

통계적 Correlation을 이용한 다중센서 영상 정합

고진신⁰ 박영태
 삼성탈레스㈜, 경희대학교
 jinshin.ko@samsung.net , ytpark@khu.ac.kr

Multi-Sensor Image Alignment By Statistical Correlation

JinShin Go⁰ YoungTae Park
 SAMSUNG THALES CO., LTD.
 Dept. of Electronic Engineering, Kyunghee University

요 약

현재 많이 연구되는 영상융합(Image fusion)에서는 필히 두 영상의 정합(alignment)이 이루어져야만 수행 된다. 각기 다른 특징을 갖는 센서(E0,IR,Radar등)로부터 얻는 영상에서는 각각 다른 특징점 정보를 가지므로, 특징점을 이용한 영상 정합 구현에는 전처리 과정이 매우 복잡하고 까다롭게 이루어져야 한다. 본 논문에서는 Correlation에 대한 통계적 상관 관계를 이용하여, 전처리 과정을 단순하게 수행 하여도 매우 강력한 영상 정합이 이루어지도록 구현 하였다. 또한, 통계적 기법에 적합하도록, 효율적인 전처리 과정을 통해 계산량이 적어 지는 방법을 제안 한다.

1. 서 론

영상정합은 두 개의 영상에서 각각의 픽셀에 대응 되도록 영상을 align하는 것이다. 두 개의 영상 데이터를 융합(fusion) 하기 위해선 필히 영상간 matching 작업이 우선 수행 되어져야 하므로 영상 융합(image fusion)을 위한 영상 정합에 관한 많은 연구가 이루어지고 있다.

각각 다른 특징을 갖는 센서(E0,IR,Radar,etc)로부터 얻은 영상에서 각 픽셀에 일치되는 밝기 값의 상관 관계는 대부분 매우 복잡하거나 계산하기 힘들다. 표1과 같이 각 파장에 따른 photon flux level이 차이가 많이 나서 각각의 센서에 따른 영상이 포함하고 있는 정보는 다양한 변수를 가질 수 밖에 없다. 즉, 가시광선 대역의 센서에서는 존재하는 feature에 대해 적외선 영역 센서에서는 존재하지 않는다. 또 반대로 적외선 영역 센서에서 존재하는 것이 가시광선 대역의 센서에서는 존재하지 않거나 미약한 신호로 나타난다. 각 다른 센서에서 얻은 두개의 영상은 각각 영상의 특징점에 대한 연관성을 알기가 난해함을 알 수 있다. 실제로 각각의 영상에 대한 특징점에 대한 정합 작업은 전 처리과정에 많은 단계의 작업을 수행한 후 영상을 registration을 할 수 있다.

본 논문에서는 각 센서의 영상에서 특징점을 찾지 않고, 상호간 영상에 대해 correlation에 대한 통계적 상관 관계를 구하여 각 영상 간의 matching을 구현하는데 있어 작업수행 단계를 대폭 줄이는 효율적인 방법을 제안 및 구현 하고자 한다.

표 1. Photon flux level

	VIS	NIR	MWIR	LWIR
파장대(um)	0.4-0.78	0.78-1.0	3-5	8-12
밝은 대낮	1.5×10^{17}	1×10^{17}	2×10^{13}	2×10^{17}
달빛하의 밤	1.5×10^{11}	-	2×10^{13}	2×10^{17}
별빛하의 밤	1.5×10^9	9×10^9	2×10^{13}	2×10^{17}

2. Motion Model

각 센서로부터 약간의 변위차이를 가지는 두개의 영상을 가질 때, 두 영상간의 displacement field는 2D parametric transformation의 term으로 model 되어 진다. 이 transformation 의 motion vector

$\vec{d}(x, y) = (u(x, y), v(x, y))^T$ 는 식(1) 과 같이 표현 된다

$$\vec{d}(x, y; \vec{p}) = X(x, y) \cdot \vec{p} \quad (1)$$

여기서, $X(x, y)$ 는 각 픽셀 (x, y) 에 상관되는 transformation matrix이다. 그리고 $\vec{p} = (p_1, \dots, p_n)^T$ 인 파라미터 벡터이다. 이것의 예로, 선형 affine transformation은 식 (2)와같다.

$$\begin{bmatrix} u(x, y; \vec{p}) \\ v(x, y; \vec{p}) \end{bmatrix} = \begin{bmatrix} p_1 + p_2x + p_3y \\ p_4 + p_5x + p_6y \end{bmatrix} \quad (2)$$

여기서, $\vec{p} = (p_1, p_2, p_3, p_4, p_5, p_6)^T$,

$$X = \begin{bmatrix} 1 & x & y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x & y \end{bmatrix} \quad (3)$$

2차 transformation으로 표현하면 식(4) 와 같다.

$$\begin{bmatrix} u(x,y;\vec{p}) \\ v(x,y;\vec{p}) \end{bmatrix} = \begin{bmatrix} p_1 + p_2x + p_3y + p_7x^2 + p_8xy \\ p_4 + p_5x + p_6y + p_7xy + p_8x^2 \end{bmatrix} \quad (4)$$

여기서, $\vec{p} = (p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8)^T$

$$X = \begin{bmatrix} 1 & x & y & 0 & 0 & 0 & x^2 & xy \\ 0 & 0 & 0 & 1 & x & y & xy & y^2 \end{bmatrix} \quad (5)$$

3. Normalized - Correlation

두 개의 image간의 normalized correlation은 영상 화소값의 평균값과 contrast의 차이에도 일관성을 가진다. 즉, 두 개의 신호가 linear한 상관관계를 가졌다면 그들의 normalized correlation은 1이다. linear한 상관관계를 가지지 않았다면, 두 개의 영상은 공간적인 변화가 있다고 할 수 있다. Normalized Correlation값을 통계적인 예측 방법을 이용해, 영상간의 상대적인 변위차이를 예측하여, 영상의 align에 이용 할 수가 있다.

$$f(x,y) \circledast g(x+u,y+v) = \frac{\sum_x \sum_y f(x,y)g(x-u,y-v)}{\sqrt{\sum_x \sum_y g^2(x-u,y-v)}} \quad (6)$$

4. Image Representation

Multi Resolution Alignment는 안정된 Matching을 하기 위해서는 모든 resolution level에서 상호 correlated 정보를 충분히 가져야 한다. 이것은 같은 센서에서 얻어진 영상은 이 조건을 충족하지만, Muti Sensor image에서는 그렇지 못하다. 일반적으로 Muti Sensor image에서는 고주파 공간의 image 성분이 비교적 양호한 Relation을 가진다. Edge 성분은 image의 사물의 외형, 윤곽 등의 정보를 가지므로, Muti Sensor image에서 동일 정보를 가질 수 있다. 그러나, 저주파 image 성분은 대부분, 조명, 센서의 특성 등의 주변환경에 좌우되는 image 성분이 많다. 따라서, Multi Sensor의 Alignment에서는 Laplacian Filter를 통과한 고주파 성분의 image을 근간으로 상호 correlation을 획득해야 한다. 그러나 Laplacian은 각 방향에 따라 정보를 달리 가지므로 상하좌우 방향의 edge를 취하여 일반적으로 전 작업을 수행 해야 한다. 본 논문에서는 각 방향 edge 대신에 Threshold 기법을 적용하여 보다 많은 정보를 하나의 Edge 영상에 포함하게 하여 통계적인 기법을 적용하는데 있어 효율성을 더욱 증대 시켰다.

Pyramid Image는 High Resolution signal 정보를 Low Resolution으로 투영시키는 구조이다.

이것은 Coarse to Fine alignment 을 가능 하도록 하여, 더욱 효과적인 alignment을 구현할 수 있도록 한다.

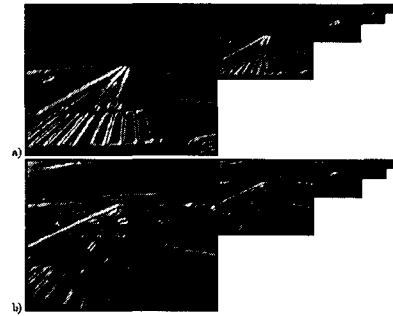


그림1.우하 방향(\)의 Edge 추출 Pyramid Image (a) EO (b) IR

5. 통계적 기법을 이용한 Alignment

주어진 두 개의 이미지 f, g 에서 normalized correlation 값이 최대가 되는 이동 vector \vec{p} 를 찾는 것이 이미지 alignment 목표 이다. 이미지 f_i 상대 적인 이미지 g_i 의 상대변위 (u,v) 를 가지는 $S_i^{(x,y)}(u,v)$ 는 식(7)과 같다.

$$S_i^{(x,y)}(u,v) = f_i(x,y) \circledast g(x+u,y+v) \quad (7)$$

\circledast : normalized correlation.

매개변수 이동 벡터 \vec{p} 에 표현되어지는 motion field를 $u = [u(x,y;\vec{p}), v(x,y;\vec{p})]$ 로 가정한다. 여기서 u 는 이동 변위, \vec{p} 는 (u,v) 의 이동 값을 가진 vector 이다. 이미지 registration 최종 목표 인 $S_i^{(x,y)}$ 가 최대인 \vec{p} vector인 것을 $M(\vec{p})$ 라 하면, $M(\vec{p})$ 는 식(8)과 같다.

$$M(\vec{p}) = \sum_{xy} \sum_i S_i^{(x,y)} [u(x,y;\vec{p}), v(x,y;\vec{p})] = \sum_{xy} \sum_i S_i^{(x,y)} [u(x,y;\vec{p})] \quad (8)$$

$M(\vec{p})$ 가 최대가 되는 \vec{p} 을 찾기 위해 통계적 기법인 Newton 방법을 적용하면, 임의의 \vec{p}_0 에 대한 $M(\vec{p})$ 의 방정식은 식 (9)와 같이 표현되어 진다.

$$M(\vec{p}) = M(\vec{p}_0) + [\nabla_{\vec{p}} M(\vec{p}_0)]^T \delta_p^p + \delta_p^{p^T} [H_M(\vec{p}_0)] \delta_p^p \quad (9)$$

여기서, $\delta_p^p = \vec{p} - \vec{p}_0$ 로 정의되어진다

$\nabla_{\vec{p}} M(\vec{p}_0)$ 는 임의의 \vec{p}_0 의 계산되어지는 $M(\vec{p})$ 의 Gradient 이고, $[H_M(\vec{p}_0)]$ 는 $M(\vec{p})$ 의 Hessian 이다.

δ_p^p 는 근사치적인 Newton 방식을 적용하여, 식 (10)과 같이 적용 할 수 있다.

$$\delta_p^* = - [H_M(\beta_0)]^{-1} \cdot \nabla_{\beta} M(\beta_0) \quad (10)$$

이것에 Chan-rule를 적용하기 위해서, 각각의 항을 표현하면 아래 식 (11) 과 같다.

$$\nabla_{\beta} M(\beta_0) = \sum_{xyi} \nabla_{\beta} S_i(u) = \sum_{xyi} [X^T \nabla_{\beta} S_i(u)]$$

$$H_M(\beta) = \sum_{xyi} [X^T H_{S_i}(u) X] \quad (11)$$

따라서, 식 (10)에 식(11)을 적용하여, 최종 δ_p^* 를 수식화 하면 식 (12) 와 같이 나타난다.

$$\beta = \sum_{xyi} [X^T H_{S_i}(u) X]^{-1} \sum_{xyi} [X^T \nabla_{\beta} S_i(u)] \quad (12)$$

이것은 Correlation항인 $S_i^{(x,y)}$ 인 것을 δ_p^* 로 재 정립하여, Correlation항이 최대인 β 을 획득할 수 있게 된다.

6. 구현 및 결과

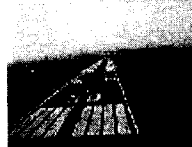
본 논문에서 상하좌우 방향의 edge성분에 대한 각 correlation 상관관계를 구하는 대신에 각 센서 영상에 적합한 Threshold 기법을 적용하여 한 방향 edge 성분에 서로 Dominate edge 성분뿐만 아니라, 다른 많은 edge를 포함 할 수 있도록 보장하여, 알고리즘을 구현 하였다.

[프로그램 Procedure]

1. 임의의 β_0 vector와 X matrix를 설정한다
2. 두 영상의 Normalized Correlation을 획득한다
3. $\nabla_{\beta} S_i(u)$ 와 $H_{S_i}(u)$ 를 구한다.
4. 식 (12)의 방정식 수행으로 β_0 를 구한다
5. β_0 를 Update : $\beta_0 := \beta_0 + \delta_p^*$ 및 영상 이동
6. Correlation이 최대가 될 때까지 2항부터 계속 반복



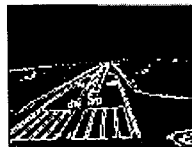
(a) IR 영상



(b) EO 영상



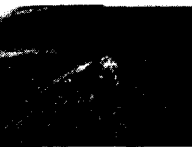
(c) IR Edge (Tr=18)



(d) EO Edge (Tr=24)

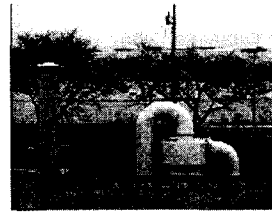


(e) Align 전 EO+IR 영상



(f) Align 후 EO+IR 영상

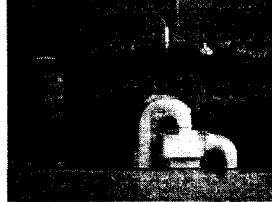
그림2. 활주로 EO, IR Image Alignment



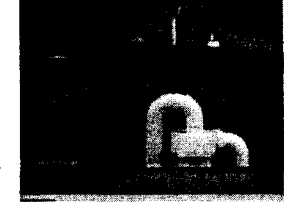
(a) CCD 영상



(b) IR 영상



(c) Align전 CCD + IR



(d) Align후 CCD+IR

그림3. 공장 구조물 CCD, IR 영상 Alignment

7. 결론

각각 다른 특징을 가지는 센서의 영상에서는 같은 특징점을 찾기가 어렵고, 이것을 이용한 영상 align를 하기 위해선 전처리 과정이 매우 복잡하고, 계산이 어렵다.

본 논문에서는 상호 영상의 특징점을 이용하지 않고 각 영상의 correlation에 대한 통계적 상관관계를 이용하여 강건한 영상 Align이 되도록 구현을 하였다. 또한 복잡한 전처리 과정이 필요 없이 다량의 edge 성분을 많이 포함 될 수 있도록 하여 통계적인 상관관계를 충실하게 만들어 수행하였다. 따라서, 각 센서 영상에 맞는 최적의 Threshold를 적용하여, Edge 성분의 네 방향성에 대해 전부 계산을 하지 않고도 정확한 align를 구현함으로써 계산량을 대폭 줄이는 방법을 제안한다.

8. 참고 문헌

- [1] M. Irani and P. Anadnan, Robust Multi-Sensor Image Alignment., *IEEE International Conference on Computer Vision (ICCV), India, January 1998.*
- [2] Paul R. Beaudet. Rotationally invariant image Operator. *In International Conference On Pattern Recognition*, pages 579-583,1978
- [3] P.J. Burt. Smart sensing with a pyramid vision machine. *Proceedings of the IEEE*, 76:1006-1015, 1988
- [4] Y. Caspi and M. Irani, A step towards sequence-to-sequence alignment. *In IEEE conference on Computer Vision and Pattern Recognition*, Pages 682-689, 2000
- [5] David L. Hall, James Llinas , An Introduction to Multisensor Data Fusion., *Proceeding of the IEEE, VOL. 85. 1, January 1997*