

실내 환경 이미지 매칭을 위한 GMM-KL프레임워크

GMM-KL Framework for Indoor Scene Matching

Jun-Young Kim* , Han-Seok Ko**

Dept. of Electronics and Computer Engineering, Korea Univ.

Abstract - Retrieving indoor scene reference image from database using visual information is important issue in Robot Navigation. Scene matching problem in navigation robot is not easy because input image that is taken in navigation process is affinely distorted. We represent probabilistic framework for the feature matching between features in input image and features in database reference images to guarantee robust scene matching efficiency. By reconstructing probabilistic scene matching framework we get a higher precision than the existing feaure-feature matching scheme. To construct probabilistic framework we represent each image as Gaussian Mixture Model using Expectation Maximization algorithm using SIFT(Scale Invariant Feature Transform).

Key Words : Scene Recognition, Gaussian Mixture Model, Expectation Maximization, Kullbak-Leibler Distance

1. 서 론

로봇이 실내환경에서의 주행에 있어서 그림 1. 과 같이 서로 같은 환경 정보를 다른 위치에서 바라보는 상황이 발생하게 된다. 그림 1. 에서 보듯이 카메라가 같은 곳을 바라보고 있지만, 카메라의 위치에 따른 두 이미지 간의 Homography 관계는 Affine하게 바뀌게 된다. 이 때 영상정보를 이용하여 데이터 베이스상에 저장된 이미지와 현재 위치에서의 입력 영상이 서로 매치가 되는지의 여부를 검색하여 대략적인 위치 정보를 알 수 있도록 한다. 어느 정도의 Affine 왜곡된 이미지에 대해서도 매칭 성능을 보장하기 위해서는 Feature의 선택과 이미지를 매칭하는 방법이 중요성을 지니게 된다.

SIFT Feature를 사용하여 데이터 베이스를 구성하고 입력 영상에 대한 각각의 SIFT Feature와의 유클리디안 거리값으로 이미지 매칭을 하는 경우 입력 영상의 Feature들과 거리가 더 가까운 잘못 매칭되는 이미지의 Feature로 매칭되는 경우가 발생하는 단점이 존재하게 된다.

이러한 단점을 극복하기 위해 본 논문에서는 기하학적인 거리등으로 Feature 대 Feature 매칭이 아닌 이미지를 Gaussian Mixture Model (GMM)로 확률적으로 모델링하여 매칭에 이용하였다. 이러한 이미지의 확률적 모델링은 이미지 세그멘테이션, 검색에 Low-level Feature를 이용하여 사용되었다 [1].

Scene 을 모델링함에 있어 SIFT Feature를 단순화하여

GMM을 구성하는데 사용하였고, 입력 영상과 데이터 베이스에 저장된 이미지간의 유사도는 일종의 상대적 Entropy를 측정하는 Kullbak-Leibler(KL) 방법을 사용한다. 이러한 확률적 모델링 방법을 사용하여 입력 이미지가 데이터 베이스 상의

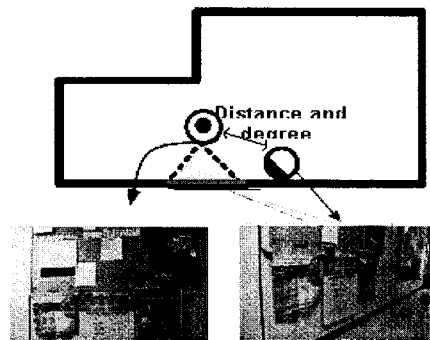


그림 1. 카메라 위치에 따른 이미지 변화

랜드마크 이미지에 대해 Affine 하게 왜곡 되었다 하더라도 강한 매치 성능을 보이도록 구성하였다. 본 논문에서는 주요한 Feature들만으로 확률적 모델링을 구성하기 위해 PCA(Principal Component Analysis) 를 이용하여SIFT Feature의 차원을 줄이는 것은 두 가지 의미를 지닌다. 첫째로 차원을 줄임으로서 계산 용량을 줄이고, 데이터 베이스에 저장해야 하는 용량을 줄일 수 있다. 둘째로 그 이미지를 가장 잘 표현하는 Feature벡터만을 쓰게 되므로 각각의 랜드마크에 대한 유일성을 높일 수 있다. 그림 2.는 위에서 설명한 실내 환경 이미지 매칭을 위해 본 논문에서 제시한 개괄도이다.

저자 소개

*高麗大學 바이오마이크로시스템기술융합과정, 碩士課程

**高麗大學 電子컴퓨터工學科 教授 · 工博

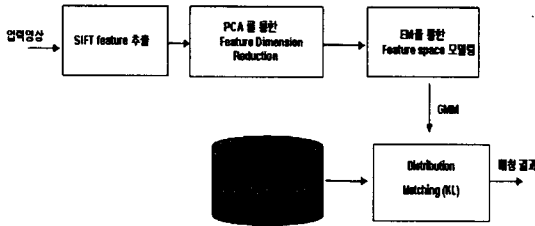


그림 2. 실내 환경 이미지 매칭 개괄도

2. 특징 추출

실내 환경 이미지는 동일한 카테고리의 이미지라고 할지라도 카메라의 위치에 따라 이미지 간의 편차가 크다. (그림 1.) 정확한 이미지 매칭을 위해서는 카메라 위치에 강인한 Feature를 뽑는 것이 중요하다. 따라서 이미지 translation, scaling, rotation, 그리고 어느 정도의 조명의 변화, affine or 3D projection 에도 강인한 Feature를 사용해야 한다. 이러한 여러 가지 조건을 만족시키면서 강인한 성능이 검증된 Feature로 D.Lowe [2]가 제안한 SIFT Feature를 사용하였다.

SIFT 키는 이미지 인텐시티로부터 가우시안 블러링을 적용 후 이러한 스케일 스페이스에서의 가우시안들의 차들로부터 최대와 최소값을 갖는 스케일을 선택한다. 여러 스케일에 대하여 스케일 스페이스를 구성하면 스케일에 강인한 Feature Point 들을 뽑을 수 있다. 키포인트 디스크립터(그림 3.) 키포인트 주위의 샘플 포인트들의 그라디언트 방향성과 크기들로부터 구해진다. 이것은 키포인트 주변의 로컬 이미지의 각 방향에 대한 크기를 구하는 것과 같다.

D.Lowe가 제안한 방법에서는 4*4 디스크립터 어레이와 8 방향의 그라디언트를 사용하여 하나의 Feature당 4*4*8=128 차원의 데이터를 사용하였다.

그러나 본 논문의 실험적 결과에 따르면 확률적 모델링을 사용하여 매칭할 경우 PCA를 통하여 32차원의 SIFT Feature를 사용하더라도 Feature-Feature매칭에 비해 더 정확한 매칭 성능을 얻을 수 있다. (그림 3.)

본 논문에서는 DB 이미지를 구성하고 각 이미지로부터 SIFT feature를 뽑은 후 PCA(Principal Component Analysis) 를 통해 각 이미지를 가장 잘 표현하는 Feature 벡터와 차원으로 Feature들의 개수와 차원을 줄인다. 이런 방식으로 여러 이미지에 대해 추가되는 Feature들의 차원을 실험적으로 고정시켜서 GMM모델링에 사용하였다.

3. 이미지 데이터 표현 및 매칭

데이터 베이스 내의 이미지와 입력 영상을 확률적으로 모델링 함에 있어 모두 가우시안 조합 모델(GMM) 으로 나타내어 진다. GMM을 구성함에 있어 Expectation Maximization(EM) 알고리즘을 이용한다. 두 이미지 간의 매칭은 확률적 분포의 유사도 판별로 볼 수 있다. 두 이미지 간의 확률적 분포의 유사도를 측정하여 매칭할 때에는 Kullbak-Leibler(KL) distance를 사용한다 [3].

3.1 이미지 데이터 모델링

데이터 베이스 상의 이미지와 입력 영상은 모두 Feature Extraction 후에 Gaussian Mixture Model로 모델링 되어진다.

이미지에서 SIFT Feature를 뽑고 PCA로 Principal 벡터들을 랜덤 변수 $X \in R^d$ 로 보고, 변수 X 가 k 개의 가우시안들의 조합으로 표현되어지면 다음과 같은 확률 분포를 갖는다.

$$f(X|\theta) = \sum_{j=1}^k a_j \frac{1}{\sqrt{(2\pi)^d |\Sigma_j|}} \exp\left\{-\frac{1}{2}(X - \mu_j)^T \Sigma_j^{-1} (X - \mu_j)\right\} \quad (1)$$

$\theta = \{a_j, \mu_j, \Sigma_j\}_{j=1}^k$ 는 이미지 하나를 규정하는 가우시안 조합에 대한 파라미터로써 세개의 엘리먼트들은 다음과 같은 조건을 만족한다 :

$a_j > 0, \sum_{j=1}^k a_j = 1, \mu_j \in R^d$ and Σ_j is $d \times d$ positive definite matrix.

a_j 는 각 가우시안 조합에 대한 Weight로서 k 개의 웨이트들의 합은 1이 되도록한다. μ_j 는 한 이미지 내에서의 각각의 가우시안에 대한 평균 벡터를 의미하고, Σ_j 는 한 이미지 내에서의 가우시안 그룹들 간의 covariance행렬이다.

각각의 이미지에 대해서 파라미터 들을 구하는 과정은 EM(Expectation Maximization) 알고리즘을 통해서 이루어지며, 파라미터 θ 에 대한 log-likelihood(2) 를

반복적으로 계산하여 미리 정의된 임계값 이가 될 때의 파라미터로 결정한다.

$$\theta_{ML} = \arg \max_{\theta} f(x_1, x_2, \dots, x_n | \theta) \quad (2)$$

하나의 이미지에 대한 Gaussian 그룹의 수 k 의 결정은 Minimum Description Length(MDL) principle 에 따라 결정된다 [4].

하나의 이미지에 위와 같은 방법으로 파라미터들을 결정하고 나면 그 이미지는 Independently Identically Distributed(IID) 된 Feature들의 집합으로 구성된다고 볼 수 있다.

3.2 이미지 매칭

3.2.1 k-d tree

k-d tree 알고리즘은 가장 가까운거리(유클리디안 거리)의 데이터 베이스의 Feature들을 빠른 검색을 위해 트리 구조에 배열시키고, 입력 영상의 Feature들이 데이터 베이스의 어떤 이미지의 Feature로 매칭되는 지를 검색하는 알고리즘이다 (Friedman et al., 1977).

데이터 베이스의 Feature 들을 트리구조에 배열 시킬 때에는 (d-차원의 Feature들이라고 가정) 1차원 엘리먼트 들에 대한 median 값을 트리루트에 넣고 다음 엘리먼트 들로 순차적으로 옮겨가면서 median 값을 보고 브랜치 노드에 넣는식으로 진행된다.

본 논문에서 확률적 모델링을 사용하였을 때와 비교실험을 위한 방법으로 사용하였다.

3.2.2 Kullbak-Leibler(KL)

Gaussian Mixture Model(GMM)을 사용하여 이미지에 대한 확률적 모델링을 완성 후 입력영상의 가장 가까운 데이터베이스 상의 이미지로 매칭을 할 때는 두 개의 확률적 분포를 갖는 개체의 매칭으로 생각할 수 있다. 두 확률적 모델 간의 상대적 엔트로피를 측정하는 방법인 Kullbak-Leibler(KL) distance를 사용한다[5]. 식 (3)으로 KL distance 를 정의한다.

$$KL(f_1 \| f_2) = E_{f_1} \log \frac{f_1(x)}{f_2(x)} \quad (3)$$

본 논문에서 쓰이는 것과 같은 불연속적인 데이터를 갖는 Feature에 대한 KL distance는 식 (4)와 같이 근사화 시켜 사용한다.

$$KL(f_1 \| f_2) \cong \frac{1}{n} \sum_{x=1}^n \log \frac{f_1(x)}{f_2(x)} \quad (4)$$

4. 실험 결과

실험을 위하여 하나의 실내환경 카테고리에 대하여 하나의 랜드마크를 GMM으로 구성하여 20개의 데이터베이스를 구성하였고, 테스트 입력 영상은 하나의 랜드마크당 Affinely transform 된 카메라 위치에서 테스트 이미지를 구성하여 총 7*20=140개의 이미지를 사용하였다. 그림 4. 는 하나의 랜드마크와 그에 따른 테스트 입력 영상들의 예를 나타내었다.

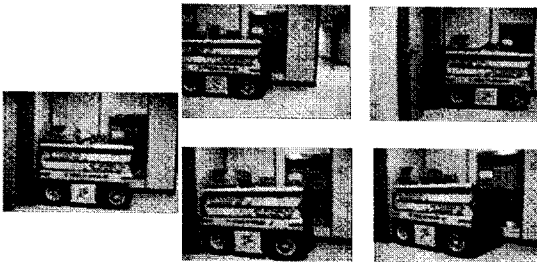


그림4.(a) 데이터베이스상의 랜드마크 이미지
(b,c,d,e) 테스트 입력 이미지

표 1. 에서는 각 클래스에 대해 k-d 트리를 이용하여 매칭한 결과와 GMM-KL 모델링을 통한 매칭 결과를 비교하였다.

5. 결론

Gaussian Mixture Model과 Kullbak-Leibler (GMM-KL)에 적합하도록 SIFT Feature을 이용하여 실내 환경 이미지에 대한 확률적 모델링-매칭을 할 경우 Feature 간 유클리디안 거리를 이용하는 방법에 비해 더 높은 정확도를 얻을 수 있었다.

PCA를 결합한 SIFT Feature를 쓰는 것 또한 각각의 이미지에 대한 독립적인 특성을 부여하여 확률적 유사도 산출시 구별성을 높여주었다.

이와 같은 이유로 단순한 거리를 이용할 경우 입력 영상의 하나의 Feature 가 올바른 랜드마크의 Feature로 매치되지 않고 더 가까운 거리에 있는 다른 랜드마크의 Feature로 매치될 가능성이 높기 때문이다. 반면에 하나의 이미지에 대해 Feature들의 확률적 분포로 표현한다면 거리에 의한 매칭보다 각각의 이미지에 대한 보다 독립적이고 정확한 설명이 가능해져 보다 높은 매칭 성공률을 보임을 알 수 있다.

표1.각이미지클래스당거리를이용한매칭과
GMM-KL방법의매칭정확도

	Class1	Class2	Class3	Class4	Class5
k-d	71.43%	57.14%	71.43%	57.16%	28.58%
GMM_KL	100%	71.43%	71.43%	57.16%	57.16%
	Class6	Class7	Class8	Class9	Class10
k-d	14.29%	57.16%	57.16%	57.16%	100%
GMM_KL	57.16%	42.87%	71.49%	57.16%	100%
	Class11	Class12	Class13	Class14	Class15
k-d	71.43%	100%	42.87%	57.16%	42.87%
GMM_KL	100%	100%	71.43%	57.16%	57.16%
	Class16	Class17	Class18	Class19	Class20
k-d	58.16%	42.87%	28.58%	71.43%	85.74%
GMM_KL	57.16%	100%	57.16%	57.16%	85.74%

참 고 문 헌

- [1] C. Carson, S. Belongie, H. Greenspan, J. Malik, Reocognition of images in large databases using color and texture, IEEE Transaction on Pattern Analysis and Machine Intelligence IEEE-PAMI, 24(8): 1026-1038, 2002
- [2] D.G. Lowe. Object recognition from local scale invariant features. In Proceedings of the Seventh International Conference on Computer Vision (ICCV '99),pages 1150-1157,Kerkyra, Greece, September 1999.
- [3] H. Greenspan, J. Goldberger, L. Ridel, A Continuous probabilistic framework for image matching, Journal of Computer Vision and Image Understanding, Vol. 84, No. 3, pp.384-406, 2001
- [4] J. Rissanen, Stochastic complexity in statistical inquiry, World scientific, 1989
- [5] S.Kullback. Information theory and Statistics. Dover,1968