

Log-polarImage를 기반으로한 이차원 물체인식을 위한 Alignment algorithm개발

손영호*†, 유범재*,오상록*, 박귀태†

* 한국과학기술연구원(KIST) 지능제어연구센터 / † 고려대학교 전기공학과

The Development of an Alignment algorithm for the Log-polar Image-based 2D Object Recognition

Young-Ho Son*†, Bum-Jae You*, Sang-Rok Oh*, and Gwi-Tae Park†

* Intelligent System Control Research Center, KIST / † Dept. of Electrical Eng., KOREA UNIV.

Abstract - 인간의 안구는 색과 모양을 식별하는 데에 관여하는 원추 세포와 물체의 명암을 구분하는 간상 세포로 구성되어지는 시세포를 가진다. 망막 위에 분포되어 있는 시세포들은 시축을 중심으로 각기 다른 밀도로 분포되어 있다. 특히 광축과 만나는 중심 지역은 Fovea라고 하는 직경 1mm 정도의 작은 우물을 이루고 있는데 원추 세포들이 고해상도로 분포되어 있고 시신경과 일대일로 연결되어 있어 시각 처리의 중심이 된다. 특히, 글자나 물체를 인식하기 위해 인간은 대상물을 응시하여 대상물의 영상이 Fovea영역에 맺히도록 추적 운동을 계속한다. 본 논문에서는 인간의 눈과 유사한 망막 모델의 하나인 Log-polar Image를 이용한 물체 인식을 위해 물체를 Log-polar Image Plane의 중심에 위치시키기 위한 모멘텀(Momentum)기반 정합 알고리즘(Alignment Algorithm)을 제시한다. 이는 눈동자 운동이 가능한 능동형 시각 장치의 Tracking 및 Pursuit 동작 중에 발생하는 추적 오차를 보상함으로써 운동 중에도 효과적인 물체 인식이 가능하게 한다. 또한, 물체를 Log-polar Image Plane의 중심에 위치시킴으로써 물체의 위치 이동, 회전이동 그리고 크기 변화에 무관하게 물체를 인식할 수 있음을 제시한다.

1. 서 론

인간의 눈은 한가지의 물체만 주시한다. 이 사실은 인간의 생물학적인 구조에서 알 수 있다. 인간의 안구 구조는 비교적 밝은 곳에서 물체를 인식하고, 색을 식별하는 데에 관여하는 원추세포와 어두운 장소에서 물체를 보는데에 관여하는 간상세포로 구성되어지는 시세포를 가진다. 본 논문에서는 인간의 눈과 유사하게 동작할 수 있는 망막 모델의 하나인 Log-polar mapping을 사용한다 [1][2][4]. 이 방법은 중앙부로부터는 해상도가 높은 영상을 획득하는데 비해, 주변부에서는 해상도가 낮아져서, 처리해야 하는 시간을 줄이는 장점을 가진다. 그리고 Log-polar mapping은 흥미로운 특성들을 가지는 데, 이미지의 크기와 회전에 대해서 불변이다. 하지만 물체의 이미지가 Log-polar space상에서 중심에 위치하지 않는다면 이들 두 특성에 만족하지 못한다. 이것은 물체가 카메라를 통해 물체의 이차원 영상이 입력될 때 이차원 영상이 좌표계에서 중심에 위치할 확률이 거의 없기 때문에 일차 모멘트[7][8]에서 얻은 도심 정보를 이용하여 이차원 영상의 물체를 Log-polar space상의 중심으로 이동시킴으로써 Log-polar mapping이 가지는 두 특성에 만족시키고 동시에 물체의 위치 이동, 회전이동 그리고 크기 변화에 무관하게 물체를 인식할 수 있음을 제시하고자 한다.

2. 본 론

2.1 The Log-polar Mapping

우리는 (x,y)를 직교좌표계로 표현하고, (ρ,θ)를 극좌표계로 표현한다. 그러면 이차원 평면에 대해서, (x,y)는 직교좌표계에서의 한 평면이며, (ρ,θ)는 극좌표에서의 한 평면이라고 할 수 있다. 또한, 직교좌표 평면에서의 한 점은 복소 평면과 극좌표평면에서의 한 점으로 표현될 수 있다. 그 식은 다음과 같다.

$$\left. \begin{aligned} z &= x + jy = \rho e^{j\theta} \\ \rho &= \sqrt{x^2 + y^2} \\ \theta &= \text{atan2}(y, x) \end{aligned} \right\} \quad (1)$$

여기서, z는 복소수이다. Log-polar mapping은 극좌표계에서의 점들을 Log-polar 평면으로 사상시키는 것을 말한다. log-polar 식은 다음과 같다.

$$w = \ln(z) = \ln(\rho) + j\theta \quad (2)$$

여기서, z ≠ 0, 그리고 w는 Log-polar평면상에서의 복소수이다.

위 식(2)는 다음과 같이 표현될 수 있다.

$$\left. \begin{aligned} \xi &= \text{Re}(w) = \ln(\rho + \alpha) \\ \eta &= \text{Im}(w) = \theta \end{aligned} \right\} \quad (3)$$

여기서, α는 작은 값의 상수이다. 위 식에서 α가 삽입되는 이유는 무한대의 해상도를 막기 위함이다. 그리고 다음 그림 1은 입력 영상(Retinal plane) I(x,y)를 Cortical plane에 매핑하는 것을 보여준다.

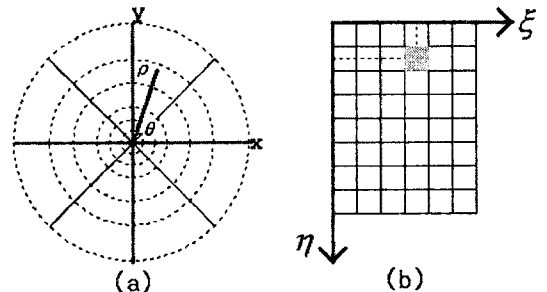


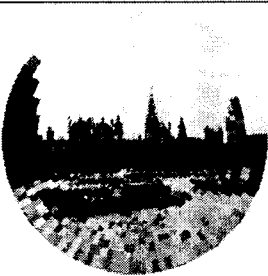
그림 1. Log-polar 변환
(a) Retinal plane, (b) Cortical Plane

2.2 Log-polar Image의 특성

카메라가 고정되어 있는 물체에 대해서 물체쪽으로 움직이거나 멀어지면, 영상된 물체의 이미지는 확대 혹은 축소가 될 것이다. 이러한 현상에 대해서 Log-polar Image는 크기에 대해서 불변이다. Log-polar Image의 크기요소는 다음과 같이 표현된다.



(a)원 이미지(256x256)



(b)재맵핑된 Log-polar 이미지(64x128)



(c)패딩된 Log-polar 이미지(64x128)

그림 2 원 이미지와 Log-polar Images

$$\left. \begin{aligned} z &= c \cdot \rho e^{j\theta} \\ w &= \ln(\rho) + \ln(c) + j\theta \end{aligned} \right\} \quad (4)$$

여기서, c 는 크기요소이다. 그리고 만약 카메라가 고정되어 있는 물체에 대해서 회전을 한다면, 영사된 물체의 이미지는 회전되어질 것이다. 이러한 현상에 대한 Log-polar Image는 회전에 대해서 불변이다. Log-polar Image의 회전각는 다음과 같이 표현된다.

$$\left. \begin{aligned} z &= \rho e^{j(\theta + \sigma)} \\ w &= \ln(\rho) + j(\theta + \sigma) \end{aligned} \right\} \quad (5)$$

여기서, σ 은 회전각이다. 다음 그림3과 4는 Log-polar Image의 특성에 대한 기하학적인 표현이다.

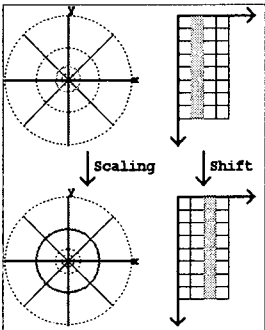


그림 3. 물체의 크기변화에 대해서 이동한 결과.

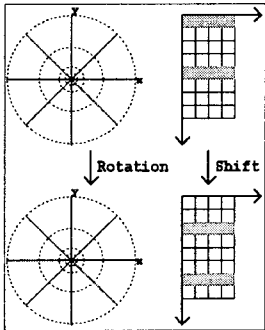


그림 4. 물체의 회전에 대해서 이동한 결과.

위에서 언급한 두 특성은 물체가 이미지 평면의 중심에 위치하였을 경우에만 해당된다. 다시 말하면 물체가 이미지 평면의 중심에 위치하지 않을 경우에는 이러한 특성을 잃어버리게 된다. 이에 대한 문제를 해결하기 위해 1차 모멘트를 이용하여, 이미지 평면의 중심에 대해서 벗어난 입력 이미지의 중심에 대한 차이만큼 이동시켜서 이미지의 크기변화와 회전변화에 대해서 무관하게 물체를 인식할 수 있도록 하고자 하는 것이다.

2.3 도심 모멘트

이차원 이미지함수 $I(x,y)$ 에 대해서, $(p+q)$ 차 모멘트 m_{pq} 는 다음과 같이 정의 된다[7][8].

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p \cdot y^q \cdot I(x,y) dx dy \quad (6)$$

위 식(6)에 대한 이산화 식은 다음과 같다.

$$m_{pq} = \sum_{x=0}^M \sum_{y=0}^N x^p y^q I(x,y) \quad (7)$$

여기서, $I(x,y)$ 는 물체는 1이고, 배경은 0으로 한다. 그리고 영차 모멘트 m_{00} 는 이미지의 총 픽셀수로 표현된다. 즉, 영차 모멘트 m_{00} 를 일차 모멘트 m_{01}, m_{10} 에 대해 나누면 이미지 함수 $I(x,y)$ 의 도심이 되는 것이다. 도심에 대한 식은 다음과 같이 표현할 수 있다.

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}} \quad (8)$$

앞에서 언급한 모멘트를 이용하여, 본 논문에서는 x - y 평면에서 이미지의 도심을 구하는 것이 아니고, Log-polar 평면에서 도심을 구하고자 한다. 그러면, Log-polar 평면에서 도심을 구하는 방법은 다음 그림을 보면서 알아보도록 한다. 그림5의(b)에서 l_1, l_2 는 그림(c)에서의 l_1, l_2 와 각각 같으므로 그림(c)에서 l_1, l_2 에 대한 각각의 선분 길이를 안다면, 그림(b)의 삼각형 면적을 구할 수 있고, 각각의 삼각형 면적을 더하면, 결과적으로, 우리는 Log-polar 평면에서 어떤 물체의 영차 모멘트를 계산할 수 있는 것이다. 또한, 이 정보를 이용하여 일차 모멘트도 구할 수 있다. 그 결과로, 우리는 Log-polar 평면의 원

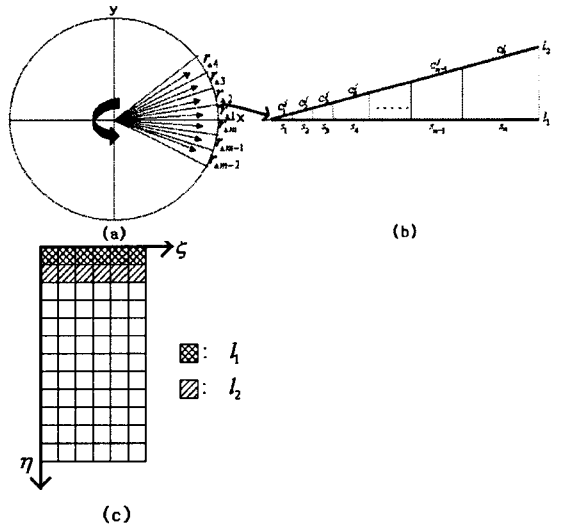


그림 5. 도심 모멘트

점에서 벗어난 물체의 도심을 찾을 수 있는 것이다. 위에서 언급된 방법의 식들은 다음과 같다.

$$S_v = \sum_{j=1}^n s_j^i I(i, j) \quad (9)$$

$$r_{vj} = \frac{\sum_{i=1}^n c_i^j s_j^i I(i, j)}{S_v} \quad (10)$$

$$S_{object} = \sum_{j=1}^m S_{Vj} \quad (11)$$

$$\xi_{Centroid} = \frac{\sqrt{\left(\sum_{j=1}^m r_{Vj} \cos \theta_{Vj}\right)^2 + \left(\sum_{j=1}^m r_{Vj} \sin \theta_{Vj}\right)^2}}{S_{object}} \quad (12)$$

$$\eta_{Centroid} = \tan^{-1} \left(\frac{\sum_{j=1}^m r_{Vj} \sin \theta_{Vj}}{\sum_{j=1}^m r_{Vj} \cos \theta_{Vj}} \right) \quad (13)$$

여기서,

$$I(i, j) = \begin{cases} 1 : Object \\ 0 : Background \end{cases}$$

$c = (r, \theta)$: 삼각형의 각 셀에 대한 도심.

$R_{centroid}$: ξ 축에 대한 물체의 도심

$\eta_{centroid}$: η 축에 대한 물체의 도심

S_{object} : 물체의 영차 모멘트

$i: 1, 2, \dots, N$ 는 ξ 축의 픽셀수

$j: 1, 2, \dots, M$ 는 η 축의 픽셀수

r_v : ξ 축에 대한 삼각형의 도심

θ_v : η 축에 대한 삼각형의 도심

S_v : 삼각형의 면적

2.3 실험 결과

앞에서 설명한 바와 같이 모멘트를 이용하여 이미지 평면의 중심에 대해서 벗어난 입력 이미지의 중심에 대한 차이만큼 이미지를 이동시키는 정합 알고리즘의 실험 결과는 그림6과 같다. 그림(a)가 x-y평면에 대해서 이미지의 중심이 x-y평면의 원점과 불일치한 이미지를 Log-polar변환에 의해서 그림(b)와 같이 변환하고 그림(b)에 대한 이미지의 영차와 일차 모멘트를 구하여 Log-polar평면의 원점에 그림(b)의 이미지 도심과 일치하게 이동시킨 것이 그림(c)이다. 그림(d)(레퍼런스 이미지)와 그림(b)에 대해서 원점으로 이동된 그림(c)와 비교해보면, 그림(c)가 (d)에 대해서 아주 유사함을 볼 수 있으므로, 앞으로 우리는 그림(c)를 이용하여 이미지의 크기변화와 회전변화에 대해서 무관하게 물체를 인식할 수 있도록 하고자 한다.

3. 결 론

우리는 모멘트를 이용하여 입력 이미지가 Log-polar 평면의 중심에 위치하지 않는 이미지를 Log-polar평면의 원점으로 얼라이먼트 시키는 과정의 결과에 대해 좋은 결과를 얻을 수 있음을 실험 결과로부터 확일 할 수 있다. 하지만, 입력이미지가 Log-polar평면의 원점으로부터 떨어진 먼 거리에서 입력이 되어진다면, 이러한 이미지에 대한 모멘트는 정확도가 떨어지고 만다. 왜냐하면, Log-polar 맵핑의 특성 때문인데, 다시 말하면, Log-polar맵핑은 Fovea의 근방에서는 해상도가 높지만, 주변부에 대해서는 해상도가 상대적으로 Fovea에서 부터 멀어질수록 해상도가 낮아지므로 이미지가 입력될 때부터 원 이미지의 정보를 잃어버리기 때문이다. 그래서 우리는 물체의 이차원 이미지를 카메라의 원점 근처에 위치시킬 필요가 있고, 또, 카메라 좌표계와 실 좌표계간의 불일치에 의한 이차원 이미지가 왜곡되는 이미지에 대해서도 무관하게 인식할 수 있는 방법을 연구할 필요가 있다고 생각한다.

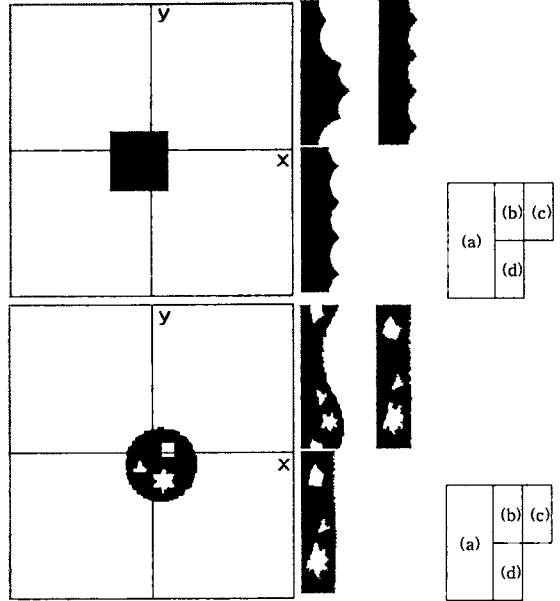


그림 6. 실험 결과
(a) x-y평면에서의 입력 이미지, (b) Log polar평면에서의 입력 이미지, (c) 이미지(b)에 대해서 원점으로 이동된 이미지 (d) 레퍼런스 이미지

[참고 문헌]

- [1] Araujo, H. and Dias, J.M., "An Introduction to the Log-Polar Mapping", Proc of the II Workshop on Cybernetic Vision, pp.139-146, 1996
- [2] J.C.Wilson and R.M. Hodgson, "Log Polar Mapping Applied To Pattern Representation And Recognition", Computer Vision and Image Processing, pp.245-277, 1992
- [3] F.L.Lim, G.A.W.West and S.Venkatesh, "use of log polar space for foveation and feature recognition", Proc of IEE Image Signal Processing, Vol. 144, No. 6, pp.323-331, 1997
- [4] Marc Bolduc and Martin D.Levine, "A Review of Biologically Motivated Space-Variant Data Reduction Models for Robotic Vision", Computer Vision and Image Understanding, Vol. 69, No. 2, pp.170-184, 1998
- [5] Herman M. Gomes and Robert B. Fisher, "Primal sketch extraction from a log polar image", Pattern Recognition Letters, Vol. 24, No. 7, pp.983-992, 2003
- [6] M. Fukumi, S. Omatu, F. takeda, and T. Kosaka, "Rotation-Invariant Neural Pattern Recognition System with Application to Coin Recognition", IEEE Transactions on Neural Networks, Vol. 3, No. 2, pp.272-280, 1992
- [7] S.Paschalakis and P.Lee, "pattern recognition in gray level images using moment based invariant features", IEE Conference on Image Processing and its Application, No 465, pp.245-249, 1999
- [8] Khotanzad.A and Hong.Y.H., "Invariant Image Recognition by Zernike Moments", IEEE Trans. on Pattern Analysis and Machine Interlligence, Vol.12, No.5, pp.489-497, 1990