

3 차원 방향 차이와 크기 차이에 강건한 템플릿 정합

한영모*

*한양사이버대학교 컴퓨터공학과
e-mail : ymhan123@hanmail.net

Template Matching robust to Differences in 3D Orientation and Size

Youngmo Han*

*Dept. of Computer Engineering, Hanyang Cyber University

요약

기존의 템플릿 정합 방법은 편리하지만, 정합 영상과 템플릿 사이의 크기와 방향 차이에 의해 성능이 많이 떨어졌다. 이러한 문제를 해결하기 위해서 본 논문에서는 정합과정 중에 크기와 방향 차이를 보정하는 방법을 설계한다. 좀더 폭넓은 사용을 위해서 3 차원 각도 차이까지 보정 가능하도록 설계하는데 주력한다.

1. 서론

영상은 분석하는 첫 번째 과정으로서, 영상에서 관심 있는 대상체를 찾아야 한다. 하지만 이 과정은 일반적으로 많은 시간을 소요하게 된다. 따라서 속도를 향상 시키려는 노력이 있었는데, 그 중에 하나가 VLSI 구현을 통한 방법이다. VLSI 구현을 통한 속도 향상은 알고리즘 개량을 통한 속도 향상과 비교가 되지 않을 만큼 큰 속도 향상을 가져올 수 있다는 점에서 주목할 만하다.

영상 인식을 위한 VLSI 는 많은 경우 템플릿 정합 영상인식 방법을 사용하는데 [1-3], 그 이유는 단순하고 규칙적인 알고리즘 모듈이 반복적으로 수행되어 VLSI 구조에 적합하기 때문이다. 이러한 취지에서 템플릿 정합 방법이 단순하지만 그 중요성이 높다고 할 수 있다.

기존 템플릿 정합 방법들은 템플릿과 정합영상에서 대상체가 같은 크기와 방향이어야만 좋은 결과를 기대할 수 있다는 한계점이 있었다. 이 문제를 극복하기 위해서, 본 논문에서는 정합영상에 맞게 템플릿을 변형하는 방법을 설계한다. 많은 실제 응용들이 3 차원 방향 차이를 포함하고 있기 때문에, 폭 넓은 응용을 위해서 본 연구에서는 3 차원 방향 차이까지 보정할 수 있도록 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 평행 이동 변위 d 를 찾으면 된다.

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

이러한 기존의 템플릿 정합 방법은 주어진 영상의 대상체가 템플릿과 크기와 각도가 일치하고, 단지 평행 이동만 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.

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

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 최소화하는 $(d, \Delta q_i)$ 쌍에서, d 가 바로 정합 위치가 된다. $\|\Delta q_i\| < \|q_i + d\|$ 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)를 사용하면, 식(2)은 식(4)와 같이 근사화된다.

$$J(\bar{d}, \Delta q_i) = \sum_{i=1}^M \| \{ E'(q_i + \bar{d}) + (\partial E' / \partial q_i)_{q_i + \bar{d}}^T \Delta q_i \} - E(q_i) \|^2 \quad (4)$$

주어진 3 차원 모델을 직교 투영 모델 카메라로 촬영하여, 템플릿 영상을 얻었고, 후에 이 3 차원 모델에

대한 영상들(정합 영상)을 얻었다고 생각하자.

다음으로, 식(4)에 포함되어 있는 Δq_i 의 값을 구해 보자. 주어진 3 차원 대상체를 카메라로 촬영하여 템플릿 영상과 정합 영상을 얻었다고 생각해 보자. 주어진 템플릿의 각 영상점 $q_i = [q_{ix}, q_{iy}]^T \in R^3$ 에 대응되는 대상체의 3 차원 위치가 $X_i = [x_i, y_i, z_i]^T \in R^3$ 이라 하자. 그리고 정합될 영상의 각 영상점 $q'_i = [q'_{ix}, q'_{iy}]^T \in R^3$ 에 대응되는 대상체의 3 차원 위치가 $X'_i = [x'_i, y'_i, z'_i]^T \in R^3$ 이라 하자. 본 논문에서는 템플릿 영상의 거리 정보인 z_i 의 값이 주어진 것으로 가정한다. 이 값은 기존의 3 차원 정보 복원 알고리즘 [4]을 사용하여 템플릿 이미지가 촬영될 때 쉽게 구할 수 있다. 그리고 템플릿의 영상 점 q_i 들의 중심점이 q_k 이고, 이에 대응되는 정합 영상의 점이 q'_k 라고 생각하자. 만일 템플릿 정합이 성공한다면, q'_k 는 정합 영상점 q'_i 들의 중심점이 될 것이다. q_k 와 q'_k 의 대상체 상의 3 차원 위치가 각각 X_k 와 X'_k 라고 생각하자. 그러면, 템플릿과 정합 영상 사이의 크기와 방향 차이는 다음과 같은 식으로 표현된다.

$$X'_i = \exp([\omega])(X_i - X_k) + X'_k \quad (5)$$

여기서, $\omega = [\omega_x, \omega_y, \omega_z]^T \in R^3$ 이고, $[\omega] = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}$ 이다.

템플릿과 정합 영상을 촬영한 카메라가 경사투영 모델로 표현된다면 식(6)과 같은 관계를 가진다.

$$q_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} s_i X_i \quad q'_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} s'_i X'_i \quad (6)$$

여기서, $s_i = 1/z_i$ 이고, $s'_i = 1/z'_i$ 이다.

만일, $\|[\omega]\| \ll 1$ [rad] (≈ 57.3 [deg])이라고 가정하면, $\exp([\omega]) \approx I + [\omega]$ 의 근사식을 얻는다. 여기서, $I \in R^{3 \times 3}$ 는 항등 행렬이다.

만일, $\Delta s_i = s'_i - s_i, i = 1, \dots, n$ 의 값이 $\Delta s_i \approx \Delta s, i = 1, \dots, n$ 와 같이 모두 비슷하고, $|\Delta s| \ll s_i$ 라고 가정하면, $s'_i \exp([\omega]) \approx (s_i + \Delta s)I + s_i[\omega]$ 와 $(s'_i / s'_k) \approx s_i / s_k$ 의 근사식을 얻는다. 이 근사식들과 관계식: $q'_k = q_k + d$, 그리고 $[\alpha]\beta = -[\beta]\alpha$ for any $\alpha, \beta \in R^3$ 라는 특성과 식(5)를 결합하면, 다음과 같은 식을 얻는다.

$$\Delta q_i = q_i' - (q_i + d) = F_i \phi + \left(\frac{s_i}{s_k} - 1 \right) d \quad (7)$$

여기서,

$$F_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} [(X_i - X_k), -s_i[X_i - X_k]] ,$$

$$s_i = 1/z_i, \quad X_i = z_i \begin{bmatrix} q_i \\ 1 \end{bmatrix}, \quad \phi = \begin{bmatrix} \Delta s \\ \omega \end{bmatrix}$$

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, $a_i = E'(q_i + d) - E(q_i) + \left(\frac{s_i}{s_k} - 1 \right) (\partial E' / \partial q_i)_{q_i+d}^T d$ 이고,

$B_i = (\partial E / \partial q_i)_{q_i+d}^T F_i$ 이다.

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

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, $q'_i = q_i + d + \Delta q_i$, Δq_i 는 식(7)에 주어져 있고, $\phi = (\sum_{i=1}^M B_i^T B_i)^{-1} (B_i^T a_i)$ 이다.

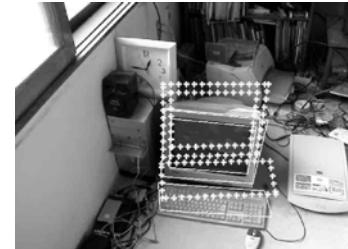
결과적으로 식(1)에서 기술된 기존의 템플릿 정합 방법과 같이 단일 벤위 d 에 대한 정합 문제로 귀결되므로, 기존의 템플릿 정합 과정과 같은 방법으로 d 값을 찾을 수 있다.

3. 결과

그림 1은 PC 의 템플릿 영상을 보여주고 있다. 그림 2~그림 5는 크기와 방향이 3 차원적으로 다른 정합 영상들의 예를 보여주고 있다.



(그림 1) 템플릿



(그림 2) 정합 영상 1



(그림 3) 정합 영상 2

제안하는 템플릿 정합 방법을 사용하여 각 정합 영상에서 PC의 위치를 구하였다. 제안하는 방법의 결과는 직선으로 표현하였고, 기존 방법의 결과는 삼자선으로 표현하였다. 그림 2~그림 5의 결과를 살펴보면, 기존 템플릿 정합은 템플릿의 3 차원 방향 차이에 의한 모양의 변화와 거리 차이에 의한 크기의 변화를 제대로 반영하지 못하여 대상체의 위치를 잘 찾지 못하고 있다. 반면에, 제안하는 방법은 주어진 영상에서의 PC 모양과 크기에 맞게 템플릿을 변형함으로써 대상체를 잘 인식하여 좀 더 정확한 정합 결과를 보여주고 있다.

참고문헌

- [1] T. Mahalakshmi, and R. Muthaiah, “VLSI implementation of an efficient template matching architecture based on feature extraction”, *Journal of Theretical and Applied Information Technology*, vol. 38, no. 2, pp. 191-195, 2012.
- [2] 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.
- [3] 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.
- [4] Y. Han, “Awareness of 3-D pose trajectory in video contents with optimal control refinement”, *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 4, issue 1, pp. 118-129, 2014.
- [5] R. M. Murray, R. M. Z. Li, and S. S. Sastry, A mathematical introduction to robotic manipulation, CRC Press, 1994.