

정규 상호정보와 기울기 방향 정보를 이용한 다중센서 영상 정합 알고리즘

주재용*, 김민재*, 구분화**, 고한석*

Multi-sensor Image Registration Using Normalized Mutual Information and Gradient Orientation

Jae-Yong Ju*, Min-Jae Kim*, Bon-Hwa Ku**, Han-Seok Ko*

요약

영상정합은 동일한 장면에 대해서 서로 다른 시점, 서로 다른 시간 혹은 서로 다른 특성의 센서로부터 얻은 영상들의 위치 관계를 대응시켜주는 기법이다. 본 논문에서는 가시광선 영상 및 적외선 영상과 같은 다중센서 영상을 정합하기 위한 방법을 제안한다. 영상정합은 두 영상에서 특징점을 추출하고, 특징점 간의 대응 관계를 구함으로써 이루어진다. 기존의 다중센서 영상 정합을 위한 방법으로 정규상호정보를 이용하여 대응 특징점을 선별하는 방법이 제안되었다. 정규상호정보 기반의 영상정합 기법은 두 영상의 통계적 상관성이 전역적이어야 한다는 가정을 전제한다. 그러나 가시광선 영상과 적외선 영상에서는 이를 보장하지 못하는 경우가 많아 대응 특징점의 정확도가 저하되기 때문에 기존의 방법은 안정적인 정합 성능을 기대하기 힘들다. 본 논문에서는 영상의 공간정보로서 기울기 방향 정보를 정규상호정보와 결합함으로써, 대응 특징점의 정확도를 향상시켰으며 이를 통해 정확성 및 안정적인 영상 정합 결과를 도모하였다. 다양한 실험 결과를 통해 제안하는 방법의 효용성을 증명하였다.

▶ Keyword : 영상정합, 다중센서 영상, 상호정보, 기울기 방향, 특징점

Abstract

Image registration is a process to establish the spatial correspondence between the images of same scene, which are acquired at different view points, at different times, or by different sensors.

• 제1저자 : 주재용 • 교신저자 : 고한석

• 투고일 : 2012. 02. 01, 심사일 : 2012. 03. 01, 게재확정일 : 2012. 04. 18.

* 고려대학교 전기전자전파공학부(Dept. of Electrical Engineering, Korea University)

** 고려대학교 영상정보처리학과(Dept. of Visual Information Processing, Korea University)

※ 본 연구는 서울시 산학연 협력사업(WR080951)의 연구결과로 수행되었습니다.

In this paper, we propose an effective registration method for images acquired by multi-sensors, such as EO (electro-optic) and IR (infrared) sensors. Image registration is achieved by extracting features and finding the correspondence between features in each input images. In the recent research, the multi-sensor image registration method that finds corresponding features by exploiting NMI (Normalized Mutual Information) was proposed. Conventional NMI-based image registration methods assume that the statistical correlation between two images should be global, however images from EO and IR sensors often cannot satisfy this assumption. Therefore the registration performance of conventional method may not be sufficient for some practical applications because of the low accuracy of corresponding feature points. The proposed method improves the accuracy of corresponding feature points by combining the gradient orientation as spatial information along with NMI attributes and provides more accurate and robust registration performance. Representative experimental results prove the effectiveness of the proposed method.

▶ Keyword : Image registration, Multi-sensor images, Mutual information, Gradient orientation, Feature

1. 서 론

최근에 감시 시스템과 의료 영상 등의 분야에서 변화 감지, 움직임 검출, 초해상도 영상 복원[1] 및 물체 인식과 추적 등의 알고리즘의 정확도와 신뢰도를 향상시키기 위해 두 개 이상의 센서를 이용한 상호 보완적인 정보 융합 연구가 활발히 이루어지고 있다. 이를 위해서는 영상정합 과정이 필수적이다. 영상정합은 같은 장면에 대한 두 개 이상의 영상들을 하나의 좌표계로 정렬시켜주는 과정으로서, 정합의 대상이 되는 영상으로는 시차를 두어 획득한 영상, 서로 다른 시점에서 획득한 영상, 혹은 서로 다른 센서로부터 획득된 영상이 있다[2].

영상정합은 정합과정에 따라 크게 영역기반 정합 기법과 특징기반 정합 기법으로 분류될 수 있다. 영역기반 (Area-based) 정합 기법은 영상 전체의 밝기 정보를 이용하여 영상의 유사성 척도를 측정하고 영상정합을 수행하는 방법이다. 다중센서 영상정합의 경우에는 영상 간의 밝기 정보의 관계가 복잡하거나 잘 알려져 있지 않은 경우가 대부분이어서 밝기 값 차이를 최소화시키는 등의 단순한 방법은 영상정합의 정확도가 매우 낮다. 따라서 밝기 정보의 통계적 특성을 이용하는 방법이 제안되어왔는데, 그 중 대표적으로 상호정보 (MI : Mutual Information)와 정규상호정보 (NMI : Normalized Mutual Information)를 유사성 척도로 이용하는 방법이 있다[3-6]. 최근에는 밝기 정보뿐만 아니라 공간정보로서 기울기 정보를 함께 이용하는 방법이 제안되었다

[7,8]. 영역기반 정합 기법은 영상 전체를 다루기 때문에 정확도가 높은 반면, 연산량이 많고 밝기 변화 및 잡음에 민감하며 국부 최대점에 빠지는 경우가 발생할 수 있는 문제점이 있다. 특징기반 (Feature-based) 정합 기법은 영상의 특징을 추출하고, 추출된 특징 간의 공통된 특성 정보를 이용한 대응 관계를 설정하여 영상을 정렬시키는 방법이다. 대표적으로 사용되는 특징들로는 코너 (corners), 에지 (edges), 외곽선 (contours), 교차점 (line intersections) 등이 있다. 특징기반 정합 방법은 특징만을 이용하기 때문에 연산량이 적고 밝기 변화에 강인하여 영역기반 정합 기법의 단점을 보완할 수 있지만, 선별된 대응 특징점의 정확도에 따라 정합 성능이 크게 좌우된다.

서로 다른 특성의 센서로부터 획득된 영상은 밝기 정보가 서로 다른 특성을 가지기 때문에 단일 센서로 획득된 영상에 비해 정합에 더 많은 어려움이 있다[9]. 예를 들어 가시광선 영상과 적외선 영상의 경우, 가시광선 영상의 밝기 정보는 물체의 색이나 주변의 밝기를 나타내는 반면에 적외선 영상의 밝기 정보는 물체의 온도를 나타내기 때문에, 두 영상이 서로 다른 특성의 밝기 값 분포를 갖고 그 관계가 복잡하다. 이러한 상황에서 다중센서 영상정합을 위한 특징기반 정합 기법과 관련된 최근의 논문에서는 영상기반 정합 기법에서 영상 간의 유사성 척도 사용되는 정규상호정보를 특징기반 정합 기법에 적용하는 방법이 제안되었다[10,11]. 그러나 가시광선 영상과 적외선 영상의 경우에는 밝기 값의 관계가 복잡할 뿐 아니라, 한 쪽 영상에서 나타나는 특징들이 다른 쪽에서는 나타나지 않는 상보적인 특성이 있다. 또한, 한 쪽 영상에서 하나의

밝기 값이 다른 쪽 영상의 여러 밝기 값에 대응되는 경우와 그 반대의 경우가 존재한다. 이러한 지역적 특성들로 인해 적외선 영상과 가시광선 영상의 밝기 값은 전역적인 통계적 상관성을 보장하지 못하는 경우가 많다[9]. 결국 기존의 방법은 실제 환경에서 대응 특징점의 신뢰도가 낮아 정확한 정합 결과를 제공하지 못하는 경우가 많다. 본 논문에서는 정규상호정보만을 이용한 다중 센서 영상정합 기법의 문제점을 보완하기 위하여 영상의 공간정보로서 기울기 (gradient) 방향 정보를 정규상호정보와 결합하고 이를 대응 특징점 선별을 위한 유사성 척도로 이용하는 방법을 제안한다. 제안된 유사성 척도를 이용하여 대응 특징점의 정확도를 향상시킴으로써 기존의 방법보다 더 정확하고 안정적인 영상정합이 가능하다. 또한 단일 문턱값 (threshold)를 이용하는 Harris 코너 검출기[12]를 통해 특징점을 추출할 경우, 영상 장면에 따라 추출되는 특징점의 수가 상이하게 된다. 특징점이 너무 적을 경우 대응 특징점이 부족하여 정합이 실패할 가능성이 높으며, 반대로 특징점이 너무 많을 경우 연산량이 증가하고 잘못된 대응 특징점이 발생할 가능성이 높아진다. 위와 같은 문제를 해결하기 위해 본 논문에서는 영상 장면에 상관없이 일정한 개수의 특징점을 추출할 수 있는 적응적 Harris 코너 검출기를 제안하였다.

본 논문의 구성은 다음과 같다. 2장에서는 제안된 다중 센서 영상정합 기법의 전체구조를 소개하고, 제안하는 적응적인 Harris 코너 검출기를 통한 특징점 추출 방법에 대해 설명한다. 그 다음에 정규상호정보와 기울기 방향 정보를 이용한 대응 특징점 선별 방법을 제안한 후 개선된 RANSAC (Random Sample Consensus) 기법을 이용한 변환함수 추정 방법에 대해 설명한다. 3장에서는 실험을 통해 제안한 영상정합 기법과 기존의 정규상호정보를 이용한 영상정합 기법의 성능을 비교하고, 마지막으로 4장에서는 결론을 내린다.

II. 본 론

1. 제안하는 영상정합 기법의 전체구조

그림 1은 제안하는 영상정합 기법의 전체 구조이다. 두 개의 다른 센서에서 획득한 영상이 입력되면, 각각의 영상에 대해 제안하는 적응적 Harris 코너 검출기를 이용하여 특징점을 추출한다. 다음으로 각 영상의 추출된 특징점을 중심으로 한 패치 (patch) 영상 간의 제안하는 정규 상호정보와 기울기 방향 정보를 결합한 유사성 척도를 계산한 후, 가장 높은

유사 척도를 갖는 대응 특징점을 선별한다. 이렇게 선별된 대응 특징점을 이용하여 기존의 RANSAC 기법[13]을 개선한 Locally Optimized RANSAC (Lo-RANSAC) 기법[14]으로 잘못된 대응 쌍을 제거하고, 변환 함수를 추정하여 영상 정합을 수행한다.

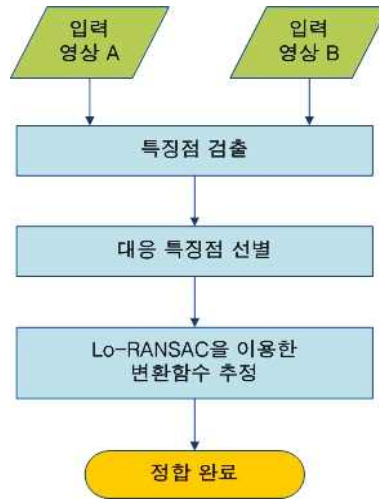


그림 1. 제안하는 영상정합 기법의 전체 구조
Fig. 1. Overall structure of the proposed image registration method

2. 특징점 검출

특징점 검출 방법은 Harris 코너 검출기[12], SUSAN[15], 및 최근에 Lowe가 제안한 SIFT[16] 등이 있다. SIFT 기법은 영상의 크기에 불변한 특징을 이용하기 때문에 물체 인식에서 많이 활용되지만, 연산량이 많고 서로 다른 센서로부터 획득한 영상의 경우에는 특징 벡터가 다르기 때문에 다중 센서 영상정합을 위해 사용하기에는 어려움이 있다. 본 논문에서는 연산량 측면을 고려하여 Harris 코너 특징을 사용하였다.

Harris 코너 검출기는 영상의 가로와 세로 방향에 대한 1차 미분값으로 구성된 행렬 $M \in R^{2 \times 2}$ 을 이용하며 이는 식 (1)과 같다.

$$M = \begin{bmatrix} I_x^2(x,y) & I_x(x,y)I_y(x,y) \\ I_x(x,y)I_y(x,y) & I_y^2(x,y) \end{bmatrix} \quad (1)$$

여기서 I_x 와 I_y 는 영상 I 의 각각 x 와 y 방향으로의 1차 미분을 나타내며, 위치 (x,y) 에서의 Harris 응답 R 은 식 (2)와 같다.

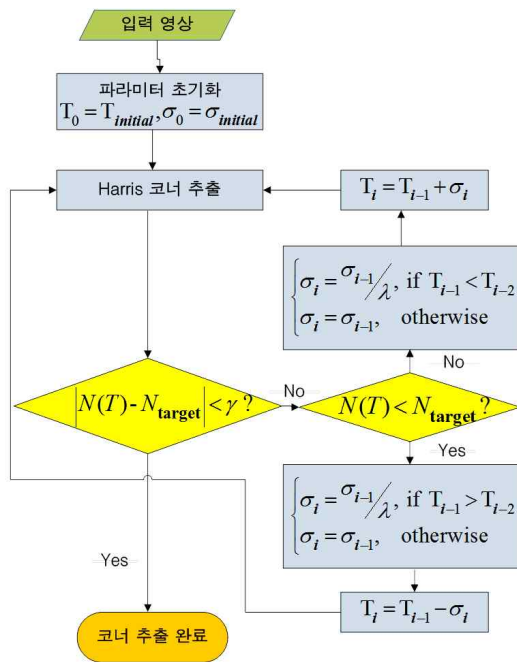


그림 2 적응적 Harris 코너 검출기의 구조
Fig. 2 Block diagram of the adaptive Harris corner detector

$$R(x,y) = \det(M) + k(\text{Trace}(M))^2 \quad (2)$$

k 는 Harris에 의해서 제안된 매개변수(0.04)이며, $\det(M)$ 과 $\text{Trace}(M)$ 은 각각 행렬 M 의 행렬식과 대각합을 의미한다. 위의 Harris 응답 R 에서 문턱값을 통해 특징점의 수가 결정되는데, 서로 다른 영상 장면에 동일한 문턱값을 적용할 경우 영상 장면에 따라 추출되는 특징점의 수가 상이하게 된다. 이 때 특징점이 너무 적을 경우 올바른 대응 특징점이 부족하여 정합에 실패할 가능성이 높으며, 반대로 특징점이 너무 많을 경우 연산량이 증가하고 잘못된 대응 특징점이 발생할 가능성이 커지게 되는 문제가 발생한다. 위와 같은 문제를 해결하기 위해서 본 논문에서는 영상 장면에 상관없이 자동으로 문턱값을 조절하여 일정한 개수의 특징점을 추출할 수 있는 적응적 Harris 코너 검출기를 제안한다.

그림 2는 제안하는 적응적 Harris 코너 검출기의 구조를 나타낸다. T 는 Harris 응답 R 의 문턱값이고, $N(T)$ 는 문턱값 T 를 적용하여 추출된 코너 개수이다. $N(T)$ 는 T 에 대한 감소함수이므로 문턱값 T 를 조절함으로써 추출되는 코너 개수를 조절할 수 있다. N_{target} 과 γ 는 각각 사용자가 추출하

고자 하는 특징점의 개수와 허용 범위로서, 최종적으로 추출되는 코너 개수는 $N_{target} - \gamma$ 와 $N_{target} + \gamma$ 사이가 된다. σ 는 문턱값 T 를 조절하는 증분량(Step size)으로서, 추출된 코너 수가 적을 경우 문턱값 T 를 σ 만큼 감소시키고 반대로 많을 경우 T 를 σ 만큼 증가시킨 후 다시 코너를 추출하는 과정을 수행한다. 이 때 이전 과정에서 T 가 감소되었고 현재 과정에서 T 를 증가시켜야 하는 경우, σ 를 λ 배만큼 감소시키며, 반대의 경우도 마찬가지이다. 여기서 λ 는 0과 1사이의 값을 갖는 상수다.

3. 대응 특징점 선별

본 논문에서는 두 입력 영상에서 검출된 특징점 중에서 서로 대응되는 특징점 쌍을 선별하기 위해, 특징점을 중심으로 한 패치 영상 간의 정규 상호정보와 기울기 방향 정보를 이용하는 방법을 제안하였다. 제안된 방식에서 각 영상에서 추출된 특징점에 대한 패치 영상을 획득한 후, 이 패치 영상 간의 제안하는 유사성 척도를 계산하여 특징점 간의 유사성을 판별한다. 이 때 제안하는 유사성 척도가 가장 큰 특징점 쌍을 대응 특징점으로 결정한다.

3.1 정규상호정보

상호정보는 두 확률변수의 통계적 상관성을 나타내는 척도로서, 영상정합에서는 두 영상의 밝기 값이 확률변수가 될 수 있다. 두 영상 A와 B에 대한 상호 정보는 다음과 같이 정의할 수 있다[3].

$$MI(A,B) = H(A) + H(B) - H(A,B) \quad (3)$$

여기서 $H(A)$ 와 $H(B)$ 는 각각 영상 A와 B의 엔트로피(shannon entropy)로 아래 식 (4)와 같이 정의되며, $H(A,B)$ 는 두 영상의 결합 확률 분포에 대한 엔트로피로 식 (5)와 같이 정의된다.

$$H(X) = - \sum_x p_X(x) \log p_X(x) \quad (4)$$

$$H(X, Y) = - \sum_{x,y} p_{XY}(x,y) \log p_{XY}(x,y) \quad (5)$$

식 (4)의 $p_X(x)$ 는 X 의 확률 분포이며, 식 (5)의 $p_{XY}(x,y)$ 는 X 와 Y 의 결합 확률 분포를 나타낸다. 영상의 경우, 엔트로피는 영상 전체 밝기 값의 불확실성(uncertainty)

의 척도가 된다. 엔트로피는 정보량으로 표현되기도 하는데, 상호정보는 두 확률변수의 공통 정보를 의미하게 된다. 영상 정합에서는 두 영상의 정합이 정확히 이루어졌을 때 두 영상의 밝기 값의 통계적 상관성이 최대가 되며 공통적인 정보량인 상호정보가 최대가 된다는 것을 이용하게 된다. 그러나 엔트로피는 배경 크기에 대한 의존도가 높기 때문에 앞서 소개한 상호정보는 중첩된 영역에 민감한 단점이 있다. 이렇게 중첩된 영역에 민감한 상호정보의 단점을 보완하기 위해 결합 엔트로피로 정규화한 정규상호정보가 식 (6)과 같이 정의된다.

$$NMI(A,B) = \frac{H(A) + H(B)}{H(A,B)} \quad (6)$$

3.3 기울기 방향 정보

영상의 기울기를 얻으려면 모든 화소 위치에서 x 와 y 방향으로의 편미분을 계산해야 한다. 편미분을 구하는 방법으로는 기본적으로 Prewitt 혹은 Sobel 마스크로 영상을 필터링하여 구하는 방법이 있는데[17], 본 논문에서는 좀 더 정확한 기울기를 구하기 위하여 Farid가 제안한 필터[18]를 이용하여 편미분을 계산하였다. 이 때 이렇게 구한 편미분은 앞에서 특징점 검출 시 Harris 코너 검출기의 Harris 응답을 구하는 데에도 이용한다.

영상 I 의 위치 (x,y) 에서의 기울기 방향의 각도는 아래 식 (7)과 같이 계산될 수 있다.

$$\theta_I(x,y) = \tan^{-1} \left[\frac{I_y(x,y)}{I_x(x,y)} \right] \quad (7)$$

여기서 I_x, I_y 는 영상 I 의 위치 (x,y) 에서 각각 x 방향과 y 방향으로의 편미분을 나타낸다. 이를 이용하여 두 영상 A와 B의 위치 (x,y) 에서의 기울기 방향의 각도 차이는 다음과 같이 나타낼 수 있다.

$$\Delta\theta(x,y) = \theta_A(x,y) - \theta_B(x,y) \quad (8)$$

이 때 두 영상에서 같은 물체에 대한 기울기 방향은 동일하며, 다중센서 영상의 경우 밝기 대비 (contrast)가 반전되는 경우가 있으므로 기울기 방향의 각도 차이 $\Delta\theta$ 는 0° 혹은 180° 가 될 것이다. 따라서 같은 물체에 대한 기울기 방향의 각도 차이 값이 불일치되는 문제를 해결하기 위해 본 논문에서는 아래 식 (9)와 같이 기울기 방향의 일치성을 나타내는 함수를 사용한다[19].

$$c\{\Delta\theta(x,y)\} = \frac{1 + \cos\{2\Delta\theta(x,y)\}}{2} \quad (9)$$

위의 $c\{\Delta\theta(x,y)\}$ 는 0에서 1사이의 값을 가진다. 값이 1일 때 두 기울기 방향의 각도 차이가 0° 혹은 180° 이고, 두 기울기 방향이 일치함을 나타낸다. 반대로 값이 0일 때 두 기울기 방향의 각도 차이가 90° 이고, 두 기울기 방향이 불일치함을 나타낸다.

3.4 제안하는 유사성 척도

앞서 언급하였듯이 가시광선 영상과 적외선 영상 정합의 경우 정규상호정보만으로는 대응점 선별 시 정확도가 저하된다. 일반적으로 기울기 정보를 이용한 영상정합 방법들은 기울기 크기 정보 (gradient magnitude)와 방향 정보 (gradient orientation)를 함께 사용한다. 그러나 다중센서 영상의 경우 대응하는 동일 물체라도 영상의 밝기 대비가 다르기 때문에 기울기 크기는 신뢰할만한 정보를 제공하지 못한다. 반면에 기울기 방향은 동일 물체에 대해서 같기 때문에 다중 센서 정합에 이용하기에 적합하다. 본 논문에서는 대응점 선별을 위한 유사성 척도로 정규상호정보와 영상의 공간정보로서 기울기 방향의 일치성 척도를 결합하는 방식을 다음과 같이 제안한다.

$$ONMI(A,B) = O(A,B) \times NMI(A,B) \quad (10)$$

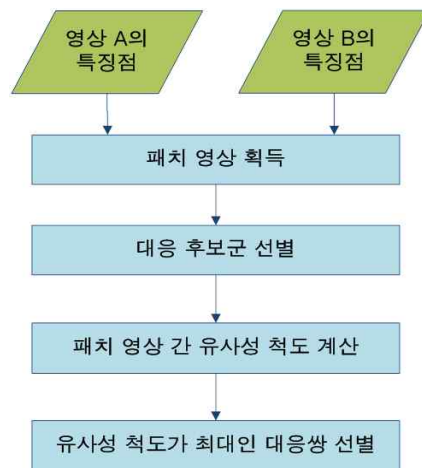


그림 3 대응점 선별 과정 개요
Fig. 3. An outline of selecting corresponding point

여기서 A, B는 각각 영상 A와 B를 나타내며, $O(A, B)$ 는 식 (11)과 같다.

$$O(A, B) = \sum_{x, y} c\{\Delta\theta(x, y)\} \quad (11)$$

대응점 선별 과정에서 위 식 (10)과 (11)의 A와 B는 각각의 입력 영상에서 추출된 특징점 주위의 패치 영상이 된다. 두 패치 영상 A와 B가 같은 물체를 나타낼수록 정규상호정보가 증가할 뿐만 아니라, 기울기 방향이 일치하는 부분이 증가함으로써 $O(A, B)$ 와 함께 제안하는 유사성 척도 $ONMI(A, B)$ 가 최대화될 것이다. 이를 이용하여 대응점을 선별하는 과정에 대해서는 다음 절에서 자세히 다룬다.

3.5 대응점 선별 과정

그림 3은 대응점 선별 과정의 전체적인 개요를 나타낸다. 두 개의 입력 영상에서 검출된 특징점들 중에서 서로 대응되는 특징점 쌍을 선별하기 위해, 특징점을 중심으로 하는 $N \times N$ 크기의 패치를 획득한다. 그리고 동일한 장면에 대해 획득된 두 영상의 경우 특징점들이 국지성(localization)을 나타내기 때문에, 대응 특징점 선별 시 서로 너무 멀리 떨어진 특징점들은 제외시킬 수 있다. 이렇게 함으로써 잘못된 대응 쌍을 줄일 수 있고, 연산량도 감소시킬 수 있다.

입력 영상 A의 i 번째 특징점을 p_i^A , 이를 중심으로 한 $N \times N$ 크기의 패치 영상을 $W_N(p_i^A)$, 그리고 입력 영상 B의 j 번째 특징점을 p_j^B 이를 중심으로 한 패치 영상을 $W_N(p_j^B)$ 라고 하면, 영상 A의 특징점 p_i^A 와 대응되는 영상 B의 특징점 cp_i^A 는 다음과 같이 구할 수 있다.

$$cp_i^A = \arg \max_{p_j^B} ONMI(W_N(p_i^A), W_N(p_j^B)), \text{ for } p_j^B \in C_i^A \quad (12)$$

여기서 C_i^A 는 식 (13)와 같이 p_i^A 와 대응될 수 있는 영상 B의 특징점 중에서 국지성을 고려한 후보들의 집합을 의미한다.

$$C_i^A = \{p_j^B | dist(p_i^A, p_j^B) < d_{th}\} \quad (13)$$

$dist(\cdot)$ 는 두 특징점 사이의 거리를 구하는 함수로 여기서는 상한 노름(maximum norm)을 사용했다.

그림 4은 대응점 선별 과정을 나타낸다. (a)와 (b)는 각각 적외선 영상 A와 가시광선 영상 B의 일부분이며, 작은 사각

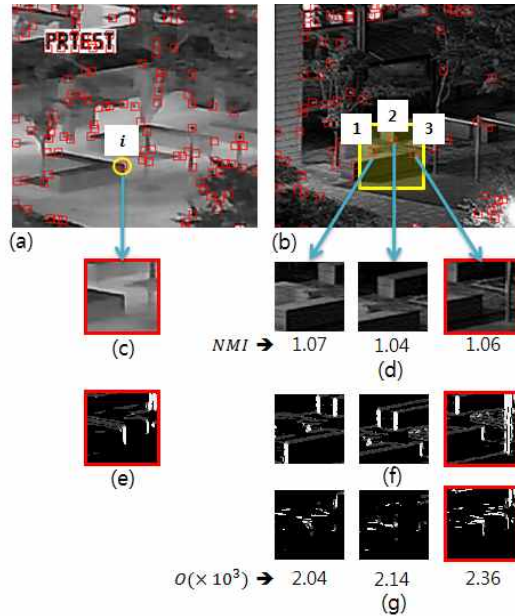


그림 4. 제안한 유사성 척도를 이용한 대응점 선별 과정. (a), (b) 적외선 및 가시광선 영상에서 추출된 특징점 (c), (d) 각 특징점의 패치 영상 (e), (f) 패치 영상의 기울기 방향 (g) <e>와 <f>간의 기울기 방향 각도 차이

Fig. 4. A process of finding corresponding points using proposed similarity measure. (a), (b) Extracted feature point from IR and EO images. (c), (d) Patch of each feature point. (e), (f) Gradient orientation of each patch. (g) Gradient orientation angle difference between <e> and each <f>

형은 추출된 특징점을 나타낸다. 이 때 (b)의 큰 사각형 영역 내의 특징점들이 (a)의 i 번째 특징점 p_i^A 에 대한 대응 후보군이다. 즉, $C_i^A = \{p_1^B, p_2^B, p_3^B\}$ 이다. (c)와 (d)는 각각의 특징점에 대한 패치 영상을 나타내고, (e)와 (f)는 각 패치 영상에 대한 기울기 방향을 나타낸 영상으로, 기울기 방향의 각도 θ 를 $100(\cos 2\theta + 1) + 50$ 의 밝기 값으로 나타냈다. 따라서 [50, 250] 범위의 밝기 값을 가지는데, 기울기가 존재하지 않는 부분은 0으로 처리했다. (g)는 (e)와 (f) 각각의 패치 간의 기울기 방향각도 차이 $c(\Delta\theta)$ 를 $100\{\cos(2\Delta\theta) + 1\} + 50$ 의 밝기 값으로 나타낸 것이다 즉, 흰색 영역이 많을수록 기울기 방향이 일치하는 부분이 많은 것이므로 특징점 간의 유사성이 더 높은 것으로 볼 수 있다. 실제 실험 시 본 논문에서 제안한 유사성 척도 $ONMI(W_N(p_i^A), W_N(p_j^B))$ 을 구하기 위해 계산된 $NMI(W_N(p_i^A), W_N(p_j^B))$ 값과 $O(W_N(p_i^A), W_N(p_j^B))$ 값을 각각 (d)와 (g)의 패치 영상 밑에 표기하였다. 육안으로 확인할 수 있듯이 실제로 특징점 p_i^A 는 p_3^B 과 대응되는데, 정규상호정

보를 최대화 하는 특징점을 대응 특징점으로 결정할 경우 p_1^B 가 대응점으로 선별되는 오류를 범하게 된다. 반면에 본 논문에서 제안한 유사성 척도 $ONMI$ 를 이용할 경우 올바른 대응점 p_3^B 를 선별할 수 있게 된다. 또한 앞에서 설명한 선별 과정을 반대로 영상 B의 특징점 p_j^B 에 대응하는 특징점 cp_j^B 를 구할 수 있는데, 이 때 앞에서 구한 특징점 쌍 (p_i^A, cp_i^A) 와 (cp_j^B, p_j^B) 가 서로 일치할 경우 이 특징점을 서로의 대응 특징점으로 최종 결정할 수 있다. 이와 같은 상호 비교 기법[10]을 수행함으로써 유사한 패턴이 반복되는 영상의 경우에 높은 신뢰도를 갖는 대응점만을 추출할 수 있다.

4. 변환함수 추정

4.1 변환 모델 및 변환 함수

변환 모델은 획득하고자 하는 장면이 평면이라는 가정 하에 몇 개의 파라미터로 정의되는데, 변환하고자 하는 영상의 특징에 따라 이 변환 모델을 선택하고, 그 변환 모델을 이루는 파라미터를 결정함으로써 변환 함수가 정의된다. 본 논문에서는 변환하고자 하는 2D 영상이 3D의 실제 물체들이 투영된 영상이므로 물체의 원근 관계를 정확하게 표현할 수 있는 원근적 변환 (projective transform) 모델을 사용하였다 [9]. 원근적 변환 모델은 다음과 같이 나타낼 수 있다.

$$\begin{aligned} x' &= \frac{h_1x + h_2y + h_3}{h_7x + h_8y + 1} \\ y' &= \frac{h_4x + h_5y + h_6}{h_7x + h_8y + 1} \end{aligned} \quad (14)$$

여기서 (x, y) 는 변환 영상의 좌표이고, (x', y') 는 기준 영상의 좌표다. 그리고 h_1, h_2, \dots, h_8 은 원근적 변환 모델의 파라미터다.

4.2 RANSAC을 이용한 변환 함수 추정

앞에서 대응점 선별 과정을 수행한 후에도 잘못된 대응점이 발생할 수 있는데, 이러한 잘못된 대응점은 정합 성능의 정확도에 영향을 미치게 된다. 따라서 이러한 잘못된 대응점을 제거하고 변환함수를 구해야 한다. 본 논문에서는 대표적인 RANSAC 기법[13]을 이용하여 변환 함수를 추정한다. RANSAC 기법을 이용한 변환함수 추정 방법은 다음과 같다.

선별된 대응점 중에서 임의로 N개의 대응점을 선택한다. 여기서 N은 변환 모델 파라미터를 정할 수 있는 최소한의 대

응점 개수로서, 원근적 변환 모델의 경우 $N=4$ 이다.

선택한 대응점을 이용하여 변환 모델 파라미터를 계산한다.

나머지 변환 영상의 대응점을 과정 2에서 계산한 변환 함수에 따라 변환시킨 후 식 (15)와 같이 변환 에러를 계산한다.

$$E = \sqrt{(x_i - x_i')^2 + (y_i - y_i')^2} \quad (15)$$

여기서 (x_i, y_i) 는 기준 영상의 대응점 좌표이고, (x_i', y_i') 는 변환 영상의 대응점이 변환된 좌표이다.

변환 에러가 사전에 설정된 문턱값보다 작은 대응점을 유효한 대응점으로 분류한다.

1번부터 3번까지의 과정을 일정 횟수 동안 반복한 후, 가장 많은 유효한 대응점을 갖는 변환 모델 파라미터를 최종 결과로 선정한다.

본 논문에서는 위의 과정 3에 국부 최적화 (Local optimization) 과정을 추가하여 기존의 RANSAC 기법의 정확도와 속도를 개선한 LO-RANSAC 기법[14]으로 변환 함수를 추정하여 영상정합을 수행한다.

III. 실험 결과

제안한 다중센서 영상정합 기법의 성능을 평가하기 위하여 동일한 장면에 대해 획득한 가시광선 영상과 적외선 영상을 대상으로 실험을 하였다. 실험에 사용한 영상의 개수는 5쌍이며, 크기는 모두 720×480 이다. 정합 수행 시 패치의 크기는 60×60 으로 설정하였다.

기존에 영상정합의 정확도를 평가하기 위한 정합 오차로서 TRE (Target Registration Error)[20]를 사용하는 방법이 제안되었는데, TRE는 추정된 변환 함수로 변환된 표적점 (target point)들과 실제 변환 함수 (ground truth)로 변환된 표적점들의 평균 위치 차이를 의미한다. 그러나 이 척도는 표적점 이외의 점들에 대한 오차가 고려되지 않기 때문에, 본 논문에서는 영상 전체에 대한 오차를 고려하기 위해 아래와 같이 정의되는 정합 오차 (Registration Error)를 이용하였다.

$$RE = \frac{1}{A} \sum_{i=1}^A \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2} \quad (16)$$

여기서 변환 영상의 i 번째 점의 좌표를 (x_i, y_i) 라 한다면, (\hat{x}_i, \hat{y}_i) 와 $(\tilde{x}_i, \tilde{y}_i)$ 는 각각 실제 변환 함수 (ground truth)와



그림 5. (a) 가시광선 영상 (b) 적외선 영상 (c) 정합 전 영상 (d) 정합 결과
Fig. 5. (a) EO and (b) IR images. (c) Image before registration and (d) Registration result

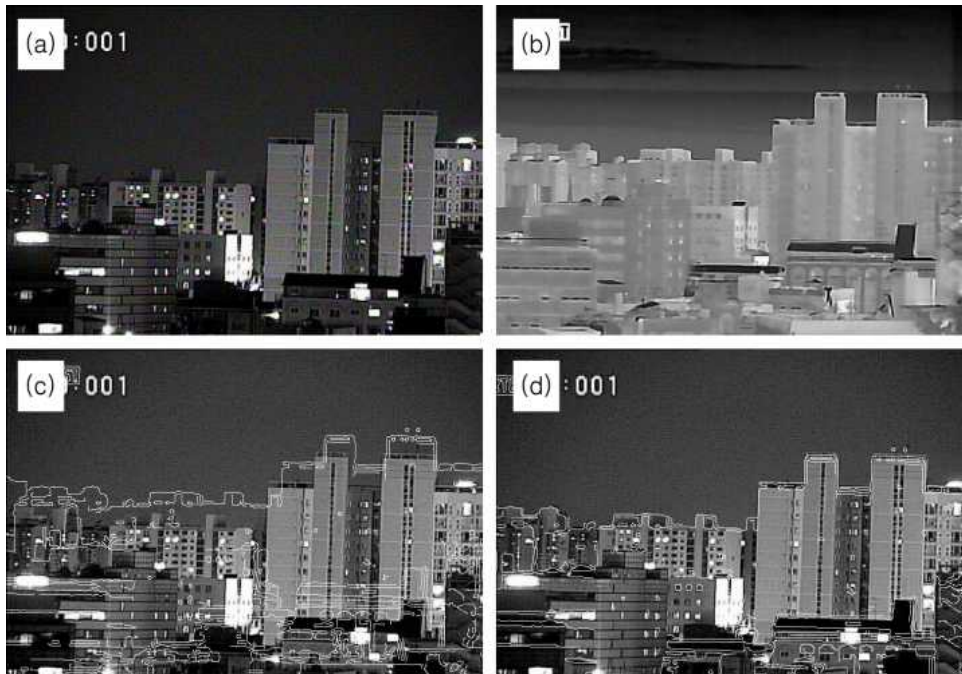


그림 6. (a) 가시광선 영상 (b) 적외선 영상 (c) 정합 전 영상 (d) 정합 결과
Fig. 6. (a) EO and (b) IR images. (c) Image before registration and (d) Registration result

추정된 변환 함수로 변환된 좌표를 나타낸다. A 는 변환 영상의 전체 픽셀수다. 즉, RE 는 변환 영상의 모든 점들에 대하여 실제 변환 함수로 변환한 점들과 추정된 변환 함수로 변환한 점들의 평균 거리 (단위: 픽셀)를 의미한다. 이 때 실험을 위해 촬영한 영상은 실제로 정확한 변환 함수를 알 수 없기 때문에, 각 영상 별로 15개~25개 사이의 대응점을 직접 선별하여 변환 함수 (ground truth)를 계산하였다. 또한 영상정합의 강인함을 평가하기 위하여 Y. S. Kim et al[7]이 평가한 방식과 같이 각 영상 별로 100번의 정합을 수행하였으며, 정합 결과의 RE 가 4픽셀 이하일 때 정합이 성공한 것으로 간주하여 정합 성공률을 계산하였다. 이 정합 성공의 판단 기준치는 실제 육안으로 정합 결과 영상을 확인하여 실험적으로 선택한 수치다.

표 1. 추출된 코너 개수에 따른 옳은 대응점의 개수와 비율, 정합 성공률 및 연산 시간 (평균)
Table 1. The average number and rate of correct corresponding pair, success rate, and computation time

N_{target}	옳은 대응점 개수	옳은 대응점 비율	정합 성공률	연산 시간 (초)
100	8.6	23.34	0.04	0.96
200	22.2	32.38	0.29	2.37
300	40	36.63	0.60	2.75
400	51.8	35.30	0.71	3.29
500	64	34.87	0.88	4.80
600	73.6	33.40	0.88	6.14
700	76.8	30.90	0.84	8.22
800	81.6	29.15	0.82	9.88
900	83	26.99	0.79	12.11
1000	85.4	25.36	0.76	14.48

우선 본 논문에서 제안한 적응적 Harris 코너 검출기의 효용성을 입증하기 위하여 추출된 코너 개수에 따른 정합 성능을 비교한다. 실험 시 N_{target} 은 100에서 1000까지 100개 단위로 설정하였고, 적응적 Harris 코너 검출기의 파라미터는 $T_{initial} = 50$, $\sigma_{initial} = 5$, $\lambda = 0.5$, $\gamma = 30$ 으로 설정하였다. 표 1은 각 영상 별로 추출된 코너 개수에 따른 옳은 대응점의 개수와 비율, 정합 성공률 및 연산 시간의 평균을 나타낸다. 만약 M 개의 대응점이 선별되었을 경우 그 중 옳은 대응점 집합 C_{true} 는 다음과 같이 나타낼 수 있다.

$$C_{true} = \{(p_i^A, cp_i^B) | dist(p_i^A, T_{gt}(cp_i^B)) < \eta \text{ for } i = 1, 2, \dots, M\} \quad (17)$$

표 2. 추출된 코너 개수에 따른 Harris 응답의 문턱값 T
Table 2. Threshold T of Harris response respect to the number of extracted corners

N_{target}	문턱값 T	
	영상 1	영상 2
100	95	25
200	45	7.5
300	15	2.5
400	10	1.875
500	5	0.9375
600	3.75	0.625
700	2.5	0.3906
800	1.5625	0.2344
900	0.9375	0.1953
1000	0.4688	0.1563

여기서 $T_{gt}(\cdot)$ 는 실제 변환함수 (ground truth)로 특징점의 좌표를 변환해주는 함수이며, 여기서 $dist(\cdot)$ 는 유클리드 노름 (euclidean norm)을 사용하였고 픽셀 단위다. 즉 선별된 대응점 (p_i^A, cp_i^B) , for $i = 1, 2, \dots, M$ 중에서 p_i^A 와 실제 변환함수로 변환된 $T_{gt}(cp_i^B)$ 의 거리가 일정 문턱값 (실험 시 $\eta = 2$) 미만인 대응점을 옳은 대응점으로 판별한다. 이 때 옳은 대응점 집합 C_{true} 의 원소 개수가 옳은 대응점 개수가 되며, 이를 선별된 대응점 개수로 나눈 것이 옳은 대응점 비율이다. 그리고 정합 성공률은 한 영상에 대한 100번의 정합 수행 중 정합이 성공한 횟수를 전체 정합 시행 횟수 100으로 나눈 수치다. 실험 결과 코너를 너무 적게 추출할 경우 옳은 대응점의 개수가 부족하여 정합 성공률이 낮은 것을 볼 수 있다. 반대로 코너를 너무 많이 추출할 경우 잘못된 대응점이 증가하면서 옳은 대응점 비율이 낮아지고, 그에 따라 정합 성공률이 떨어질 뿐만 아니라 연산 시간이 대폭 증가하는 것을 알 수 있다.

표 2는 적외선 영상 1과 2에 대하여 각 코너 개수 별로 추출하기 위한 Harris 응답의 문턱값 T 를 나타낸다. 만약 기존의 Harris 코너 검출기를 사용할 경우, $T = 5$ 로 설정하였을 때 영상 1이 입력될 경우 약 500개의 코너가 추출되어 정합 성능이 좋지만, 영상 2의 경우 200~300개 사이의 코너가 추출되어 정합 성능이 현저히 저하되는 문제가 발생할 것이다. 그리고 $T = 1$ 로 설정할 경우 영상 2에 대해서는 정합 성능이 좋지만, 영상 1의 경우에는 코너가 너무 많이 추출되어 정합의 정확도가 떨어질 뿐만 아니라 연산시간이 대폭 증가할 것이

다. 따라서 영상정합 수행 시 본 논문에서 제안한 적응적 Harris 코너 검출기를 통해 적절하게 코너를 추출한다면, 특징점을 추출하기 위해 소요되는 시간은 기존 Harris 검출기에 비해 약 10% 정도로 미미하게 증가되지만 정합 성공률을 보장하는 적정수의 옳은 대응점을 선별하게 되어 전체 정합 과정에 소요되는 시간이 과도하게 증가하는 것을 방지함과 동시에 안정적인 영상 정합을 수행할 수 있다.

다음으로 대응 특징점 선별 시 본 논문에서 제안한 유사성 척도를 이용한 방법과 기존의 정규상호정보를 이용한 방법 [10]과의 성능을 비교한다. 이 때 기존의 방법은 변환함수 추정 시 최소자승법 (Least Squares Fitting)을 이용한 업데이트 방법을 사용하였는데, 동일한 조건 및 환경에서 대응점 선별 성능을 비교하기 위해서 변환 함수 추정부분은 본 논문에서 사용한 Lo-RANSAC 기법을 기존 방법에 동일하게 적용하여 비교하였다. 정합을 위한 특징점의 수 N_{target} 은 500으로 설정하였다. 표 3에서 알 수 있듯이 가시광선 영상과 적외선 영상은 전역적으로 통계적 상관성을 보장하지 못하기 때문에 기존의 정규상호정보만을 이용한 방법은 대응 특징점의 정확도가 낮아 정합의 안정성이 떨어진다. 반면에 제안한 방법은 기술기 방향 정보를 결합함으로써 기존 방법에 비해 대응 특징점의 정확도가 높고, 그에 따라 정합의 안정성도 향상된 것을 확인할 수 있다. 연산 시간은 제안한 방법이 기존의 방법보다 평균 약 30% 증가했다.

표 3. 기존 방법과 제안한 방법의 정합 성능
Table 3. Registration performance of conventional and proposed method

영상	옳은 대응점 수		옳은 대응점 비율		평균 정합 오차 (RE)		정합 성공률	
	기존	제안	기존	제안	기존	제안	기존	제안
1	31	47	28.70	31.13	6.17	1.56	0.44	0.91
2	48	85	35.04	45.70	4.20	2.81	0.71	0.86
3	83	113	47.16	48.29	3.51	2.28	0.73	0.92
4	29	41	22.31	23.03	4.75	3.88	0.41	0.69
5	32	43	28.07	27.92	3.66	2.20	0.80	0.94
평균	44.6	65.8	32.26	35.21	4.46	2.55	0.62	0.86

그림 5와 6는 각각 영상 1과 영상 2에 대한 영상정합 결과를 보여준다. 각 그림 별로 (a)와 (b)는 입력된 가시광선과 적외선 영상이며, (c)는 적외선 영상의 Canny 엣지[21]를 가시광선 영상에 겹친 영상이고, (d)는 정합된 적외선 영상의 Canny 엣지를 가시광선 영상에 겹친 영상을 나타낸다.

IV. 결 론

본 논문에서는 서로 다른 센서로부터 획득한 영상에 대해 특징점을 추출하고, 제안한 유사성 척도를 이용하여 대응 특징점 쌍을 선별한 후, RANSAC 기법으로 변환함수를 추정하는 영상 정합 기법을 제안하였다. 특징점 추출 시 제안한 적응적 Harris 코너 검출기를 사용하여 영상 장면이 상관없이 일정한 수의 특징점을 추출함으로써, 안정적이고 효율적인 정합 수행이 가능 하였다. 또한 대응 특징점 선별 시 기술기 방향 정보를 정규상호정보와 결합하여 특징점 간의 유사성을 판별하는 방식을 제안함으로써, 기존의 정규상호정보만을 이용한 방법보다 대응 특징점의 정확도와 정합 성능을 향상시켰다. 현재의 연구에서는 특징점 개수를 적절하게 임의로 설정하였으나, 영상에 따라 최적의 특징점 개수를 결정할 수 있는 이론적인 연구를 향후 진행할 예정이다.

참고문헌

- [1] M. J. Kim, H. S. Ko, et al, "Robust Video Super Resolution Algorithm Using Measurement Validation Method and Scene Change Detection," EURASIP Journal on Advances in Signal Processing, Vol. 2011, No. 103, Nov. 2011.
- [2] B. Zitova and J. Flusser, "Image Registration Methods : A Survey," Image and Vision Computing, pp. 977-1000, June. 2003.
- [3] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-Information-Based Registration of Medical Images : A Survey," IEEE Transactions on Medical Imaging, Vol. 22, No. 8, pp. 986-1004, Aug. 2003.
- [4] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality Image Registration by Maximization of Mutual Information," IEEE Transactions on Medical Imaging, Vol. 16, No. 2, pp. 187-198, Apr. 1997.
- [5] H. Xie, L. E. Pierce, and F. T. Ulaby, "Mutual Information based Registration of SAR images," in Proc. of IEEE International Geoscience and Remote Sensing Symposium, Vol. 6,

- pp.4028-4031, July. 2003.
- [6] Zeger F. Knops, J. B. A. Maintz, M. A. Viergever and J. P. W. Pluim, "Registration using Segment Intensity Remapping and Mutual Information," International Conference on Medical Image Computing Computer-Assisted Intervention, pp. 805-812, Sep. 2004.
- [7] Y. S. Kim, J. H. Lee, and J. B. Ra, "Multi-sensor Image Registration based on Intensity and Edge Orientation Information," Pattern Recognition, Vol. 41, pp. 3356 - 3365, Apr. 2008.
- [8] J. H. Lee, Y. S. Kim, et al, "Robust CCD and IR Image Registration Using Gradient-Based Statistical Information," IEEE Signal Processing Letters, Vol. 17, No. 4, Apr. 2010.
- [9] M. Irani and P. Anandan, "Robust Multi-Sensor Image Alignment", Proceedings of IEEE International Conference on Computer Vision, pp. 959-966, Jan. 1998.
- [10] W. C. Choi, S. N. Choi, et al, "Automatic Image Registration Based on Extraction of Corresponding-Points for Multi-Sensor Image Fusion," Journal of the Korea Institute of Military Science and Technology, Vol. 12, No. 4, pp. 524-531, Aug. 2009.
- [11] Juan Kang, et al, "Image Registration Based on Harris corner and Mutual Information," International Conference on Electronic & Mechanical Engineering and Information Technology, Aug. 2011.
- [12] C. Harris and M. J. Stephens, "A Combined Corner And Edge Detector," Alvey Vision Conference, pp.147-151, 1988.
- [13] Fischler M A, Bolles R C, "Random Sample Consensus: A Paradigm For Model Fitting With Applications To Image Analysis And Automated Cartography," Communications of the ACM, pp. 381-395, June. 1981.
- [14] O. Chum, J. Matas, and J. Kittler, "Locally Optimized RANSAC," Proc. Ann. Symp. German Assoc. for Pattern Recognition (DAGM '03), Sep. 2003.
- [15] S. M. Smith, J. M. Brady, "SUSAN-A New Approach to Low Level Image Processing," International Journal of Computer Vision 23, Vo. 23, No. 1, pp. 45-78, Nov. 1997.
- [16] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60, No. 2, Oct. 2004.
- [17] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing-3/E, PEARSON, 2009.
- [18] H. Farid, and E. P. Simoncelli, "Differentiation of Discrete Multidimensional Signals," IEEE Transactions on Image Processing, Vol. 13, No. 4, Apr. 2004.
- [19] J. P. W. Pluim, J. B. A. Maintz, M. A. Viergever, "Image Registration by Maximization of Combined Mutual Information and Gradient Information," IEEE Transactions on Medical Imaging Vol. 19, pp. 809-814, Feb. 2000.
- [20] C.R. Maurer Jr., R.J. Maciunas, J.M. Fitzpatrick, "Registration of Head CT Images to Physical Space using a Weighted Combination of Points and Surfaces," IEEE Transactions on Medical Imaging, Vol. 17, pp. 753-761, Oct. 1998
- [21] J. Canny, "A Computational Approach to Edge Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, pp. 679-698, Nov. 1986

저자 소개



주재웅

2011: 고려대학교 전기전자전파공학부 공학사
 현 재: 고려대학교 전기전자전파공학과 석박
 사통합과정
 관심분야: 영상신호처리
 Email : jyju@ispl.korea.ac.kr



김민재

2007: 고려대학교 전기전자전파공학부 공학사
 현 재: 고려대학교 전기전자전파공학과 석박
 사통합과정
 관심분야: 영상신호처리
 Email : mjkim@ispl.korea.ac.kr



구 본 화

2000: 광운대학교 제어계측공학과 공학사
2002: 고려대학교 전자공학과 공학석사
2008: 고려대학교 영상정보처리학과 공학박사
현 재: 고려대학교 연구 교수
관심분야: 표적 탐지/추적, 데이터 융합
Email : bhku@ispl.korea.ac.kr



고 한 석

1982: Carnegie-Mellon University 전기
공학 공학사
1988: Johns Hopkins University 전자
공학 공학석사
1992: Catholic Univ. of America 전자
공학 공학박사
현 재: 고려대학교 전기전자전파공학부 교수
관심분야: 영상 및 음성 신호처리, 패턴 인식
Email : hsko@korea.ac.kr