 가지고 계산되는 행렬 값을 의미한다. 그러나 복비  $\lambda_1, \lambda_2$ 는 변화에 민감하므로 점들의 치환에 강건한 안정적인  $J$ 벡터를 사용한다. 2차원 불변  $J$ 벡터의 각 요소 (3.4)는 식(3.3)에 의해 계산되어지는데, 이를 사영과 점치환 불변 벡터( $PPIV$  : Projective and Point Permutation Invariant Vector)라 부른다.

$$J[\lambda] = \frac{2\lambda^6 - 6\lambda^5 + 9\lambda^4 - 8\lambda^3 + 9\lambda^2 - 6\lambda + 2}{\lambda^6 - 3\lambda^5 + 3\lambda^4 - \lambda^3 + 3\lambda^2 - 3\lambda + 1} \dots (3.3)$$

$$J^{(1)} = J[\lambda_1], J^{(2)} = J[\lambda_2], J^{(3)} = J\left[\frac{\lambda_1}{\lambda_2}\right] \dots (3.4)$$

$$J^{(4)} = J\left[\frac{\lambda_2 - 1}{\lambda_1 - 1}\right], J^{(5)} = J\left[\frac{\lambda_1(\lambda_2 - 1)}{\lambda_2(\lambda_1 - 1)}\right]$$

5개 각각의 점  $c_i$ 는 구해진 2차원 불변벡터  $J^{(i)}$  각각과 1:1 매핑 관계에 있다. 논문에서는  $J$ 벡터를 계산하고 이를 크기 순(ascending order)으로 정렬하여 정합단계에서 이용한다.

### 3.2 두드러진 영역구성과 코너점 추출

논문에서는 영상내의 두드러진 영역의 값이 평활한 영역의 값보다 더 큰 값을 보유하게 되고, 이 영역에 특징이 존재한다는 이론[5]에 근거하여 식(3.5)를 이용하여 saliency map을 구성하였다.

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}} \dots (3.5)$$

위 식에서  $\sigma$ 는 윈도우의 표준편차를 나타내고,  $N$ 은 윈도우의 총 픽셀수,  $\mu$ 는 전체 영상의 명암 값의 평균,  $x_i$ 는 윈도우내 각 픽셀의 명암 값을 나타낸다. 두드러진 영역을 찾는 후에는 이 영역 안에서 KLT-CD(Kanade-Lucas-Tomasi Corner Detector) 알고리즘[6]을 사용하여 코너점을 추출하고 추출된 코너 점으로 이루어지는 5개의 순열 조합을 구성하는데, 순열 조합은 블록 내 코너점 중에서 5개를 뽑되, 중복을 허용하지 않고 순서대로 배열하는 경우의 수의 집합이다. 이렇게 구한 순열조합에 대해 위에서 정의한 사영과 점 치환불변 벡터인  $PPIV$ 를 계산한다.

### 3.3 특징모델 후보 생성단계

$PPIV$ 가 사영변환(projective transformation)하에서도 보존된다면, 첫째 동일선상에 세 개의 점이 존재하지 않아야 하고, 둘째 동일평면상의 점들로 구성되어야 한다. 동일선상에 세 개 점이 놓이게 되면, 세 점의 좌표로 구한 행렬값

(determinant)이 0가 되어 복비가 0이나 ∞의 값이 되기 때문이며, 동일평면상에 존재해야 하는 이유는 복비, 결과적으로 PPIV의 값이 불변하기 위한 제약사항이 된다. 실험에서는 동일선상의 3개점인지를 검사하기 위해 그람 매트릭스 [7]를 사용하였고, 동일평면상의 점인지를 검사하기 위해서 벡터 값이 2.0~2.8사이의 값인지를 검사하였다. 또한 검사를 거친 5개 점 집합을 특징모델이라 지칭하였다.

### 3.4 강건 특징모델 구성

추출된 특징모델은 전체 특징모델 후보 집합을 형성하게 되는데, 인식단계의 강건한 정합을 위해 추출된 특징 중 가장 바람직한 특징 모델들을 벡터 오더링(vector ordering) 기법[8]을 사용하여 추려낸다. 먼저 평균 불변 벡터값을 식(3.6)에 의해 구하고, 각 특징모델의 불변 벡터 값과 평균 불변 벡터값의 차이를 식(3.7)과 같이 계산한 후 정해진 임계치를 기준으로 바람직한 특징모델을 선택한다.

$$PPIV_{mean} = (1/q) \sum_{i=0}^{q-1} PPIV_i \dots\dots\dots (3.6)$$

$$d_i = \sqrt{\sum_{j=0}^4 (PPIV_i^j - PPIV_{mean}^j)^2} \dots\dots\dots (3.7)$$

## IV. 정합

### 4.1 유사도함수를 이용한 대응블럭 추출과 PPIV 불변검사

오정합율을 줄이고 일치되는 특징모델을 추출하기 위해 참조영상과 입력영상을 일정한 블록으로 나누고 유사도가 큰 두 블럭을 정합단계를 수행[9][10], 계산 속도를 향상시켰다. 그리고 대응블럭내의 특징모델이 입력영상의 하나의 특징모델과 일치하는지는 식(4.1)과 같은 임계치( $t_{PPIV} = 0.08$ )를 적용하여 비교하였다.

$$PPIV - PPIV' \leq t_{PPIV} \dots\dots\dots (4.1)$$

### 4.2 블록면검사와 특징모델 후보의 대응성확인

PPIV는 수치적으로 불안정한 요소가 포함되어 있어 이 수치만으로 정합을 수행하고 대응성을 결정할 경우 잘못된 정합 결과를 가져올 수 있다. 즉, 특징모델의 정합 후보 집합은 완전한 정합 쌍이 아니고 잡음 등으로 이상치(outlier)가 포함될 수 있으므로, 실험에서는 블록면 검사(convex hull test)를 통해 잘못된 정합 후보들을 걸러내고 대응점들을 수정하였다. 이것은 블록면상의 점들의 수는 보존되며, 블록면 상에 존재하는 이웃점들과의 관계 역시 보존된다는 이론에 근거한다. 블록면 검사는 Graham 탐색 알고리즘을 사용하였다. 그리고 블록면 검사를 통해 정확히 대응성이 획득된 4, 5개 점 특징모델의 변환 행렬을 사용하여 후보 집합에 대한 검증을 실시하였다[11]. 즉, 입력영상의 5개의 점과 정합된 참조영상의 5개의 점이 이러한 변환 행렬을 계산하는데 사용된다. 변환 행렬은 8개의 방정식이 이루는 선형시스템의 해(solution)로부터 값이 구해진다. 5개중 4개의 점이 사영기저(projective basis)를 구성하고 사영기저는 사영 변환식을 계산하는데 사용되어진다. 예를 들어 블록면 검사를 통과한 참조영상의 5개중 4개의 점 집합을  $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), (X_4, Y_4)$ 이라 하고, 입력영상의 4개의 점 집합을  $(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)$ 이라 하면, 변환행렬  $t$ 는 식(4.2)와 같이 표시되어진다.

$$\begin{bmatrix} X_1 & Y_1 & 1 & 0 & 0 & 0 & -x_1 X_1 & -x_1 Y_1 \\ 0 & 0 & 0 & X_1 & Y_1 & 1 & -y_1 X_1 & -y_1 Y_1 \\ X_2 & Y_2 & 1 & 0 & 0 & 0 & -x_2 X_2 & -x_2 Y_2 \\ 0 & 0 & 0 & X_2 & Y_2 & 1 & -y_2 X_2 & -y_2 Y_2 \\ X_3 & Y_3 & 1 & 0 & 0 & 0 & -x_3 X_3 & -x_3 Y_3 \\ 0 & 0 & 0 & X_3 & Y_3 & 1 & -y_3 X_3 & -y_3 Y_3 \\ X_4 & Y_4 & 1 & 0 & 0 & 0 & -x_4 X_4 & -x_4 Y_4 \\ 0 & 0 & 0 & X_4 & Y_4 & 1 & -y_4 X_4 & -y_4 Y_4 \end{bmatrix} \begin{bmatrix} t_{11} \\ t_{12} \\ t_{13} \\ t_{21} \\ t_{22} \\ t_{23} \\ t_{31} \\ t_{32} \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ x_3 \\ y_3 \\ x_4 \\ y_4 \end{bmatrix} \dots\dots\dots (4.2)$$

식(4.2)는  $AX = B$ 의 형태를 지니고 있으며, 변환행렬  $X$ 인  $t_{11}, t_{12}, t_{13}, t_{21}, t_{22}, t_{23}, t_{31}, t_{32}$ 를 구하는 것은  $A$ 행렬의 역행렬을 구하는 문제와 같아지게 된다.

만일 원근투영하의 3차원 좌표와 대응되는 영상 좌표가  $(x, y)$  이면, 이동과 회전에 의한 영상좌표  $(x', y')$  는 식 (5.3)와 같은 8개의 변수로 표현될 수 있다.

### V. 카메라 동작인수 산출

$$x' = \frac{a_1x + a_2y + a_3}{a_7x + a_8y + 1}, y' = \frac{a_4x + a_5y + a_6}{a_7x + a_8y + 1} \dots\dots (5.3)$$

본 장에서는 정합단계에서 얻어진 특징모델 집합을 가지고 제안하는 정확성과 수렴도를 향상시킨 2단계 카메라 동작인수 측정방법에 대해 기술한다.

카메라 외부인수 측정을 위해 사용되는 최소화기법 중 LM 방법은 결과 값이 정확하게 목표 값에 수렴하지만, 수렴을 위해 아주 작은 스텝크기(step size)로 최소화를 진행하므로 지역 최소점(local minima)에 빠질 위험이 있고, 이 경우 소요 시간이 길게 된다. 따라서 본 논문에서는 정확도와 수렴시간을 고려해서 초기 해를 호모그래피를 이용하여 일차적으로 구하고 그 해를 다시 LM에 입력해서 수렴도와 정확도를 동시에 향상시킨 최종적인 해를 구하고자 한다.

또한 카메라의 회전이 작고 초점거리가 일정하면서 크다면 대응되는 영상 좌표사이의 관계는 식(5.4)와 같은 이파인 모델로 표현[12]할 수 있다.

$$x' = x + \gamma y - \beta f = a_1x + a_2y + a_3 \dots\dots\dots (5.4)$$

$$y' = -\gamma x + y + \alpha f = a_4x + a_5y + a_6$$

식(5.4)의 이파인 인수로부터 카메라의 회전정보인 초점거리  $f$ , 상하각(tilt angle)  $\alpha$ , 좌우각(pan angle)  $\beta$ , 경사각(swing angle)  $\gamma$ 가 유도될 수 있다.

#### 5.1 카메라 모델과 좌표변환

실세계 좌표계에서의 한 점의 좌표를  $P_w(x_w, y_w, z_w)$  이라 하고, 그 점을 카메라 좌표계로 표시할 때의 좌표를  $P_c(x_c, y_c, z_c)$  이라 하면, 실세계좌표와 카메라 좌표계 사이의 관계는 식(5.1)과 같이 나타낼 수 있다.

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T, T = \begin{bmatrix} T_x & T_y & T_z \end{bmatrix}^T \dots\dots (5.1)$$

여기서  $R$ 은 실세계 기준 좌표계에 대한 카메라 좌표계의 회전을 나타내는 식(5.2)와 같은 회전 행렬이고  $T$ 는 기준 좌표계의 원점에 대한 카메라 좌표계로의 이동 벡터이다.

#### 5.2 2D 호모그래피를 이용한 초기값 산출

원근 투영하의 3차원 공간에서 같은 평면에 존재하는 점들에 대한 2차원 점들에 대한 이동은 3x3의 호모그래피 행렬로 표현되기 때문에 영상을 통한 카메라 동작은 투영된 2차원 호모그래피를 찾는 문제로 볼수 있다[13]. 본 논문에서는 정합된 특징모델 중 4쌍의 점들에 대한 대응관계로부터 직접 선형 변환인 DLT(Direct Linear Transformation) 알고리즘에 의해 호모그래피 행렬을 계산한다. (그림 4)와 같이 첫 번째 영상에 대한 동차좌표를  $X = (X, Y, W)^T$ , 두 번째 영상에 대한 동차좌표를  $x = (x, y, w)^T$  라 하고, 두 좌표간의 대응관계를  $X \leftrightarrow x$  라 할때, 식(5.5)의 관계식을 만족하면 두 영상 사이에는 호모그래피 행렬  $M$ 이 존재한다.

$$R = R_\alpha R_\beta R_\gamma \dots\dots\dots (5.2)$$

$$= \begin{bmatrix} r_0 & r_1 & r_2 \\ r_3 & r_4 & r_5 \\ r_6 & r_7 & r_8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

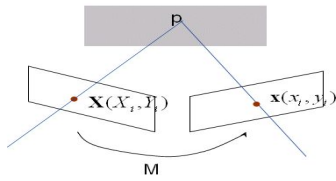


그림 4. 2D 호모그래피  
Fig 4. 2D Homography

$$X = Mx, \quad M = [r_1 \ r_2 \ t] \dots\dots\dots (5.5)$$

내부 파라미터를 알고 있다고 가정하면, 처음 두개의 열 벡터  $r_1$  과  $r_2$  를 얻어낼 수 있으므로 식(5.6)과 같은 카메라의 회전행렬  $R$  의 직교성을 이용하여 나머지 열벡터  $r_3$  는 외적  $r_1 \times r_2$  로부터 얻어서 투영행렬  $P$  를 완성한다.

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \quad P = [r_1 \ r_2 \ r_1 \times r_2 \ t] \dots\dots\dots (5.6)$$

호모그래피로부터 회전행렬  $R$  이 계산되면, 각 축에 대한 회전 변환 각도를 추정[14]할 수 있고, 앞서의 식(5.2)와 식(5.6)은 식(5.7)과 같은 관계를 가지므로, 최종적으로 카메라의 회전정보  $\alpha, \beta, \gamma$  는 식(5.8)과 같이 복원한다.

$$\frac{r_{12}}{r_{11}} = \frac{\sin\alpha}{\cos\alpha} = \tan\alpha, \quad \frac{r_{23}}{r_{33}} = \frac{\sin\gamma}{\cos\gamma} = \tan\gamma, \quad r_{13} = -\sin\beta \quad (5.7)$$

$$\alpha = \text{atan}\left(\frac{r_{12}}{r_{11}}\right), \quad \beta = \text{asin}(-r_{13}), \quad \gamma = \text{atan}\left(\frac{r_{23}}{r_{33}}\right) \quad (5.8)$$

그러나 식(5.8)으로부터 구해진  $\alpha, \beta, \gamma$  는 주기 함수인  $\sin$  과  $\tan$  함수로부터 유도되기 때문에 비슷한 값에 대해서도 다른 각도가 나오는 모호성을 지니게 된다. 따라서 논문에서는 이 해를 다시 LM에 입력해서 수렴도와 정확도를 향상시킨 최종적인 해를 구하고자 한다.

5.3 LM 방법을 이용한 최소화기법

LM 방법은 비선형 최소 제곱 추정을 이용한 최소화 기법의 하나로 최소화 기법의 반복시 목적 함수가 지역적인 2차원 함수로 올바르게 근사될 경우에는 가우스 뉴턴법을 적용하고, 목적함수가 2차원 함수로 올바르게 근사되지 못할 경우에는 기울기 강하 방법을 적용한다. 식 (5.9)에서 인수  $a$  는 최소화 기법을 통해 최종적으로 구하고자 하는 모델의 인수를 말한다. LM 알고리즘은 평균 제곱 오차(mean square error)를 정의함으로써 출력과 실제 출력과의 근접성을 평가하므로 식(5.9)와 같은 모델을 기반으로 구하고자 하는 문제에 적합한 목적함수를  $\chi^2(a)$  라 하면 목적함수는 식(5.10)과 같이 설정될 수 있다.

$$y = y(x, y; a) \dots\dots\dots (5.9)$$

$$\chi^2(a) = \sum_{i=1}^N w_i \left[ \frac{y' - y(x_i, y_i, a)}{\sigma_i} \right]^2 = \sum_{i=1}^N w_i \|r_i\|^2 \dots\dots (5.10)$$

식(5.10)에서  $y'_i$  는  $i$  번째 입력 자료를 의미하고,  $\sigma_i$  는  $i$  번째 자료의 분산,  $w_i$  는  $i$  번째 입력 자료의 가중치로서 0과 1의 값을 가지는 이진 가중치(binary weight)이다. 식 (5.10)의 목적 함수를 테일러 급수(Taylor series)를 이용하여 모델의 인수  $a$  에 대한 2차 방정식의 형태로 근사시키면 목적 함수는 식(5.11)과 같이 나타낼 수 있다.

$$\chi^2(a) \approx r - d \cdot a + \frac{1}{2} a \cdot D \cdot a \dots\dots\dots (5.11)$$

식(5.11)에서  $d$  는 목적 함수에 대한 인수  $a$  의 1차 미분 행렬을 의미하고,  $D$  는 목적 함수에 대한 인수  $a$  의 2차 미분 행렬인 헤시안 행렬을 의미한다. 식(5.10)의 인수  $a$  는 식 (5.12)과 같이 역 헤시안 방법을 통해 계산할 수 있다. 반대로, 식(5.10)이 2차 방정식의 형태로 근사 할 수 없다면 식 (5.13)와 같이 기울기 강하 방법을 통해 해결한다.

$$a_{min} = a_{cur} + D^{-1} \cdot [-\nabla \chi^2(a_{cur})] \dots\dots\dots (5.12)$$

$$a_{next} = a_{cur} - \text{constant} \times \nabla \chi^2(a_{cur}) \dots\dots\dots (5.13)$$

LM 방법은 해를 구하기 위해 식(5.12)과 식(5.13)을 결합한 식 (5.14)와 같은 선형 방정식을 정의한다.

$$\sum_{l=1}^M \alpha'_{kl} \delta a_l = \beta_k \dots\dots\dots (5.14)$$

$$\alpha'_{kl} = \begin{cases} \alpha_{kl}(1 + \lambda) & \text{if } (k = j) \\ \alpha_{kl} & \text{if } (k \neq j) \end{cases}$$

식(5.14)에서  $\alpha$  는  $0.5 \times \mathbf{D}$  를 의미하며,  $\beta$  는  $-0.5 \times \mathbf{d}$  를 의미한다. 또한  $\delta a_l$  는 인수  $a$  의 변화량을 의미한다. 식(5.14)는 목적 함수가 인수에 대한 2차 방정식의 형태로 올바르게 근사될 경우에는 식(5.14)의 선형방정식에서  $\lambda$  를 감소한 후 역 헤시안 방법을 이용하여 해를 구하고, 근사되지 못할 때에는  $\lambda$  를 증가한 후 기울기 강하 방법을 이용하여 해를 구하게 된다. 실험에 사용된 측정모델은 식(5.4)를 이용한 식(5.15)과 같은 아파인(affine) 모델이다.

$$y(x, y, \mathbf{a}) = \begin{bmatrix} u(x, y) \\ v(x, y) \end{bmatrix} = \begin{bmatrix} a_1x + a_2y + a_3 \\ a_4x + a_5y + a_6 \end{bmatrix} \dots\dots\dots (5.15)$$

위 식에서  $u(x, y)$  와  $v(x, y)$  는 카메라의 회전에 따른 특징모델의 이동 성분(수평 성분과 수직 성분)을 의미한다. 식(5.15)을 기반으로 한 목적 함수는 식 (5.16)과 같다.

$$\chi^2(\mathbf{a}) = \sum_{i=1}^N w_i \left[ \frac{y'_i - y(x_i, y_i, \mathbf{a})}{\sigma_i} \right]^2 + \sum_{i=1}^N w_i \left[ \frac{(x'_i - \hat{x}_i(x_i, y_i, \mathbf{a}))^2}{\sigma_{x_i}^2} + \frac{(y'_i - \hat{y}_i(x_i, y_i, \mathbf{a}))^2}{\sigma_{y_i}^2} \right] + \sum_{i=1}^N w_i \|r_i\| \dots\dots\dots (5.16)$$

위 식에서  $w_i$  는  $i$  번째 입력 자료의 가중치를,  $x'_i$  와  $y'_i$

는  $i$  번째 입력 자료의  $x, y$  성분을 의미하며,  $\hat{x}_i(x_i, y_i, \mathbf{a})$  와  $\hat{y}_i(x_i, y_i, \mathbf{a})$  는 모델의  $x, y$  성분을 의미한다.

### VI. 실험결과

실험을 위해 사용된 컴퓨터는 IBM사의 인텔 2.6GHz, 256 램을 장착한 컴퓨터이며, 운영체제로는 MS사의 윈도우 XP를, 컴파일러는 MS사의 비주얼C++를 사용하였다. 영상은 실내에서 Nikon사의 Coolpix 3200로 촬영한 후 640×480 화소의 크기로 정규화하였다.

(그림 5)는 학습을 통해 선택된 강건 특징모델 3쌍의 예이다. 영상에서 총 25개의 코너점을 추출하였고, 이중 블록 유사도 검사를 거쳐 5184개의 특징모델 후보 집합이 추출되었다. 그리고 3개의 점이 동일선상에 놓인 경우를 제외하고 동일평면 검사를 거쳐 101개의 후보가 남게 되었으며, 최종적으로 블록면 검사와 대응성 검사를 거친 25개의 코너 점 중 벡터오더링 방법을 통해 3개의 특징모델을 선택하였다.

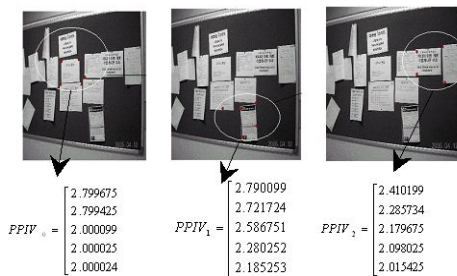


그림 5. 학습을 통해 선택된 강건 특징모델 3쌍  
Fig 5. Selected Three Robust Feature Models by Learning

(그림 6)은 뷰가 다른 입력영상을 사용하여 3개의 강건 특징모델이 인식된 예와 계산된 PPIV값이다.

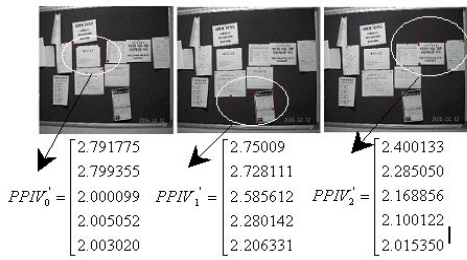


그림 6. 인식된 특징모  
Fig 6. Recognized Feature Models

(그림 7)은 카메라 동작인수 산출을 위해 사용된 영상으로 (a)는 참조영상에 대해 Y축을 기준으로 오른쪽으로 2도 회전(pan angle)한 영상이고 (b)는 (a)에서 특징을 추출한 영상이다. 3쌍의 특징모델이 정합단계에서 추출되어, <표 1>의 총 15개의 정합된 좌표 값을 가지고 <표 2>의 호모그래피 행렬을 산출하였다. <표 3>은 호모그래피로 구한 값을 초기치로 사용하여 LM 방법으로 산출된 카메라 동작인수의 추출결과이다.



그림 7. 입력영상(a)과 특징이 추출된 영상(b)  
Fig 7. Input Image(a) and Extracted Features in Input Image(b)

표 1. 참조영상과 입력영상의 대응점 집합  
Table 1. Correspondence point set of Reference and Input Image

정합된 특징모델 수	참조영상과 입력영상의 대응점 집합									
	참조영상의 좌표(x,y)			입력영상의 좌표(x,y)						
1	143	227	134	239	253	141	229	138	242	255
	115	114	172	173	156	116	117	175	179	158
2	188	273	332	280	342	190	276	335	282	345
	237	235	227	308	305	239	240	229	312	308
3	288	328	424	328	434	290	340	440	330	438
	105	103	100	156	176	110	102	98	157	178

표 2. 2D 호모그래피 행렬을 이용해 산출된 카메라 동작인수 결과  
Table 2. Extracted Camera Motion Parameter using 2D Homography Matrix

(a) 호모그래피 행렬결과										(b) 추출된 카메라회전인수				
2D 호모그래피 행렬										Method				
r1	r2	r4	r1	r2	r4	r1	r2	r4	r2	Method	$\alpha$ (tilt)	$\beta$ (pan)	$\gamma$ (swing)	
0.914	0.006	0.595	-0.021	0.948	0.066	-0.000	0.000			이상적	값	0.000	+2.000	0.000
										이상적	오차	0.000	0.000	0.000
										2D 호모그래피	값	0.007	1.006	1.571
											오차	0.007	0.994	1.571

표 3. 호모그래피와 LM을 이용하여 추출된 카메라 동작인수  
Table 3. Extracted Camera Motion Parameter by Homography and LM Method

Method		$\alpha$ (tilt)	$\beta$ (pan)	$\gamma$ (swing)
이상적	값	0.000	+2.000	0.000
	오차	0.000	0.000	0.000
LM	값	-0.354	1.9780	0.000
	오차	-0.354	-0.0220	0.000

(그림 8)은 10개의 영상을 이용하여 호모그래피의 값을 초기치로 사용한 경우와 랜덤한 값을 사용한 경우에 LM의 속도를 그래프로 나타낸 것이다. 속도가 1/2 정도로 줄어들었음을 알 수 있다. 따라서 동작인수를 산출하는데 있어, 제안한 2단계 동작인수 산출 방법이 LM이 갖는 수렴하는데 오래 걸리는 단점을 호모그래피의 값을 초기치로 사용함으로써 해소하고 호모그래피로 구한 값의 부정확성을 LM 방법을 이용하여 보완할 수 있음을 알 수 있다.

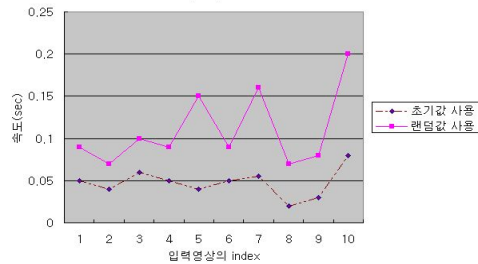


그림 8. 속도비교  
Fig8. Time Comparision

## Ⅶ. 결론 및 향후 과제

본 논문에서는 효율적인 기하학의 불변 특징벡터를 사용하여 카메라의 뷰포인트에 무관한 강건한 특징을 추출하고 이를 이용하여 카메라의 동작인수를 산출하는 방법을 제안하였다. 본 논문에서 제안하는 특징추출은 실내영상을 대상으로 하였으며, 영상을 분석하기 위해 블럭 단위 연산을 수행하였다. 카메라 동작인수 산출 단계에서 제안한 방법은 기존 방법인 LM 알고리즘의 단점을 보완하여 수렴도와 정확도를 향상시켜 최적의 카메라 동작인수를 추출할 수 있게 해준다. 실험영상은 실내 영상을 대상으로 실험하였으므로 실외영상을 이용할 경우에는 특징추출에 어려움이 있을 수 있다. 따라서 실외의 잡음에 강한 특징추출 연구가 지속될 필요가 있고, 또한 3차원 정보를 이용한 실험도 계속되어야 하겠다.

## 참고문헌

- [1] Jeffrey Hightower and Gaetano Borriello, "A Survey and Taxonomy of Location Systems for Ubiquitous Computing," Technical Report, Computer Science and Engineering, University of Washington, Aug. 2001.
- [2] K.Sugihara, "Some Location Problems for Robot navigation Using a Single Camera," Computer Vision, Graphics and Image Processing 42, pp. 112-129, 1988.
- [3] O. Faugeras and G. Toscani, "The Calibration Problem for Stereo," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 15-20, Miami Beach, FL, June 1986.
- [4] R.Lenz, P.Meier, "Efficient Invariant Representation," International Journal of Computer Vision 26(2) p. 137-152, 1998.
- [5] Panos E. Trahanias, Savvas Velissaris and Thodoris Garavelos, "Visual Landmark Extraction and Recognition for Autonomous Robot Navigation," Proc. IROS 97, pp. 1036-1042, 1997.
- [6] Carlo Tomasi and Takeo Kanade, "Detection and Tracking of Point Features," Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.
- [7] K. Kanatani, "Computational Projective Geometry," CVGIP:Image Understanding Workshop, Washington, DC, pp. 745-753, 1993.
- [8] V. Barnett, "The Ordering of Multivariate Data," Journal of Royal Statistical Society A, Part 3 139 pp. 318-343, 1976.
- [9] 김성곤. "전역탐색 알고리즘을 이용한 이동 무선통신 네트워크의 최적화에 대한 연구," 한국컴퓨터정보학회 논문지 제9권 제1호, 2004.3.
- [10] 지정규, 김용근, "가변탐색창을 이용한 실시간 객체추적 시스템," 한국OA학회 논문지 제7권 제4호, 2002.12.
- [11] Fishler, M.A and Bolles, R.C., "Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography," Commumination ACM, vol. 24, no. 6, pp. 381-395, 1981.
- [12] 장석우, "카메라의 동작을 보정한 장면 전환 검출," 숭실대학교 박사학위 청구논문, 2000.
- [13] Hartley, R. I, Zisserman, A, "Multiple View Geometry in Computer Vision," Cambridge University Press, 2000.
- [14] Michael A. Penna, "Determining Camera Parameters From The Perspective Projection Of A Quadrilateral," Pattern Recognition, Vol. 24. No. 6, pp. 553-541, 1991.

저 자 소개



차 정 희

1983년 숭실대 전자계산학과 졸업  
(학사)  
1983~1998 신용보증기금 전산실  
근무  
1999~2001년 동국대 대학원 컴퓨  
터공학과 졸업(공학석사)  
2005. 8월 숭실대 대학원 컴퓨터공  
학과 졸업(공학박사)  
<관심분야> 로봇비전, 영상처리, 지문  
인식, 가상현실, 3D모델링 등



이 근 수

1983년 숭실대 전자계산학과 졸업  
(학사)  
1988년 숭실대 대학원 컴퓨터공학  
과 졸업(공학석사)  
1993년 숭실대 대학원 컴퓨터공학  
과 졸업(공학박사)  
1992~1993년 군산대학교 컴퓨터  
과학과 교환교수  
2003~2004년 미국 George mason  
University 전자계산학과 객  
원 교수  
1989~현재 한경대학교 컴퓨터 공학  
과 교수  
<관심분야> 패턴인식, 퍼지이론, 로봇비  
전, 지식기반시스템, 동작이  
해, 비디오검색등

