

# Chamfer 거리 보간과 서브픽셀 탐색을 이용한 개선된 Chamfer Matching

조태훈\*

\*한국기술교육대학교 정보기술공학부

e-mail : [thcho@kut.ac.kr](mailto:thcho@kut.ac.kr)

## An Improved Chamfer Matching Using Chamfer Distance Interpolation and Subpixel Search

Tai-Hoon Cho\*

\*School of Information Technology, Korea University of Technology and Education

### 요 약

chamfer 정합(matching)은 많은 응용분야에서 사용되는 에지(edge) 기반 정합기법으로, 변환된 모델에지와 영상에지간의 거리를 최소화하는 과정이다. 이 거리는 보통 거리변환을 이용하여 픽셀분해능으로 계산된다. 본 논문에서는 서브픽셀(subpixel) 거리계산을 위해 거리계산시 보간법을 사용하여, 정확한 chamfer 정합을 구현한, 개선된 chamfer 정합 방법을 제안한다. 또한, 보다 정밀한 정합을 위해, 최소거리를 갖는 픽셀위치의 주변영역을 이용하여 최적의 위치를 추정하지 않고, 서브픽셀로 실제 거리계산을 통해 최소 거리를 찾기 위해 Powell의 최적화기법을 이용하였다. 실험결과는 제안된 방법의 타당성을 보여주었다.

### 1. 서론

한 영상에 있는 물체를 또 다른 영상에서 찾는 물체정합(object matching)은 컴퓨터 비전과 영상 해석에서 매우 중요한 문제이다. 이것은 어떤 물체를 포함하는 영상에서 이 물체의 모델(model)을 정렬시켜야 하는 산업적 응용에서는 특히 유용하다. 이 물체정합에서 얻는 변환 혹은 자세(pose)는 pick and place 동작에서 자동 자세 보정을 비롯한 다양한 작업을 위해 사용될 수 있다. (대부분, 물체의 모델은 물체의 영상으로부터 생성된다.)

물체정합은 크게 영상기반 정합과 특징기반 정합의 두 가지 방법으로 나눌 수 있다. 특징기반 정합은 영상 밝기 값을 직접 사용하는 대신 영상과 모델에서 추출한 에지(edge), 코너(corner)등의 특징(feature)을 이용한다. 에지는 가장 중요한 저수준(low-level) 특징으로, 에지기반 정합은 전형적인 산업적 응용에서 나타날 수 있는 비 균일 조명환경에서 영상기반 정합보다 강건하다. 또한, 영상에서 추출된 에지는 영상에 비해 보다 간결한 표현방법이므로, 에지기반 정합은 영상기반 정합보다 처리시간에서 일반적으로 보다 효율적이

다.

에지 기반 정합 기법 중 주목할 만한 것은 Barrow 등[1]에 의해 처음 제안되고, [2]에서 더욱 개선된 chamfer 정합(matching) 기법이다. 그 후, 많은 응용분야에서 활발히 적용되었다[3][4][5]. chamfer 정합과정은 모델에지와 영상 에지간 chamfer 거리(chamfer distance)를 최소화하는 것이다. 이 방법의 큰 단점은 변환된 모델에지와 영상에지간의 거리를 구할 때 보통 픽셀수준(pixel-level)에서 거리 변환(distance transform)을 이용하여 계산하기 때문에 정합정확도가 떨어지게 된다. 픽셀수준의 거리변환을 사용하면, 정합에 서브픽셀알고리즘을 사용해도 서브픽셀 변화에 대한 둔감함 때문에 정합정확도가 크게 개선되지 않는다.

본 논문에서는 서브픽셀 거리계산을 위해 거리계산시 보간법을 사용하여, 정확한 chamfer 정합을 구현한, 개선된 chamfer 정합 방법을 제안한다. 또한, 보다 정밀한 정합을 위해, 최소거리를 갖는 픽셀위치의 인접영역을 이용하여 최적의 위치를 추정하지 않고, 서브픽셀에서 함수 평가를 실행하여 함수의 최소값을 구할 수 있는 Powell[6]의 최적화기법을 이용하였다. 이 방법의 실효성을 보여주는 실험결과가 제시된다.

## 2. Chamfer Matching

### 2.1 개요

Chamfer 정합[2]은 2-D 물체의 에지윤곽선(edge contour)으로부터 물체를 인식하는 한 방법이다. 미리 알고 있는 물체의 에지윤곽선 영상을 "prepolygon image", 인식하려는 물체의 에지윤곽선 영상을 "predistance image"이라고 하자. 에지 윤곽선 영상은 Sobel 과 같은 에지검출 연산자에 의해서 검출된 에지 영상으로 이진 영상이다

predistance image 에서 에지가 아닌 각 픽셀(pixel)은 가장 가까운 에지픽셀(edge pixel)과의 거리가 주어진다. 에지픽셀은 0 을 갖는다. 정확한 Euclidean distance 는 계산상 시간과 메모리 자원을 크게 요하므로, 근사화가 사용된다. 이진에지영상을 근사 거리영상(distance image)으로 변환하는 것을 거리변환(distance transformation: DT)이라고 한다. 정합알고리즘에서 사용되는 DT 는 Euclidean distance 의 상당히 좋은 근사화가 되어야 한다. 그렇지 않으면, 거리값(distance values)으로부터 계산되는 정합척도(matching measure)의 분별력이 떨어지게 된다. 또한, 영상내의 전역거리(global distance)가 지역거리(local distance; 즉, 인접 pixel 간 거리)를 영상영역에 전파(propagation)함으로써 근사화 되는 것이 바람직하다.

3x3 인접영역(neighborhood)의 두 지역거리, 수직/수직인접화소간 거리와 대각선방향 인접화소간 거리로, 각각 3 과 4 를 사용하는 3-4 DT 는 Euclidean distance 와 비교하여 최대 8%의 오차만을 가져, Euclidean distance 의 훌륭한 근사화라고 할 수 있다[2]. 에지 점들은 잡음에 영향을 받으므로, Euclidean distance 가 일반적으로 꼭 필요하지는 않기 때문에, 거리영상을 구할 때, Euclidean distance 를 사용하는 것보다 3-4 DT 를 사용하는 것이 훨씬 계산을 빠르게 수행할 수 있다. 3-4 DT 보다 더 정확한 DT 계산을 위해서는 더 복잡하지만 5x5 인접영역기반의 5-7-11 DT[10] 가 사용될 수 있다.

3-4 DT 에 의한 거리영상은 순차적인 알고리즘으로 매우 효율적으로 구해질 수 있다[2]. 이진 에지영상에서, 각 에지화소는 먼저 0 으로, 에지가 아닌 각 화소는 "무한대" 값으로 초기화된 후, 두 번의 pass, "forward" (좌에서 우로, 위에서 아래로)와 "backward" (우에서 좌로, 아래에서 위로) pass 가 영상에 대해서 이루어진다.

prepolygon image 에서 에지화소들이 추출되어 (x,y) 좌표 쌍의 리스트(list)로 변환된다 (x : 행, y : 열). 이 리스트로부터 실제로 사용되는 에지점들이 선택된다. 선택된 점들의 리스트는 "polygon"이라고 한다.

polygon 이 거리영상위에 중첩될 때, polygon 의 좌표에서의 거리영상의 화소값들의 평균은 에지들간의 정합정도를 나타내며 이를 "에지거리(edge distance)"라고 한다. 두 에지들 간의 완벽한 정합은, 각 polygon 점이 거리영상의 에지화소와 일치하게 되어, 에지거리가 0 인 결과를 낼 것이다. 실제의 정합은 에지거리를 최소화하는 과정으로, 최적의 정합위치는 에지거리

가 최소화되는 위치가 된다. 정합척도로는 중간값(median), 산술평균, rms(root mean square), 최대값 등이 있다.

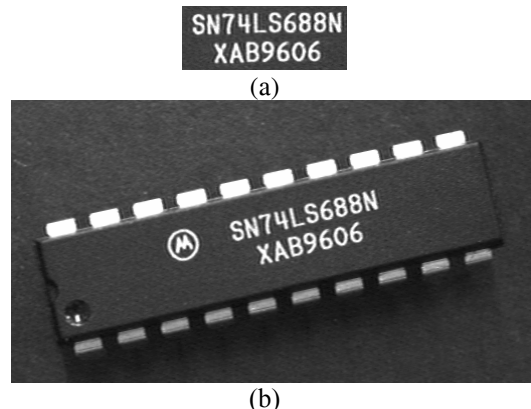
변환식에 의해 결정되는 polygon 의 각 위치에서 에지거리가 산출되어, 이 에지거리가 최소가 되는 위치가 최적정합위치가 된다. (x,y)를 polygon 좌표, (X,Y)를 거리영상의 위치좌표라 하면, 예를 들어, 이동과 회전을 포함하는 변환식은  $X = cx + x \cos\theta - y \sin\theta$ ,  $Y = cy + x \sin\theta + y \cos\theta$  이 된다. 여기서,  $\theta$  는 회전각, cx, cy 는 각각 X, Y 방향의 이동파라미터를 뜻한다. (X,Y)좌표는 보통 정수가 아니므로, 가까운 정수값으로 근사화한다.

최적의 polygon 위치를 찾는 것은 변환식의 각 파라미터가 하나의 차원에 해당하는 다차원 함수의 전역 최소점(global minimum)을 찾는 것과 같다. 따라서 지역최소점에 빠지는 것을 피하기 위해, 찾고자하는 최적 위치와 매우 가까운 지점에서 전역 최소화가 시작되어야 한다. 효율적인 탐색을 위해, 피라미드를 이용한 계층적인 구조의 알고리즘을 사용하는 것이 보편적이다. 발견한 최소 에지거리가 일정크기(보통, 1~1.5)이하이면, 모델이 탐색영상에 있다고 판단하고, 그렇지 않으면, 모델이 없다고 판단한다.

### 2.2 구현

에지를 추출하는 기법은 Canny 기법 [7]을 포함하여, 많이 있으나, 여기서는, 계산의 효율성과 우수한 성능을 갖는 Sobel 에지연산자 [8]를 적용하여, 에지크기영상과 에지각도영상을 추출하였다.

Sobel 연산자를 적용하여 얻은 에지 그래디언트(edge gradient) 크기영상을 thresholding 하여, 에지 영상을 얻으면 보통 굵은 에지들이 많이 나오게 된다. 에지 수가 증가하면, 정합속도가 느려지기 때문에, 에지 수를 줄이기 위해, 비최대점 억제(non-maxima suppression) 기법 [9]을 이용하여 에지들을 세선화 하였다. 비최대점 억제기법은 에지그래디언트 크기영상에서 그래디언트 방향 선상에서 지역최대점이 아닌 값들을 제거한다.



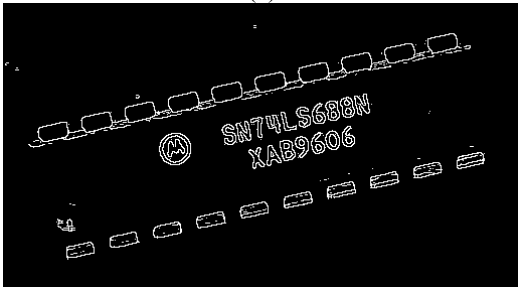
(그림 1) (a) 모델 영상 (b) 탐색영상

그림 1 은 모델영상과 탐색영상의 한 예이다, 탐색영상은 모델이 추출된 원 영상을 12 도 시계반대방향으로 회전시켜 얻은 것이다. 그림 2 은 그림 1 영상에

Sobel 에지검출과 비최대점 억제기법을 적용하여 추출된 에지들을 보여주고 있다. 그림 3 은 그림 2(b)에 3-4 DT 를 적용하여 얻은 거리영상이다. 에지픽셀은 완전히 검고, 에지로부터 멀리 떨어져 있는 픽셀일 수록 더 밝다.



(a)

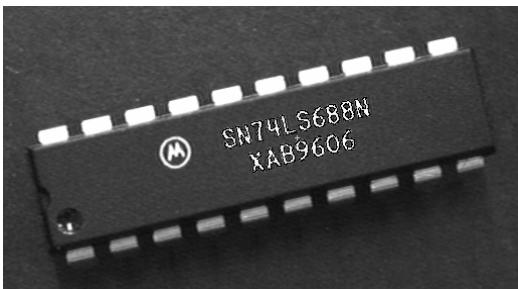


(b)

(그림 2) (a) 그림 1(a)의 에지 (b) 그림 1(b)의 에지



(그림 3) 그림 2(b)의 거리영상



(그림 4) 그림 2(a)의 모델을 이용한 정합결과

정합척도로는 중간값(median), 산술평균, rms(root mean square), 최대값 등이 있으나, rms 값이 에지거리에서 현격히 적은 오인식율을 내는 것으로 알려져 있어[2], 여기서는 rms 값을 사용하였다.

rms 값은 아래와 같이 주어진다.

$$d_{rms} = \frac{1}{C} \sqrt{\frac{1}{n} \sum_{i=1}^n d_i^2}$$

여기서,  $d_i$  는 polygon 에 있는  $n$  개의 에지점 중  $i$  번째 에지점의 거리값이고,  $C$  는 DT 에서 단위거리를 보상하기 위한 상수이다. (3-4 DT 의 경우  $C=3$  이며, 5-7-11 DT 의 경우  $C=5$  이다.)

그림 4 는 그림 2(a) 모델을 이용한 매칭결과를 보여준다. 모델에지들이 테스트영상에서 정확히 포개어 있어, 정확하게 위치와 각도를 인식한 것을 잘 나타내

주고 있다. (에지들의 위치를 검정색으로 표시하고, 중앙부근의 흰색십자표시는 기준점을 나타낸다.)

### 2.3 개선된 알고리즘

polygon 좌표  $(x,y)$ 로부터 거리영상 좌표  $(X,Y)$ 로의 변환식에 의해 결정되는 polygon 의 각 위치에 대해 에지거리를 계산할 때, 변환된 좌표  $(X,Y)$ 는 보통 정수가 아니므로, 가까운 정수값으로 반올림하여 근사화된다.

하지만, 보다 정확한 정합을 위해, 여기서는  $(X,Y)$ 에서의 거리값(distance value)을 구하는데 거리영상  $(X,Y)$  주위의 네 인접픽셀을 이용하여 거리보간하는 방법을 제안한다. 이를 위해 양선형 보간법(bilinear interpolation) 을 사용하였다. 거리영상(distance image) 을  $D$  라 하면,  $(X,Y)$ 에서의 보간된 거리값  $D_{XY}$  는 다음과 같이 계산된다.

$$D_{XY} = (1-\alpha)(1-\beta)D(\lfloor X \rfloor, \lfloor Y \rfloor) + (1-\alpha)\beta D(\lfloor X \rfloor, \lfloor Y \rfloor + 1) + \alpha(1-\beta)D(\lfloor X \rfloor + 1, \lfloor Y \rfloor) + \alpha\beta D(\lfloor X \rfloor + 1, \lfloor Y \rfloor + 1)$$

여기서,  $\alpha = X - \lfloor X \rfloor$ ,  $\beta = Y - \lfloor Y \rfloor$ , 그리고  $\lfloor X \rfloor$  는  $X$  의 내림한(truncated) 정수값을 뜻한다.

표 1 은 거리영상에서 양선형 보간법을 사용한 경우와 사용하지 않은 경우의 Chamfer 정합 결과를 보여준다. 거리보간을 사용하면, 거리값이 크게 줄어드는 것을 알 수 있다.

보다 정확한 정합파라미터를 얻기 위해, 간단한 1 차원 파라볼릭 보간법 같은 서브픽셀 알고리즘이 사용될 수 있다. 어떤 평가함수  $f(x)$ 의 값이 정수 픽셀 위치  $x=H$  에서 최대값  $f_m$  을 갖고,  $f(H-1) = f_l$ ,  $f(H+1) = f_r$  이라 하면,  $f(x)$ 가 최대가 되는  $x$  값  $H_s$  는 다음과 같이 1 차원 파라볼릭 보간법에 의해 서브픽셀 정밀도로 구할 수 있다.

$$H_s = H + \frac{f_r - f_l}{2(f_m - f_r - f_l)}$$

<표 1> chamfer matching 을 적용한 결과

	위치[pixel]	각도[도]	거리값
거리보간 하지 않은 경우	(292, 130)	12	0.537
거리보간 한 경우	(292, 130)	12	0.444

평가함수 파라미터가  $x, y$ , 각도 등 세 개라면, 위의 1 차원 보간법을 각 파라미터에 독립적으로 적용하여, 각 파라미터 값을 서브픽셀로 구한다. 하지만, 이 방법은 실제의 평가함수 값(여기서는 거리값)을 서브픽셀로 계산하지 않고, 매칭위치를 보간에 의해 추정한다. 따라서, 보다 정확한 매칭을 위해 최소 chamfer 거리 값 탐색에서 보간된 거리값을 이용하여 최적화하는 기법이 바람직하다. 이를 위해, 평가함수의 미분값 계산을 요하지 않는 Powell 의 최적화 기법 [6]이 사용될 수 있다.

표 2 는 거리보간없이 그림 4 에 서브픽셀을 적용한 결과이며, 표 3 은 거리보간과 서브픽셀을 적용한 결

과이다. 서브픽셀 알고리즘을 적용하면 거리값이 크게 줄어들음을 알 수 있다. 서브픽셀 알고리즘으로는 Powell 방법이 파라볼릭 보간법보다 나은 결과를 내는 것으로 보인다.

<표 2> 거리보간없는 chamfer 정합과 서브픽셀 알고리즘 적용 결과

서브픽셀	위치[pixel]	각도[도]	거리값
적용 전	(292, 130)	12	0.537
적용 후 (파라볼릭 보간)	(291.82, 129.80)	12.08	0.478
적용 후 (Powell)	(291.76, 129.73)	12.00	0.463

<표 3> 거리보간 chamfer 정합과 서브픽셀 알고리즘 적용 결과

서브픽셀	위치[pixel]	각도[도]	거리값
적용 전	(292, 130)	12	0.444
적용 후 (파라볼릭 보간)	(291.82, 129.79)	12.08	0.386
적용 후 (Powell)	(291.72, 129.70)	12.00	0.378

### 3. 개선된 알고리즘의 정밀도 시험

개선된 매칭 알고리즘의 정밀도를 측정하였다. 먼저, x, y 정밀도를 측정하기 위해서, 그림 1(a)을 포함하는 테스트영상(640x480)을 x, y 축 방향으로 각각 -0.8~0.8 픽셀 범위에서, 0.2 픽셀 간격으로 이동시키면서, 얻은 총 80 개의 영상에 대해 실행하여 위치오차(표준편차, 최대오차)를 구하였다. 또한, 테스트영상을 영상의 중심기준으로 +30 도 범위 내에서, 1 도 간격으로 회전시켜 얻은 60 장에 대해서도 알고리즘을 적용하여, 각도오차를 구하였다.

이 결과를 표 4 과 표 5 에 보인다. (사용된 모델영상은 그림 1(a)이고, 이동/회전시 선형보간법을 이용하였다.) 파라볼릭 보간법에 의한 서브픽셀알고리즘은, x,y 정밀도는 최대 0.14 픽셀, 각도는 0.15도 오차를 보이며, Powell의 방법은 최대 0.06 픽셀의 x 오차, 0.08 픽셀의 y 오차, 0.04 도의 각도 오차를 보인다. 거리보간을 하지 않은 경우, 서브픽셀탐색으로 Powell 방법을 사용하면 안 좋은 결과를 낳았는데(표 4), 이것은 거리보간없는 거리값계산은 x, y, 각도에 매우 작은 변화에 둔감하여, 지역적 최소점에 빠지기 쉽기 때문인 것으로 추측된다. 따라서, chamfer 거리 보간과 Powell의 서브픽셀알고리즘을 사용하는 것이 가장 정밀한 것으로 나타났다.

위의 실험에서는 3-4 DT 를 사용했으나, 보다 정확한 5-7-11 DT 를 사용해도 거의 같은 결과를 얻었다. 하지만, 응용분야에 따라서는 5-7-11 DT 가 3-4 DT 보다 더 나올 수 있을 것으로 판단된다.

<표 4> 위치 및 각도 정밀도 (chamfer 거리보간 없음)

서브픽셀 방법		x [pixel]	y [pixel]	각도 [도]
파라볼릭 보간	표준편차	0.09	0.09	0.07
	최대오차	0.14	0.12	0.19
Powell 최적화	표준편차	0.30	0.32	0.06
	최대오차	0.42	0.71	0.31

<표 5> 위치 및 각도 정밀도 (chamfer 거리보간)

서브픽셀 방법		x [pixel]	y [pixel]	각도 [도]
파라볼릭 보간	표준편차	0.09	0.09	0.07
	최대오차	0.14	0.12	0.15
Powell 최적화	표준편차	0.03	0.03	0.02
	최대오차	0.06	0.08	0.04

### 4. 결론

본 논문에서는 정밀한 chamfer 정합을 위해, 서브픽셀 거리계산이 가능하도록 거리계산시 보간법을 사용하는 개선된 방법이 제시되었다. 또한, 보다 정밀한 정합을 위해, 최소거리를 갖는 픽셀위치의 주변영역을 이용하여 최적의 위치를 추정하지 않고, 서브픽셀로 실제 거리계산을 통해 최소 거리를 찾기 위해 Powell의 최적화기법을 이용하였다. 실험 결과로 이 방법의 실효성을 보였다.

### 참고문헌

- [1] H.G. Barrow, J.M. Tenenbaum, R.C. Bolles, and H.C. Wolf, "Parametric correspondence and chamfer matching: Two new techniques for image matching," *Proc. 5th Int. Joint Conf. Artif. Intelli.*, Cambridge, MA, pp.659-663, 1977.
- [2] G. Borgefors, "Hierarchical chamfer matching: a parametric edge matching algorithm," *IEEE Trans. Pattern Analy. and Mach. Intelli.*, vol.10, no.6, pp. 849-865, 1988.
- [3] D. Chetverikov and Y. Khenokh, "Matching for Shape Defect Detection," *LNCS 1689*, pp. 367-374, 1999.
- [4] D. Gavila, "Pedestrian detection from a moving vehicle," *Proc. ECCV*, pp.37-49, 2000.
- [5] A. Thayananthan, B. Stenger, P.H.S. Torr, and R. Cipolla, "Shape context and chamfer matching in cluttered scenes," *Proc. CVPR 2003*, Madison, Wisconsin, pp.127-135, 2003.
- [6] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, 2nd Ed., Cambridge University Press, 1992.
- [7] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analy. and Mach. Intelli.*, vol.8, no.6, pp.679-698, 1986.
- [8] E.R. Davies, *Machine Vision*, 3rd Ed., Morgan Kaufmann, 2005.
- [9] R. Jain, R. Kasturi, and B.G. Schunck, *Machine Vision*, McGraw-Hill, 1995.
- [10] G. Borgefos, "Distance transformations in digital images," *Computer Vision, Graphics, and Image Processing*, vol.34, no.3, pp.344-371, 1986.