

정규논문 (Regular Paper)

방송공학회논문지 제17권 제4호, 2012년 7월 (JBE Vol. 17, No. 4, July 2012)

http://dx.doi.org/10.5909/JBE.2012.17.4.634

## 다수의 영상 특징점 정합을 위한 비선형 최적화 기법

안 양 근<sup>a)</sup>, 홍 지 만<sup>a)†</sup>

### Nonlinear Optimization Method for Multiple Image Registration

Yangkeun Ahn<sup>a)</sup> and Jiman Hong<sup>a)†</sup>

#### 요 약

본 논문에서는 다수의 영상에서 발견된 특징점의 정확한 정합을 위한 비선형 최적화 기법을 제안한다. 영상에서 발견된 특징점은 선형 해법에 의해 다수의 영상간의 변환을 구할 수 있지만 큰 오차를 수반하게 된다. 이는 영상이 생성되는 모델이 비선형이며, 다수 시점간의 운동역시 비선형의 형태를 띠기 때문이다. 하지만 다수의 영상의 비선형 최적화는 일반적인 비선형 해법을 도입하였을 때에는 복잡도가 지수적으로 증가하는 단점이 있다. 본 논문에서는 Levenberg-Marquardt 비선형 최적화 방법의 희박해법(Sparse solution)을 이용하여 다수의 특징점간의 변환을 구하는 방법을 보인다.

#### Abstract

In this paper, we propose nonlinear optimization method for feature matching from multiple view image. Typical solution of feature matching is by solving linear equation. However this solution has large error due to nonlinearity of image formation model. If typical nonlinear optimization method is used, complexity grows exponentially over the number of features. To make complexity lower, we use sparse Levenberg-Marquardt nonlinear optimization for matching of features over multiple view image.

Keyword : Image registration, Nonlinear optimization

#### 1. 서 론

다수의 영상간의 변환을 구하는 것은 카메라의 교정 및 외부 파라미터 추정에 반드시 필요한 과정이다<sup>[2-7]</sup>. 두 장의 영상간의 변환을 구하는 것은 homography 변환을 구하는 것으로 Direct Linear Transform (DLT) 방법이 제안되어 왔다<sup>[1]</sup>. 이 방법은 4개 이상의 특징 점 대응을 통하여 2차 변

환 행렬 H를 추정하는 것이다.

하지만 선형 해법을 통하여 계산된 해는 오차를 수반하게 된다. 이는 영상의 생성 모델이 (1.1)과 같이 근원적으로 비선형이기 때문이다. (1.1)에서  $u, v$ 는 영상좌표를  $X, Y, Z$ 는 3차원 좌표를 의미한다.

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} R & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1.1)$$

(1.1)에서 각각의 변수는 (1.2)와 같다.

a) 송실대학교 컴퓨터학과 (School of Computing, Soongsil University)

† 교신저자 : 홍지만 (Jiman Hong)

E-mail: jiman@ssu.ac.kr

Tel: +82-2-828-7168 Fax: +82-2-820-0920

· 접수일(2012년3월21일), 수정일(2012년6월7일), 게재확정일(2012년7월3일)

$$K = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}, t = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (1.2)$$

(1.2)에서  $K$ 는 내부 파라미터(intrinsic parameter),  $R, t$ 는 각각 회전과 이동량을 기술하는 외부 파라미터(extrinsic parameter)이다.

따라서 선형 해를 초기해로 설정하고 (1.1)의 영상모델을 이용한 비용함수 최소화를 통한 비선형 최적해법이 이용되어 왔다. 하지만 이러한 비선형 최적화는 시점의 개수  $m$ 과 특징점의 개수  $n$ 에 대한 복잡도가  $O((m+n)^3)$ 에 비해 하는 단점이 있다<sup>[1]</sup>.

하지만 하나의 영상에서 존재하는 특징 점은 입력된 모든 영상 집합에 존재하는 것이 아니라 일부에서만 관찰되게 된다. 그림 1에서 보면, 특징점 F2는 시점 V1에서 관찰되지 않는다. 이는 특징 점 F2와 시점 V1은 아무런 연관이 없음을 의미한다.

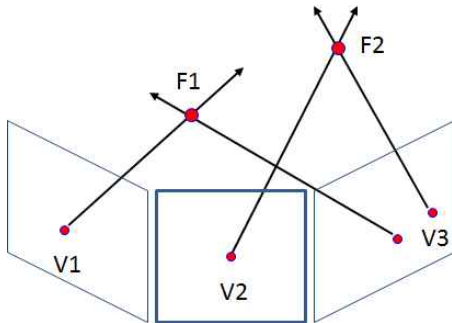


그림 1. 영상의 특징점간의 관계는 관찰 시점에 따라 유한다  
 Fig 1. Each feature has different view point

본 논문에서는 이러한 특성을 이용하여 잘 알려진 최적화 방법 Levenberg-Marquardt(LM) 방법의 희박해법(Sparse solution)을 이용하여 다수 시점의 최적화의 계산량을 줄이는 방법을 보인다.

## II. 두 시점 영상의 관계

평면위의 특징점이 두 시점에서 관찰되면 각 특징점 사

이에는 (2.1)의 관계가 성립한다.

$$x' = \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = Hx \quad (2.1)$$

(2.1)에서 행렬  $H$ 는 2차원 영상 좌표에서 기술된 특징점  $x', x$  간의 변환을 나타낸다. 행렬  $H$ 를 구하기 위하여 (2.2)의 관계를 이용한다.

$$x' \times Hx = 0 \quad (2.2)$$

(2.2)에서 각 요소는 식 (2.3)과 같다.

$$x' = \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix}, Hx = \begin{bmatrix} h^{1T}x \\ h^{2T}x \\ h^{3T}x \end{bmatrix} \quad (2.3)$$

변환행렬  $H$ 에 의한  $x$ 의 변환이  $x'$ 과 같다는 관계로부터 자기 자신과의 외적이 0임을 이용하면 (2.4)를 설정 할 수 있다.

$$x' \times Hx = \begin{bmatrix} v' h^{3T}x - h^{2T}x \\ h^{1T}x - v' h^{3T}x \\ u' h^{2T}x - v' h^{1T}x \end{bmatrix} = 0 \quad (2.4)$$

(2.4)로부터 (2.5)의 선형 시스템을 설정 할 수 있다.

$$\begin{bmatrix} 0 & -x^T & v' x^T \\ x^T & 0 & -u' x^T \\ -v' x^T & u' x^T & 0 \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = Ah = 0 \quad (2.5)$$

여기서 행렬  $H$ 는 행렬  $A$ 의 특이치 분해(Singular Value Decomposition;SVD)로 구할 수 있다. 이 두 시점의 관계로부터 역시 다수의 시점역시 이 두 시점의 확장된 형태로 표현 할 수 있음을 알 수 있다. 예를 들어 그림 1의 3개 시점 V1,V2,V3는 각각 V1-V2의 변환, V2-V3의 변환 그리고 V1-V3의 변환으로 관계를 기술 할 수 있다.

### III. 비선형 최적화 방법

이 절에서는 기존의 비선형 최적화 방법에 대하여 소개하고 LM의 회박해법에 대해서도 소개한다.

#### 1. 일반적인 비선형 최적화 방법

(3.1)의 비선형 함수를 고려하자.

$$X = f(P) \tag{3.1}$$

여기서  $P$ 는 파라미터,  $X$ 는 측정치이다. 비선형 최적화는 파라미터 추정치  $\hat{P}$ 에 대하여 (3.2)의 오차  $\|\epsilon\|$ 을 최적화 하는 방향으로 구하는 것이다.

$$X = f(\hat{P}) - \epsilon \tag{3.2}$$

만약 현재 시점에서 추정된 파라미터  $P$ 를  $P_0$ 라 할 때, 파라미터에 미소변위  $\Delta$ 를 고려하면 (3.3)과 같은 관계를 설정 할 수 있다.

$$f(P_0 + \Delta) = f(P_0) + J\Delta, \quad J = \frac{\partial f}{\partial P} \tag{3.3}$$

여기에서  $J$ 는 함수  $f$ 의 Jacobian이다. 반복법에 의하여 다음 순차의 파라미터  $P_1$ 은 (3.4)의 관계에 의하여 구한다.

$$P_1 = P_0 + \Delta \tag{3.4}$$

$$f(P_1) - X = f(P_0) + J\Delta - X = \epsilon_0 + J\Delta \tag{3.5}$$

(3.5)에서 미소변위  $\Delta$ 는  $\|\epsilon_0 + J\Delta\|$ 를 최적화 하여야 한다. 여기에서  $\Delta$ 는 (3.6) 및 (3.7)의 식을 만족한다.

$$J^T J \Delta = - J^T \epsilon_0 \tag{3.6}$$

$$\Delta = - J^+ \epsilon_0 \tag{3.7}$$

여기서  $J^+$ 는 pseudo inverse이다. 이를 normal equation이라 한다<sup>[8]</sup>.

#### 2. LM 최적화 방법

LM 최적화 방법은 다음의 normal equation을 고려한다.

$$(J^T J + \lambda I) \Delta = - J^T \epsilon \tag{3.8}$$

여기서  $\lambda$ 는 해로 근접하는 속도를 조정하는 damping 항  $\lambda$ 로 LM 최적화가 빠른 속도로 해를 추정하는데 도움을 준다. 하지만  $J^T J$ 의 역행렬을 구하는 연산은 파라미터의 개수의 제곱에 비례하는 단점이 있다. 예를 들어 H행렬을 구하는 경우  $n$ 개의 점에 대하여  $2n + 9$ 개의 파라미터가 필요하는데, 점의 개수가 늘어날수록 연산량은 제곱으로 늘어나게 된다. 하지만 그림 1과 같은 관계를 이용하면 이를 블록 파라미터 형태로 표현하여 회박 LM 최적화 방법을 적용 할 수 있다.

#### 3. 회박 LM 최적화 방법

파라미터  $P$ 를 (3.9)와 같은 두 가지 파라미터 집합으로 구성되어 있다고 하자.

$$P = [a^T, b^T]^T \tag{3.9}$$

$a$ 는 재투영 오차(reprojection error)를 위한 점 집합이고,  $b$ 는 homography라고 하자. 이때 Jacobian역시 (3.10)과 같이 블록 행렬 형태로 표현 할 수 있다.

$$A = [\partial X / \partial a], \quad B = [\partial X / \partial b] \tag{3.10}$$

따라서 normal equation은 (3.11)의 형태로 바뀐다.

$$J\delta = [A|B] \begin{bmatrix} \delta_a \\ \delta_b \end{bmatrix} = \epsilon \tag{3.11}$$

이때 normal equation의 역행렬  $J^T \Sigma_X^{-1} J\delta = J^T \Sigma_X^{-1} \epsilon$

은 (3.12)와 같은 블록행렬로 표현 될 수 있다.

$$\begin{bmatrix} A^T \Sigma_X^{-1} A & A^T \Sigma_X^{-1} B \\ B^T \Sigma_X^{-1} A & B^T \Sigma_X^{-1} B \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_b \end{bmatrix} = \begin{bmatrix} A^T \Sigma_X^{-1} \epsilon \\ B^T \Sigma_X^{-1} \epsilon \end{bmatrix} \quad (3.12)$$

이를 (3.13)의 축약 형태로 표현하면

$$\begin{bmatrix} U^* & W \\ W^T & V^* \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_b \end{bmatrix} = \begin{bmatrix} \epsilon_a \\ \epsilon_b \end{bmatrix} \quad (3.13)$$

$\delta_a, \delta_b$ 는 Schur complement에 의하여 (3.14) 및 (3.15)와 같이 구할 수 있다.

$$(U^* - W V^{*-1} W^T) \delta_a = \epsilon_a - W V^{*-1} \epsilon_b \quad (3.14)$$

$$V^* \delta_b = \epsilon_b - W^T \delta_a \quad (3.15)$$

다음 단계에서의 파라미터 벡터  $P_1$ 은  $((a + \delta_a)^T, (b + \delta_b)^T)^T$ 로 갱신될 수 있다. 그림 1과 같은 관계를 가질 때, normal equation의 Jacobian은 (3.16)의 구조를 가지게 된다.

$$J = \begin{bmatrix} A_1 & B_1 & & \\ A_2 & & B_2 & \\ \vdots & & & \ddots \\ A_n & & & B_n \end{bmatrix} \begin{bmatrix} \delta_a \\ \delta_{b_1} \\ \vdots \\ \delta_{b_n} \end{bmatrix} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad (3.16)$$

이를 자세히 보면  $B = [\partial X / \partial b]$ 가 대각성분만 가지게 된다. 이는 각 특징점들이 서로 연관이 없기 때문에 가능한 것이다. 각 시점에 등장하는 특징점들을 잘 관리하여 매개 변수화 한다면 (3.16)과 같은 행렬 대각구조의 이점을 이용할 수 있다. 이 방법은 시점의 개수  $m$ 과 특징점의 개수  $n$ 에 대한 복잡도가  $O(m^3 + mn)$ 에 비례한다. 이 방법은 기존의 비선형 최적화법의 복잡도  $O((m+n)^3)$ 에 비하여 많은 계산량 감소가 존재한다.

#### 4. 희박 LM을 이용한 Homography 추정

측정벡터를  $X_i = (x_i^T, x_i'^T)^T$ 로 설정하고, 파라미터 벡터

$P = (h^T, \widehat{x}_{1,2,\dots,n}^T)$ 로 설정하자. 여기에서  $\widehat{x}_i^T$ 는 초기치  $H$ 를 이용한 첫 번째 시점의 이미지 좌표의 추정치이다. 즉  $\widehat{x}_i = H^{-1} x_i'$ 이다. 이때 각 파라미터에 대한 Jacobian은 다음과 같다.

$$A_i = \partial \widehat{X}_i / \partial h = \begin{bmatrix} 0 \\ \partial \widehat{x}_i' / \partial h \end{bmatrix} \quad (3.17)$$

$$B_i = \partial \widehat{X}_i / \partial \widehat{x}_i = \begin{bmatrix} I \\ \partial \widehat{x}_i' / \partial \widehat{x}_i \end{bmatrix} \quad (3.18)$$

(3.17)의 행렬요소를 이용하여 (3.16)의 normal equation의 해를 구한다.

## IV. 실험결과 및 결론

### 1. 실험 결과

본 논문에서 제안한 방법으로 서로 다른 시점에서 촬영된 두 영상을 정합해 보았다. 그림 2는 선형 해법을 통해 정합한 영상이고 그림 3은 이후 비선형 최적화를 통해 보정한 영상이다.

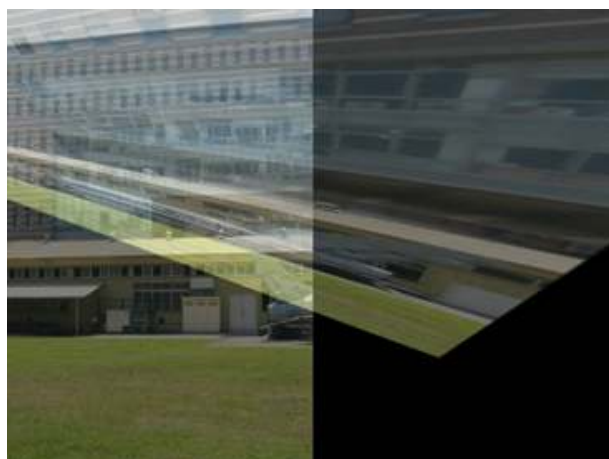


그림 2. 선형 해법으로 정합한 영상. 선형 해법은 큰 오차를 수반하게 된다.  
 Fig 2. Image stitched by linear solution. Linear solution has large error



그림 3. 제안된 비선형 최적화를 통하여 정합한 영상. 하나의 영상처럼 정합된 것을 볼 수 있다.  
Fig 3. Image stitched by proposed method

비선형 최적화를 거친 이후 그림에서 보는 것과 같이 한 장의 영상으로 정합된 것을 볼 수 있다. 그림 4와 그림 5의 비교를 통하여 비선형 해법은 선형 해법 이후에 오차를 줄이는 좋은 방법임을 알 수 있다.

## 2. 결론 및 향후 연구 계획

본 연구에서는 선형 추정을 이용하여 합성된 영상에서 발생하는 정렬 오차를 최소화 하기위한 비선형 최적화 방법을 제안 하였다. 제안된 방법은 기존의 비선형 방법에 비하여 복잡도가 월등히 감소된 방법으로 특징점의 개수가 늘어날수록 기존의 방법에 비하여 우위를 가진다. 따라서 SIFT<sup>[9]</sup>나 SURF<sup>[10]</sup> 같은 특징점을 많이 검출해내는 특징점



그림 4. 선형해법으로 정합한 영상. 어긋난 부분을 볼 수 있다.  
Fig 4. Image stitched by linear method



그림 5. 제안된 비선형 최적화를 통해 정합한 영상. 정확히 정렬된 사진을 볼 수 있다.  
Fig 5. Image stitched by nonlinear optimization.

알고리즘과 잘 동작한다. 제안된 방법은 실제 영상을 이용한 실험으로 검증 되었다. 선형 추정법에 비하여 이미지 정렬 오차가 매우 줄어든 것을 확인 할 수 있었다. 향후에는 이 방법을 **structure from motion**에 적용하여 다수의 영상에 대한 비선형 최적화를 진행해 볼 예정이다.

### 참 고 문 헌

- [1] R. Hartley and A. Zisserman, "Multiple-view Geometry in Computer Vision" Cambridge University Press, March 2004.
- [2] 김성용, 주명호, 강행봉, "파노라마 영상 생성을 위한 빠른 특징 매칭 방법", 대한전자공학회 학술대회
- [3] 이용현, 김윤중 "파노라마 사진을 위한 가중치 적용 정합 알고리즘에 관한 연구", 정보통신전문대학원 논문집, Vol.3 No.1 2005.
- [4] 이강호, "실린더 파노라마 영상의 계구성을 위한 단순화된 사영 변환", 한국컴퓨터정보학회, Vol.11 No.2 2006.
- [5] 신용득, 백문홍, 장가람, 박재한, "깊이 감지 카메라의 깊이 이미지를 이용한 3차원 점군 데이터 획득", 한국 정밀공학회 2011년도 춘계학술대회 논문집(상), 2011.6, page 3-825
- [6] 강윤석, 호요성, "다시점 카메라와 깊이 카메라를 이용한 3차원 장면의 깊이 정보 생성 방법", 전자공학회논문지 48권 3호. 2011.5 page 1-131.
- [7] 진광식, 윤태성, "스테레오 카메라 캘리브레이션 및 구간분할 Newton 보간법을 이용한 개선된 깊이 정보 추출", 대한전자공학회 학술대회 논문집 (신호처리합동) 제10권 1호, 1997.1, page(s): 3-1144
- [8] William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, "Numerical Recipes 3rd edition: The art of scientific computing", Cambridge University Press. 2007.
- [9] Lowe, D.G. "Object recognition from local scale-invariant features", Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, 1999.
- [10] Bay, H. and Tuytelaars, T. and Van Gool, L. "Surf: Speeded up robust features", Computer Vision –ECCV 2006.

### 저 자 소 개

#### 안 양 근



- 1994년 : 한국해양대학교 제어계측공학과 학사
- 1997년 : 한국해양대학교 제어계측공학과 석사
- 1997년 ~ 2000년 : 한국전력공사 전력연구원(KEPRI) 연구원
- 2000년 ~ 2005년 : (주)해동정보통신 선임연구원
- 2005년 ~ 현재 : 전자부품연구원(KETI) 책임연구원
- 주관심분야 : 임베디드시스템, 컴퓨터비전, HCI, UI/UX, 3D

#### 홍 지 만



- 1994년 : 고려대학교 컴퓨터공학과 학사
- 1997년 : 서울대학교 컴퓨터공학과 석사
- 2003년 : 서울대학교 컴퓨터공학과 박사
- 2000년 ~ 2003년 : (주)지멘텍 기술연구소장
- 2004년 ~ 2007년 : 광운대학교 컴퓨터공학과 조교수
- 2007년 ~ 현재 : 송실대학교 컴퓨터학과 교수
- 주관심분야 : 시스템소프트웨어, 임베디드시스템