

2 차원 자동 변형 템플릿을 사용하는 영상인식에 대한 연구

한영모*

*한양사이버대학교 컴퓨터공학과

e-mail : ymhan123@hanmail.net

A Study on Image Recognition using 2D Auto-tuning Template

Youngmo Han*

*Dept. of Computer Engineering, Hanyang Cyber University

요 약

템플릿 정합을 사용하는 영상인식 방법은 사용이 편리한 장점이 있지만, 템플릿과 정합 영상의 자세가 맞지 않으면 좋은 결과를 기대하기 어렵다. 이와 같은 단점을 보완하기 위해서 본 논문에서는 템플릿과 정합 영상 사이에 2 차원 방향과 크기에 차이가 있어도 성능이 유지될 수 있는 방안을 제시한다. 사용의 편의성을 고려하여, 템플릿 이외의 추가정보, 예를 들어 직교투영상의 거리 정보가 없어도 사용 가능하도록 알고리즘을 설계하는데 주력한다.

1. 서론

인공위성 영상, 의학 영상, 가정용 로봇 비전의 영상 등을 분석하는데 영상 인식이 매우 유용하게 사용되고 있다[1]. 영상인식은 일반적으로 시간이 많이 걸리는 것이 난제인데, 이를 해결하기 위한 많은 연구가 진행되고 있다. 그 중에 한 방법이 VLSI 구현을 통해 속도를 향상시키는 하드웨어적인 방법이다. 다른 영상인식 알고리즘에 비해 템플릿 정합은 VLSI 로 구현하기 편리한 단순하고 규칙적인 알고리즘 모듈의 반복 형태이다. 이 때문에 템플릿 정합방법이 VLSI 구현을 통한 영상인식의 속도 향상 목적으로 많이 사용되고 있다[1-3].

템플릿 정합방법의 단점은 템플릿과 정합 영상의 자세가 맞지 않으면 좋은 결과를 기대하기 어렵다는 것이다. 이와 같은 단점을 보완하기 위해서, 본 논문에서는 2 차원 방향 차이와 크기 차이에 대하여 정합 영상에 맞도록 템플릿의 방향과 크기를 자동적으로 변형하여 정합을 수행하는 방법을 제시한다. 3 차원 모델을 활용하는 기존의 영상정합(image registration)과 달리, 본 논문의 방법은 사용의 편의성을 고려하여 템플릿 영상 이외의 추가의 정보(템플릿 영상이 직교투영 카메라로 촬영 되었다고 가정할 때, 거리 정보 등)을 필요로 하지 않도록 알고리즘을 설계한다.

2. 제안하는 템플릿 정합 알고리즘

먼저 기존의 템플릿 정합 방법을 살펴보자. M 개의 영상점 q_i 에서의 영상 밝기가 $E(q_i)$ 인 모델 템플릿이 주어졌다고 하자. 이 모델 템플릿을 N 개의 영상점

q'_i 에서 영상 밝기가 $E'(q'_i)$ 로 이루어진 영상에 정합(matching)하려고 한다. 그러면 기존의 템플릿 정합은 식(1)에 주어진 매칭 척도를 최소화 하는 영상의 평행이동 변위 d 를 찾으려 한다.

$$J(d) = \sum_{i=1}^M \|E'(q_i + d) - E(q_i)\|^2 \quad (1)$$

이러한 기존의 템플릿 정합 방법은 주어진 영상의 대상체가 템플릿과 크기와 각도가 일치하고, 단지 평행이동만 한다는 가정을 사용하고 있다. 따라서, 주어진 영상의 대상체가 크기나 각도 등이 템플릿과 다른 경우에는 이러한 기존의 방법으로는 템플릿과 같은 물체로 인식하기 어렵고, 정합이 되었다 해도 템플릿 모델 오차가 정합 오차로 이어지게 된다.

이 문제를 해결하기 위해서, 본 논문에서 제안하는 방식은 주어진 영상의 대상체와 템플릿 간의 오차를 보정하는 템플릿 변형 Δq_i 를 포함하도록 식(1)의 정합 척도를 식(2)과 같이 수정하는 것이다.

$$J(d, \Delta q_i) = \sum_{i=1}^M \|E'(q'_i) - E(q_i)\|^2 \quad (2)$$

여기서, $q'_i = q_i + d + \Delta q_i$ 이다. d 와 Δq_i 는 각각 주어진 영상의 대상체와 템플릿 간의 평행이동 변위와 모양 변형 변위를 나타낸다. 그러면 식(2)을 최소화 하는 $(d, \Delta q_i)$ 쌍에서, d 가 바로 정합 위치가 된다.

$\|\Delta q_i\| \ll \|q_i + d\|$ 인 경우에, Taylor 급수의 첫 번째 근사를 사용하여 식(3)과 같이 쓸 수 있다.

$$E'(q_i + \Delta q_i + \bar{d}) \approx E'(q_i + \bar{d}) + (\partial E' / \partial q_i)_{q_i + \bar{d}}^T \Delta q_i \quad (3)$$

여기서, $\partial E' / \partial q_i$ 를 열(column) 벡터로 정의한다. 식(3)을

사용하면, 식(2)은 식(4)와 같이 근사화된다.

$$J(\bar{d}, \Delta q_i) = \sum_{i=1}^M \left\| \{E'(q_i + \bar{d}) + (\partial E / \partial q_i)^T_{q_i + \bar{d}} \Delta q_i\} - E(q_i) \right\|^2 \quad (4)$$

주어진 3 차원 모델을 직교 투영 모델 카메라로 촬영하여, 템플릿 영상을 얻었고, 후에 이 3 차원 모델에 대한 영상들(정합 영상)을 얻었다고 생각하자. 템플릿의 각 영상점 q_i 들의 중심점을 q_c 라고 하고, 정합 영상의 각 영상점 q'_i (식(2)를 살펴보면, $q'_i = q_i + d + \Delta q_i$ 의 형태로 표현됨을 알 수 있다)의 중심점을 q'_c 라고 하자. 그러면 템플릿과 정합영상 사이의 평행 변위는 $d = q'_c - q_c$ 로 표현될 것이다. 만일, 두 영상 사이에 평행 변위뿐만 아니라 방향의 변위도 있는 경우를 생각해 보자. 분석을 위해서, 템플릿 영상의 직선 $\vec{l}_i = q_i - q_c$ 과 정합영상의 직선 $\vec{l}'_i = q'_i - q'_c$ 을 생각해 보자. 그리고 두 직선 사이의 각도 차이가 θ 로 표현된다고 생각해 보자 (경사 투영 모델 카메라로 촬영된 영상이라면 이 두 직선들의 각도 차이는 각각 다를 것이다. 하지만 본 논문에서 가정한 것과 같이 직교 투영 모델 카메라 하에서 영상 평면상의 2 차원 회전각들은 모두 같다고 생각할 수 있다.) 그러면 물리적인 배치를 고려해 볼 때, 직선 $\vec{l}_i = q_i - q_c$ 의 회전 후의 값과 직선 $\vec{l}'_i = q'_i - q'_c$ 은 서로 평행함을 알 수 있다. 따라서, 다음의 관계가 성립한다.

$$k \vec{l}'_i = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \vec{l}_i, \quad k \in R \quad (5)$$

그런데 이 식의 양변의 크기는 같아야 항등식이 성립하므로 k 값을 다음과 같이 얻을 수 있다.

$$k = \|\vec{l}'_i\| / \|\vec{l}_i\| \quad (6)$$

$\Delta l_i \in R$ 의 값을 $\Delta l_i = l'_i - l_i$ 의 형태로 정의하고, $\Delta l_1 \approx \Delta l_2 \approx \dots \approx \Delta l$ 이며 $\|\Delta l / l\| \ll 1$ 라고 가정해 보자. 그리고 $\|\theta\| \ll 1$ [rad] 이라고 가정해 보자. 그러면, $q'_i = q_i + d + \Delta q_i$ 의 관계와 식(1), 식(2)을 사용하여, 다음의 결과를 얻을 수 있다.

$$\Delta q_i = q'_i - q_i - d = +F_i \phi \quad (7)$$

여기서, $F_i = \begin{bmatrix} 1 \\ l_i \end{bmatrix} (q_i - q_c) \begin{bmatrix} -(q_{iy} - q_{cy}) \\ (q_{ix} - q_{cx}) \end{bmatrix}$, $\phi = \begin{bmatrix} \Delta l \\ \theta \end{bmatrix}$,
 $l_i = \|q_i - q_c\|$, $d = q'_c - q_c$

그러면, 식(4)에 주어진 $J(\bar{d}, \Delta q_i)$ 의 Δq_i 에 대한 최소화 문제는 식(8)에 주어진 $J(\bar{d}, \phi)$ 의 ϕ 에 대한 최소화 문제로 고쳐 쓸 수 있다.

$$J(\bar{d}, \phi) = \sum_{i=1}^M \|a_i + B_i \phi\|^2 \quad (8)$$

여기서, $a_i = E'(q_i) - E(q_i + \bar{d})$ 이고, $B_i = (\partial E / \partial q_i)^T_{q_i + \bar{d}} F_i$ 이다.

$J(\bar{d}, \phi)$ 를 최소화 하는 ϕ 값은 $\phi = (\sum_{i=1}^M B_i^T B_i)^{-1} (\sum_{i=1}^M B_i^T a_i)$

이므로, 제안하는 템플릿 정합 방법은 식(9)에 주어진 템플릿 정합 척도를 최소화 하는 $d \in R^2$ 를 찾으면 된다.

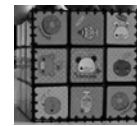
$$J(d) = \sum_{i=1}^M \|E'(q'_i) - E(q_i)\|^2 \quad (9)$$

여기서, $q'_i = q_i + d + \Delta q_i$, $\Delta q_i = \bar{d} + F_i \bar{\phi}$, F_i 와 \bar{d} 는 식(7)에 주어졌고, $\bar{\phi} = (\sum_{i=1}^M B_i^T B_i)^{-1} (\sum_{i=1}^M B_i^T a_i)$ 이다.

결과적으로 식(1)에서 기술된 기존의 템플릿 정합 방법과 같이 단일 변위 d 에 대한 정합 문제로 귀결되므로, 기존의 템플릿 정합 과정과 같은 방법으로 d 값을 찾을 수 있다.

3. 결과

본 실험에 사용된 템플릿이 그림 1에 주어져 있다.

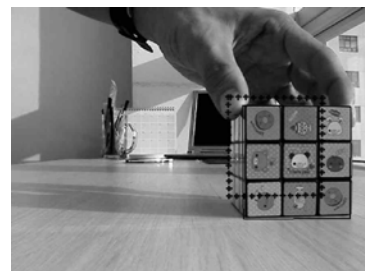


(그림 1) 템플릿

본 논문에서 제안한 템플릿 정합 방법을 사용하여 그림 2와 그림 3에 주어져 있는 정합 영상의 예에서 그림 1에 주어진 큐브를 찾는 실험을 진행한다. 정합 영상들에서 실험 조건을 통제하기 위해서 (정합 영상의 크기와 각도를 정확하게 통제하여 결과를 분석하기 위해서), 상용화된 사진 편집 도구를 이용하여 정합 영상의 방향과 크기를 인위적으로 조정해가며 실험을 진행하였다. 그림 2와 그림 3은 각각 큐브의 시계 방향 10도 회전, 큐브의 10% 크기 축소(거리가 멀어짐)를 포함한 영상이다.



(그림 2) 시계 방향 10도 방향 차이



(그림 3) 10% 작은 크기 차이

각 영상들에서 템플릿 정합을 통해 큐브의 위치를 구하였다. 제안하는 방법의 결과는 직선으로 표현하

였고, 기존 방법의 결과는 십자선으로 표현하였다. 그림 2~그림 5의 결과를 살펴보면, 일반 템플릿 정합은 템플릿의 시계 방향 회전, 크기 확대, 크기 축소를 반영하지 못하는데 비해, 제안하는 방법은 주어진 영상에서의 큐브 형태에 맞게 템플릿을 변형함으로써 좀 더 정확한 정합 결과를 보여주고 있다.

참고문헌

- [1] K. Ahuja, and P. Tuli, "Object recognition by template matching using correlations and phase angle method", *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 3, pp. 1368-1373, 2013.
- [2] T. Mahalakshmi, and R. Muthaiah, "VLSI implementation of an efficient template matching architecture based on feature extraction", *Journal of Theoretical and Applied Information Technology*, vol. 38, no. 2, pp. 191-195, 2012.
- [3] M. Ikebe, and T. Asai, "A digital vision chip for early feature extraction with rotated template-matching CA", *Journal of Robotics and Mechatronics*, vol. 17, no. 4, pp. 372-377, 2005.
- [4] N. Ranganathan, and S. Venugopal, "A VLSI chip for template matching", *Proceedings, IEEE International Conference on VLSI in Computers and Processors*, pp. 542-545, 1994.