

개선된 chain code와 HMM을 이용한 내용기반 영상검색

조완현*, 이승희*, 박순영**, 박종현**

*전남대학교 통계학과

**목포대학교 전자공학과

Content-based Image Retrieval using an Improved Chain Code and Hidden Markov Model

WanHyun Cho*, SeungHee Lee*, SoonYoung Park**, JongHyun Park**

*Department of Statistics, Chonnam National University

**Department of Electronics Engineering, Mokpo National University

e-mail : {whcho, u9898126}@chonnam.chonnam.ac.kr

{syPark, jhpark}@chungkye.mokpo.ac.kr

Abstract

In this paper, we propose a novel content-based image retrieval system using both Hidden Markov Model(HMM) and an improved chain code. The Gaussian Mixture Model(GMM) is applied to statistically model a color information of the image, and Deterministic Annealing EM(DAEM) algorithm is employed to estimate the parameters of GMM. This result is used to segment the given image. We use an improved chain code, which is invariant to rotation, translation and scale, to extract the feature vectors of the shape for each image in the database. These are stored together in the database with each HMM whose parameters (A, B, π) are estimated by Baum-Welch algorithm. With respect to feature vector obtained in the same way from the query image, a occurring probability of each image is computed by using the forward algorithm of HMM. We use these probabilities for the image retrieval and present the highest similarity images based on these probabilities.

I. 서론

최근에 급격한 멀티미디어 정보의 증가와 활용으로 다양한 내용기반 영상검색 기법들이 개발되어 소개되고 있다. 대표적인 것으로 IBM사의 QBIC 시스템, Berkely 대학의 CBIR 시스템, Columbia 대학의 Visual Seek와 WebSeek, 그리고 MIT의 Photobook 등을 생각할 수 있다. 이들 대부분의 방법들은 영상정보의 모양, 색상, 또는 질감 등을 사용한 내용기반 특징벡터를 추출하여 검색에 이용하였다.

본 논문에서는 개선된 chain code와 HMM을 사용한

새로운 내용기반 영상검색으로의 접근 방법을 제안하였으며 그림 1에서 시스템의 구조를 보여주고 있다. 따라서 제안된 영상검색 시스템에서는 각 영상의 색상정보에 대하여 가우시안 혼합모형(GMM)을 가정하고 결정적 어닐링 EM(DAEM) 알고리즘을 이용하여 GMM의 모수를 추정하며 동질의 성질을 갖는 영상의 군으로 객체를 분할하는 방법을 사용한다[1][2]. 그리고 객체 분할된 영상들로부터 회전, 이동, 축척에 불변인 개선된 chain code를 추출하여 HMM의 관측값으로 사용하며 Baum-Welch 알고리즘을 사용하여 HMM의 모수들을 추정한다. 또한 질의 영상의 특징벡터를 데이터베이스에 색인된 각각의 HMM에 대입하여 추정된 모수가 주어졌을 때 관측치가 출현할 확률을 계산하여 영상 검색에 이용하는 방법을 기술한다. 본 논문의 구성은 II장에서 개선된 chain code를 이용한 모양 특징 추출에 대하여 기술하고, III장에서는 추출된 특징벡터를 HMM에 이용한 영상검색 방법에 대하여 설명한다. IV장에서는 제안된 방법의 실험결과에 대하여 논하였으며, 마지막 V장은 제안된 방법의 결론과 향후 연구과제에 대하여 기술하였다.

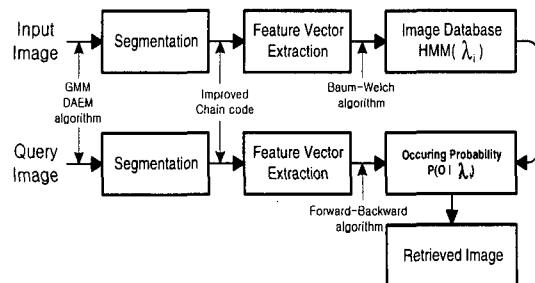


그림 1. 제안된 영상검색 시스템의 구조

II. 개선된 chain code를 이용한 모양 특징추출

2.1 GMM과 DAEM 알고리즘을 이용한 영상분할

영상이 K개의 독립된 객체로 구성되어 있다고 가정하면 칼라 특징벡터의 확률분포로 다음과 같은 K개의 확률밀도함수의 선형결합으로 표현할 수 있는 혼합모형(GMM)을 생각할 수 있다.

$$g(\mathbf{x}|\boldsymbol{\theta}) = \sum_{i=1}^K \pi_i f(\mathbf{x}|\boldsymbol{\theta}_i) \quad (1)$$

또한 RGB 모델로 표현된 입력영상을 HSI 색상모델로 변환후 각 화소의 H와 S성분을 이용하여 2차원 특징벡터로 표현하고, 각 색상의 확률분포로 가우시안 모형을 사용한다.

$$f(\mathbf{x}|\boldsymbol{\theta}_i) = \frac{1}{(2\pi)(\det \Sigma_i)^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)' \Sigma_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)} \quad (2)$$

이때 가정된 가우시안 혼합모형의 모수를 추정하기 위하여 다음과 같은 로그 우도함수를 생각하고, 이 우도함수를 최대화하는 방법으로 결정적 EM 알고리즘을 이용한다.

$$\log L(\mathbf{x}, \mathbf{z}|\boldsymbol{\theta}) = \sum_{i=1}^K \sum_{j=1}^N Z_{ij} \cdot \log g(\mathbf{x}_j|\boldsymbol{\theta}_i) \quad (3)$$

여기서 벡터 \mathbf{z} 는 각 화소가 어떤 확률모형에 속하는가를 나타내는 지시벡터이다.

다음으로 추정된 모수를 가지고 다음의 사후확률을 계산할 수 있고, 이 사후확률이 가장 높은 값을 갖는 확률모형에 화소를 할당하므로써 영상내의 동질의 특성을 갖는 객체들을 분할 할 수 있다.

$$p(i|\mathbf{x}, \boldsymbol{\theta}) = \frac{\pi_i f_i(\mathbf{x}|\boldsymbol{\theta}_i)}{\sum_{k=1}^K \pi_k f_k(\mathbf{x}|\boldsymbol{\theta}_k)}, \quad i=1, \dots, K \quad (4)$$

그림 2는 GMM을 적용하여 영상내의 객체를 분할한 결과를 보여주고 있다.

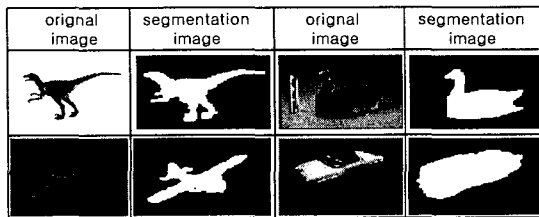


그림 2. GMM과 DAEM을 이용한 영상의 객체 분할

2.2 개선된 chain code

유사한 모양을 갖는 객체들로 분할된 영상으로부터 표준 chain code를 사용하여 영상의 윤곽선에 대한 특징벡터를 간단히 추출할 수 있으나 이것은 영상의 회전, 이동, 축척에 대한 단점을 가지고 있다[3]. 따라서

이러한 문제를 해결하기 위해 본 논문에서는 객체로부터 효율적인 윤곽 특징을 추출하기 위해서 개선된 chain code를 제안하였으며 그림 3에 구성도를 보여주고 있다. 분할된 객체의 경계는 잡음에 의한 작은 교란(disturbance)과 불완전한 객체분할 때문에 발생하는 코드의 불필요한 반복, 그리고 chain code의 길이가 일반적으로 매우 길기 때문에 resampling을 적용하여 이러한 문제를 제거하였다. 또한 chain code는 시작점에 따라 영상의 방향, 이동 등에 의존하기 때문에 shape number를 적용하여 방향과 이동에 불변인 chain code를 추출하였다. 그리고 Gaussian 필터를 사용하여 임펄스성 코드를 제거함으로써 윤곽선에 대한 코드를 smooth하게 만들어 최종적으로 개선된 8-D chain code를 추출하였다.

그림 4는 그림 3의 dinosaur 영상에 제안된 알고리즘을 적용하여서 그래프를 이용하여 시각적으로 보여주고 있으며 4(a)는 기존 8-D chain code, 4(b)는 resampled 8-D chain code, 그리고 4(c)는 smoothed 8-D chain code를 보여주고 있다.

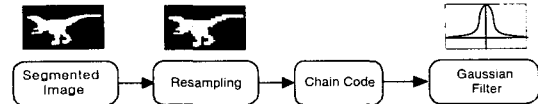
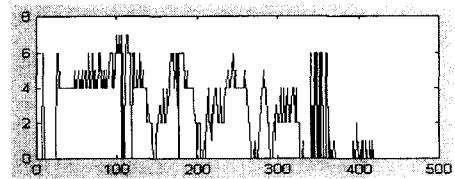
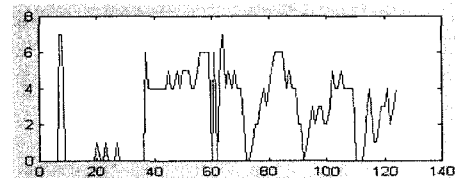


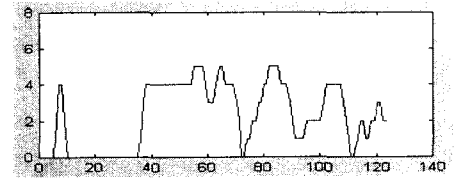
그림 3. 개선된 chain code의 구성도



(a)



(b)



(c)

그림 4. dinosaur 영상에 대한 기존의 chain code와 개선된 chain code의 비교:(a)기존 8-D chain code, (b)resampled 8-D chain code, (c)smoothed 8-D chain code

III. HMM을 이용한 영상검색

3.1 EM 알고리즘을 사용한 HMM 모수집합의 학습

HMM은 유한개의 상태집합 $\{s_1, \dots, s_N\}$ 이 존재하고, 각 상태는 특정한 확률분포를 갖는 이산형 확률과정이다. 이때 각 상태의 이동은 추이확률에 의해서 결정되며, 각 상태에서 M 가지 관측값 $\{o_1, \dots, o_M\}$ 이 주어지는데 이것은 특정한 확률분포를 따른다. 따라서 HMM은 다음과 같은 모수집합 $\lambda = (A, B, \pi)$ 들에 대해서 정의될 수 있다[4][5][6].

① A : 상태 추이 확률 행렬

$$A = (a_{ij}), \quad a_{ij} = P\{S_t = j | S_{t-1} = i\}, \quad 1 \leq i, j \leq N \quad (5)$$

② B : 관측 확률 분포 행렬

$$B = (b_j(k)), \quad b_j(k) = P\{O_t = k | S_t = j\}, \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \quad (6)$$

③ π : 초기 상태 확률분포

$$\pi = (\pi_1, \dots, \pi_N), \quad \pi_i = P\{S_1 = i\}, \quad 1 \leq i \leq N \quad (7)$$

이때 HMM의 모수들을 추정하기 위해서 EM 알고리즘으로 알려진 Baum-Welch 알고리즘을 이용할 수 있다.

① 초기화 단계

HMM의 모수집합 $\lambda = (A, B, \pi)$ 에 대하여 적절한 초기 모수값 λ_0 을 가정한다.

② 반복단계

먼저 영상으로부터 추출된 특징벡터 $O = (o_1, \dots, o_T)$ 와 전 단계에서 얻어진 모수값 λ_{old} 들을 이용하여 전향변수들과 후향변수들을 계산한다.

$$\alpha_t(i) = P(O_1 = o_1, \dots, O_t = o_t, S_t = s_t | \lambda_{old}) \quad (8)$$

$$\beta_t(i) = P(O_{t+1} = o_{t+1}, \dots, O_T = o_T | S_t = s_t, \lambda_{old}) \quad (9)$$

또한 이들 값과 베이즈 법칙을 사용하여 사후확률과 결합확률을 계산한다[7].

$$\gamma_t(i) = \left[\frac{\alpha_t(i)\beta_t(i)}{\sum_{i=1}^N \alpha_t(i)\beta_t(i)} \right] \quad (10)$$

$$\xi(i, j) = \frac{\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{P(O|\lambda)} = \frac{\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)} \quad (11)$$

최종적으로 위에서 주어진 두 가지 확률을 사용하여 새로운 모수집합 $\lambda_{new} = (A, B, \pi)$ 의 추정값을 다시 계산한다.

$$\hat{\pi}_i = \gamma_1(i), \quad 1 \leq i \leq N \quad (12)$$

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \quad 1 \leq i \leq N, \quad 1 \leq j \leq N \quad (13)$$

$$\hat{b}_j(k) = \frac{\sum_{t=1}^T \delta_{o_t, o_k} \cdot \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)}, \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \quad (14)$$

③ 종료단계

전체적으로 모수 집합에 추정치들의 변화량이 무시할 만큼 작아지면 반복을 종료하고, 그렇지 않으면 이런 과정을 계속하여 모형을 완성한다.

3.2 특징벡터와 HMM를 이용한 영상검색

영상검색 시스템의 목적은 가장 유사한 영상이 효과적으로 검색될 수 있도록 영상을 조직화하고, 색인하는 작업이다. 즉, 영상검색 시스템의 두 가지 중요한 과제는 영상의 적절한 색인과 검색 과정에서 유사성의 비교라고 할 수 있다.

따라서 우리는 데이터베이스에 들어있는 각 영상들에 대하여 개선된 chain code 특징벡터를 추출하고, 이들을 기초로 하여 HMM의 모수집합 $\lambda = (A, B, \pi)$ 을 추정한다. 그리고 이들을 사용하여 각 영상을 색인한다. 다음으로 질의 영상으로부터 추출된 개선된 chain code 특징벡터 $O = (o_1, \dots, o_T)$ 에 대하여 데이터 베이스에 색인되어 있는 각 영상의 HMM을 사용하여 이 특징벡터가 출현할 확률을 계산한다.

$$P(O | \lambda) = \sum_{i=1}^T \alpha_T(i) \quad (15)$$

그런데 이 확률은 일반적으로 아주 작은 값으로 주어진다. 따라서 우리는 다음과 같이 정의되는 값을 질의 영상과 데이터베이스내의 색인 영상간에 비유사도로 사용할 수 있다.

$$D = -\log P(O | \lambda) \quad (16)$$

그러므로 우리는 두 영상간의 거리를 계산하여 이것이 최소가 되는 HMM, $\lambda_{min} = (A, B, \pi)$ 에 대응되는 영상으로 검색한다.

IV. 실험 및 결과

본 논문에서 제안한 개선된 chain code와 HMM을 이용한 내용기반 영상검색 방법의 효율성을 실험하기 위하여 영상들의 객체분할, 특징벡터추출, HMM학습, 질의 영상에 대한 출현 확률 등을 수행하였다.

실험을 위해 사용된 영상 데이터베이스는 다른 구조와 색상으로 구성된 Corel photo 칼라 영상 105장을 혼합하여 사용하였으며 그림 5에 일부를 보여주고 있다. 데이터베이스의 영상은 120×80의 화소들로 이루어진 256 칼라 영상으로 dinosaur, airplane, car, china, doll, duck, rose의 모두 7개의 클래스로 구성되어 있다.

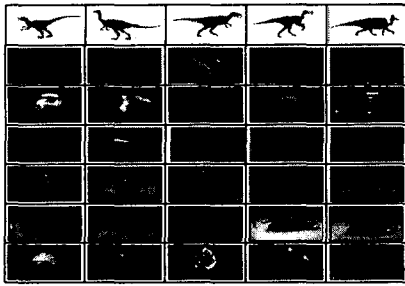


그림 5. 실험에 사용된 database의 일부분

HMM의 실험에서 상태는 8개인 ergodic 모델을 사용하였으며 데이터베이스에 있는 각각의 영상에 대하여 추출된 shape number 특징벡터를 이용하여 HMM의 학습을 수행하였다. 이때 특징벡터들은 HMM의 관측값으로 생각할 수 있으며 Baum-Welch 알고리즘을 이용하여 HMM의 파라미터를 추정하였다. 질의 방법은 데이터베이스 영상들로부터 랜덤하게 하나의 질의 영상을 선택하여 특징벡터를 추출하고 이들을 데이터베이스에 학습되어 있는 각각의 HMM에 대입하여 관측치가 출현할 확률 $P(O|\lambda)$ 을 계산한다.

그림 6은 제안된 알고리즘의 수행 결과를 보여주고 있는데 임의의 질의 영상을 각 클래스별로 한 개씩 선별하여 관측치가 출현할 확률이 높은 상위 3개의 영상을 검색 결과로써 보여준다. 그림에서 객체의 모양에 대한 클래스 사이의 유사성이 작은 dinosaur, airplane, duck 영상은 검색 결과가 우수하게 나타났으나, car, china, doll, rose 영상은 객체들이 클래스 사이에 서로 유사성을 가지고 있어서 검색 결과가 빈약하게 나타남을 볼 수가 있다.

실험 결과에서 볼 수 있듯이 질의 영상에 대하여 영상이 가지고 있는 객체의 모양이 임의의 방향성을 가지고 있어도 개선된 chain code를 사용하여 특징벡터를 추출하였기 때문에 클래스 사이의 패턴에 대한 유사성이 작은 경우에는 잘 동작함을 알 수 있다.

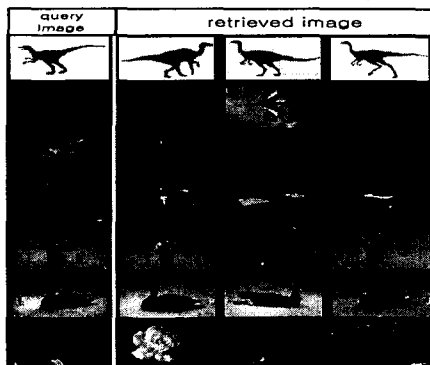


그림 6. 제안된 내용기반 영상검색의 검색 결과

V. 결론 및 향후과제

본 논문에서는 영상이 포함하고 있는 객체의 모양에 대하여 개선된 chain code를 이용하여 특징 벡터를 추출하고 여기에 HMM을 적용하여 영상검색을 수행하는 방법을 제안하였다. 영상에 포함된 객체의 모양특징을 추출하기 위하여 여러 색상으로 이루어진 자연 영상에 GMM을 적용하여 통계적 분포특성에 따라 농도의 영역을 포함하는 샘플들의 군을 객체로 분할하였으며 분할된 객체에 개선된 chain code를 이용하여 회전, 축척, 이동에 불변한 모양 특징을 추출하였다. 추출된 특징벡터들은 HMM의 관측값으로 생각하고 학습을 하였으며 Baum-welch 알고리즘에 의해 HMM의 파라미터를 추정하였다.

유사성에 의한 정합은 입력된 질의 영상의 특징벡터를 이용하여 학습된 HMM에 대한 관측치 출현 확률을 계산하였으며 관측값이 가장 높게 나오는 영상을 질의 영상과 가장 유사한 영상으로 검색할 수가 있었다.

결과적으로 본 논문에서는 개선된 chain code에 의해 추출된 모양 특징벡터를 HMM에 적용함으로써 주어진 패턴에 대하여 잘 모델링 할 수 있어 내용기반 영상검색에 효율적으로 응용할 수 있음을 볼 수 있었다. 그러나 내용기반 영상검색에의 응용에 있어서 보다 더 검색 능력 향상을 위하여 영상의 공간적인 색상 특징을 추가하여 복합적으로 검색을 수행하는 것과 학습시간의 단축에 대한 연구가 차후 계속 진행되어야 할 것이다.

참고문헌

- [1] 박종현, 박순영, 조완현, "결정적 어닐링 EM 알고리즘을 이용한 칼라영상의 분할", 대한전자공학회 제 22권, 제2호, 1999, pp. 569-572.
- [2] 박종현, 박순영, "객체의 모양과 색상특징을 이용한 내용기반 영상검색 기법", 제11회 영상처리 및 이해에 관한 워크샵 발표논문집, 1999, pp. 50-55.
- [3] B.M. Mehtre, M.S.Kankanhalli, W.F.Lee, "Shape Measures for Content based Image Retrieval: a Comparison", Information Processing and Management, Vol. 33, No. 3, 1997, pp. 319-337.
- [4] Y. He and A. Kundu, "2-D Shape Classification Using Hidden Markov Model," IEEE Trans. on PAMI, vol 13, No. 11, 1991, pp. 1172-1191.
- [5] L.R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceeding of the IEEE, Vol 77, No. 2, Feb. 1989, pp. 257-286.
- [6] S. Muller, S. Erickeler and G. Rigoll, "Image Database Retrieval of Rotated Objects by User Sketch", IEEE workshop on CBAIVL, 1998, pp. 40-44.
- [7] P.McKenzie, M.Alder, "Initializing the EM Algorithm for Use in Gaussian Mixture Modelling" Pattern recognition in practice IV, Vlieland, 1994.