

확률적 부울(Boolean) 모델과 연관성 학습을 통한 내용기반 영상 검색 성능 향상

고병철⁰ 변혜란

연세대학교 컴퓨터과학과
{soccer1, hrbyun}@aipiri.yonsei.ac.kr

Performance Improvement For Content-Based Image Retrieval Using Probabilistic Boolean Model And Relevance Learning

ByoungChul Ko⁰ Hyeran Byun

Dept. of Computer Science, Yonsei University

요약

전체 영상을 이용하지 않고 영상 안에 포함된 특정 객체 혹은 영역만을 이용하는 "영역에 의한 질의(query-by-region)" 방법은 내용기반 영상 검색 중 상위개념의 방법이지만, 영상 분할의 한계, 여러 개로 분할된 영역을 모두 검색하기 위한 인덱싱 문제, 유사성 측정 시 선형적으로 분리되지 않는 특징 값들에 대한 무리한 선형 조합으로 인한 검색 오류와 같은 많은 문제점을 안고 있다. 따라서 본 논문에서는 영역 기반 영상 검색 시스템인 FRIP에 대하여 영상 분할의 한계를 극복하고, 사용자의 주관성을 영상 검색에 적용하기 위해 확률적 연관성 학습 모델(MPFRL)을 유사성 측정 단계에서 적용 하였고, 아울러 검색 모델로는 기존에 일반적으로 사용되어 오던, 선형 모델을 사용하지 않고 선형 모델보다 유연한 검색 결과물 보여주는 확률적 이점 부울 모델(PDB)을 사용하였다. 또한, 검색 시간을 단축 시키기 위해, 선형 검색 방법에 부울 AND 연산자를 적용 시킴으로써, 검색 시간을 상당부분 단축 할 수 있었다. 실험 결과, 본 논문에서 제안하는 방법(MPFRL+PDB)을 사용할 경우 검색 결과가 선형 조합 보다 향상되는 것을 알 수 있었다. 아울러 사용자 피드백을 통해 사용자가 특정 가중치를 일일이 조절하지 않더라도 단순한 몇 번의 클릭만으로 사용자의 주관성을 반영하고 보다 정확한 검색 결과물을 보여 줄 수 있는 시스템을 설계 할 수 있었다.

1. 서론

디지털 영상, 비디오와 같은 멀티미디어 데이터의 급격한 증가는 정확한 검색 결과와 단순한 사용자 인터페이스를 가지는 검색 시스템을 요구하고 있다. 초기 텍스트 기반 영상은 좋은 검색 결과를 보여주는데 불구하고 많은 한계점으로 인해, 대부분의 영상 검색 시스템은 색상, 질감, 모양 정보와 같은 영상의 내용을 추출하여 유사성 측정에 적용하는 방법을 사용하고 있다. 특히, 전체 영상을 이용하지 않고 영상 안에 포함된 특정 객체 혹은 영역만을 이용하는 "영역에 의한 질의(query-by-region)" 방법은 내용기반 영상 검색 중 상위개념의 방법으로 아직까지는 많은 한계점에 부딪히고 있다. 무엇보다도, 전처리 단계에서 대부분의 자연영상을 인간의 시각 체계와 비슷한 수준으로 분할 한다는 것은 현재의 컴퓨터 비전 기술로는 많은 어려움이 따름으로, 영역기반 영상검색은 이러한 단점에 대한 보완책이 필요하다. 따라서 본 논문에서는, 본 논문의 영역기반 영상 검색 시스템인 FRIP (Finding Region In the Pictures) [1]을 기반으로, 인간의 상호작용을 이용한 가중치 학습 개념을 통해 사람이 일일이 가중치를 조절하지 않더라도 몇 번의 피드백 동작만으로 최적의 가중치를 찾아내고, 보다 유사한 영역을 포함하는 영상을 찾도록 하는데 연구의 초점을 두었다. 아울러, 영상검색의 비교 횟수를 줄이고 정확한 검색을 위해 5개의 특징 벡터 중, 최적의 2개 특징 벡터를 선출하여 검색 단계에서 부울(Boolean) AND 연산자를 적용시켜 검색 시간을 줄일 수 있는 방법을 제안한다. 마지막으로, 대부분의 영상 검색에서 많이 사용되고 있는 특징 벡터들의 선형 조합 모델이 실제로 선형(linear)적으로 분리되지 않는 특징 벡터들을 한 개의 특징 벡터로 선형 조합하는 단점이 있으므로 본 논

문에서는 특정 벡터들의 거리차를 선형 조합하지 않고, 확률적 부울 이점 모델 (Probabilistic Boolean Disjunctive Model : PDB)에 적용시켜 보다 유연하고 정확한 검색 결과를 얻을 수 있도록 하였다.

본 논문의 구성은 다음과 같다. 2장에서는 FRIP 시스템에 대한 간략한 설명과 부울 AND 연산자를 이용한 유사성 측정 방법에 초점을 맞춘다. 3장에서는 사용자 연관성 피드백을 통해 검색 결과에 대한 가중치 조절 방법을 언급한다. 4장에서는 선형 조합이 아닌, 부울 이점 모델을 사용하여, 각 특징 벡터들을 조합하는 방법을 설명하고, 5장에서 실험 결과를 통해 선형 조합 방법에 비해 부울 이점 모델이 보다 유연하고 정확한 결과를 보여줌을 실험을 통해 제시한다.

2. 관련연구

FRIP [1] 시스템은 영역 기반 검색 시스템이다. FRIP 시스템에서는 영상을 색상 및 질감 성분으로 따라 영역으로 분할하고, 특히 원형 필터를 개발하여 적용 함으로써, 일반적인 영역 병합 방법에서 호랑이의 줄무늬와 같은 자연영상을 한 개의 영역으로 병합시키지 못하는 문제점을 해결하였다. 영역이 분할 된 후, 색상, 질감, 모양, 크기, 위치 정보의 5개 특징 벡터가 추출되고 가우시안 정규화에 의해 0-1 사이의 값으로 조정 되어 특징 값으로 저장된다.

2.1 부울 AND 연산자와 계단식 정합 방법을 이용한 유사성 측정

일반적으로 선형 검색(Linear Search)방법은 데이터베이스가 1,000 ~ 2,000개 정도로 작을 경우 "실시간"의 검색 결과를 얻을 수 있다. 또한, 특징 값의 dimension이 10 이상일 경우, multi-dimension 데이터를 인덱스 하는 비용은 데이터베이스 영상이 증가함에 따라 exponential 하게 증가 함으로, 이 경우 선형 검색은 R-tree, X-tree, B-tree 등 보다 오히려 효율적인 인덱싱 방법이 될 수 있으므로, 대부분의 영상 검색 시스템에서는 이 방법을 사용하고 있다 [2]. 하지만, 영역 기반 영상 검색에서는 영상의 영역이 내용에 따라 여러 개의 영역들로 분할 되고, 마찬가지로 각 영역들로부터 여러 개의 특징 벡터들이 추출되어야 한다. 따라서, 검색 시 질의 영역은 데이터베이스의 각 영역들과 모두 비교해야 함으로, 단순한 선형 검색 방법을 사용할 경우, 많은 검색 시간을 필요로 하게 된다. 따라서 본 논문에서는 5개의 특징 벡터 중 색상과 모양 정보를 이용하여 첫번째 단계에서 후보 영역만을 찾아내고, 후보 영역들에 대해서만 특징 벡터들 간의 거리차를 계산하여 최종적으로 부울 이접 모델에 적용하는 방식을 취한다. 본 논문에서 제안하는 부울 AND 연산자와 계단식 정합 방법을 이용한 유사성 측정 방식 단계는 아래와 같다.

- ① 첫번째 단계에서, 사용자가 color-care 와 shape-care 조건 중 하나 또는 2개 모두 선택했을 경우, 질의 영역과 데이터베이스 영상의 각 영역들간에 우선적으로 색상과 모양에 대한 거리(distance)가 구해진다. 각 데이터베이스 영상으로부터 최소 거리를 갖는 영역을 찾은 뒤에 해당 영역의 두 거리 값이 모두 임계 값을 만족할 경우 해당 영역을 포함하는 영상은 후보(candidate) 영상으로 선언된다. 본 논문에서는 임계 값을 데이터베이스의 10~20%의 영상만을 선출하도록 조정하였다.
- ② 나머지 사용자 질의 조건(texture-care/not, size-care/not, location-care/not)에 따라, 후보 영역과 질의 영역 사이에 거리 값이 구해진다.
- ③ 계산된 거리 값으로부터 PDB 모델에 의해 최종 점수가 계산 되고(수식 - (5)), 내림차순으로 최상위 K개의 영상이 사용자 인터페이스에 검색 결과로 보여진다.

3. 확률적 부울 이접 모델(Probabilistic Boolean Disjunctive Model)

일반적으로 대부분의 내용기반 영상 검색은 유사성 측정을 위해서 각각의 특징 값들 사이에 차이 값을 계산하고 이를 이용하여 가중치 선형 조합하는 형태를 취한다.

$$Sim(X, Y) = \sum_{i=1}^n w_i D_i(x_i, y_i) \quad (1)$$

w_i : 선형 조합을 위한 특징 벡터의 가중치

D_i : 각 특징 벡터사이의 거리

하지만, 이러한 방법은 각 특징 벡터들이 항상 선형적으로 분리 가능하다는 가정을 가짐으로, 실제 유사성 측정에서, 선형 분리 가능하지 않은 많은 특징들을 잘못 분리할 수 있는 단점이 있다. 이와는 대조적으로, 대부분의 상업적 정보 검색 시스템들은 부울 모델(Boolean AND, OR, NOT)을 사용하고 있다. 하지만 부울 모델의 경우 검색 결과에 대한 측정치를 제공하지 못하므로, 가중치 방법이나, 퍼지, 또는 확률적 모델을 부울 모델에 적용함으로써 검색 결과의 순위를 결정 할 수 있다 [3].

본 논문에서는 부울 모델 중에서, 확률적 부울 이접 모델(Probabilistic Boolean Disjunctive Boolean model)을 검색 결과의 우선 순위 결정을 위한 membership 함수(function)로 적용 하였다. 이때, 각 특징 벡터들의 확률 분포에 대한 가중치를 적용하고, 사용자의 상호 작용적인 피드백을 제공하기 위해 본 논문에서는 PDB 모델에 본 논문에서 제안하는 개선된 확률적 연관성 학습 모델(Modified Probabilistic Feature Relevance Learning : MPFRL) [4]을 결합하였다.

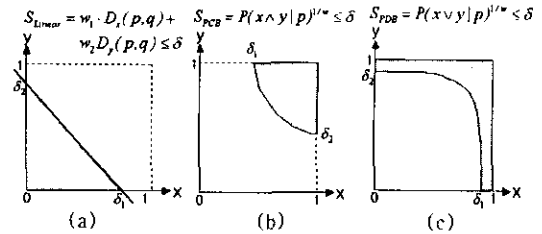


그림 1. 세가지 검색 모델: (a)선형 모델, (b) 부울 결합(AND) 모델(PCB) (c) 부울 이접(OR) 모델(PDB)

PDB 모델에서 $Q(f_1, f_2, \dots, f_n)$ 질의 영역이라고 가정하고, I 를 데이터베이스 영상이라고 가정하자. 이 모델에서 질의 영역의 특징 벡터와 I 의 한 특징 벡터 사이의 거리를 D_i 라 할 때, 둘 사이의 확률 $P(f_i | I)$ 는 수식 (2)에 의해 0-1 사이의 값으로 변환된다. 이때, 주어진 영상에서 각각의 특징들은 서로 독립적이라고 가정 한다.

이런 가정 하에서, 특징 벡터 f_i, f_j 에 대한 수식 (3)은 유효하게 된다.

$$P(f_i | I) = \left(\frac{1}{1 + D_i} - \frac{1}{2} \right) \times 2 \quad (2)$$

$$P(f_i \wedge f_j | I) = P(f_i | I) \times P(f_j | I) \quad (3)$$

다음으로, 각 특징 거리 값에는 가중치가 적용되므로 수식 (2)로부터 구해진 각 특징 거리의 확률 값은 수식 (4)에 의해 가중치가 적용된 형태로 변환되어야 한다.

$$P(f_i | I)' = P(f_i | I)^{w_i} \quad (4)$$

5개의 가중치가 적용된 확률 값을 가지고 PDB 모델에 의해 전체적인 특징 값의 거리는 수식 (5)에 의해 결정된다.

$$P(D_1 \vee \dots \vee D_5 | I) = P(f_1 | I)' + \dots + P(f_5 | I)' - P(f_1 \wedge \dots \wedge f_5)' \quad (5)$$

$$P(f_1 \wedge \dots \wedge f_5)' = \prod_{i=1}^5 P(f_i | I)' \quad (6)$$

4. 가중치 갱신을 위한 확률적 특징 값의 연관성 학습

본 논문에서는 결과 영상에 대한 사용자의 상호 작용적인 피드백을 제공하기 위해 사용자의 피드백 영상을 training data로 설정하고 질의 영상(z)과 training data의 특징 값(x_i) 사이의 지역적 연관성(local feature relevance)을 이용하여 가중치를 갱신하는 방법을 사용한다(수식 - (7))[4].

$$r_i(z) = E[f(x) | x_i = z] \quad (7)$$

각각의 특징 값에 대한 연관성(r_i)을 구한 뒤에, 각 특징 값에 대한 가중치는 수식 (8)과 같이 전체 특징 값의 연관성으로, 각 특징 값들의 연관성을 나누는 형태로 결정된다. 이때, T 는 단순히 가중치에 대한 연관성의 영향을 최대화 혹은 최소화($T=1$) 하기 위한 파라메타이다.

$$w_i(z) = \exp(Tr_i(z)) / \sum_{i=1}^n \exp(Tr_i(z)) \quad (8)$$

각 단계(iteration)에서, 사용자는 주관에 따라 수식 (5)에 의해 결정된 상위 k만큼의 영상을 연관(relevant) 혹은 비연관(irrelevant) 영상으로 선택하고, 이렇게 선택된 영상들은 연관성 여부에 따라 positive / negative history에 저장된다. 다음으로, 수식 (8)에 의해 계산된 각 가중치들은 이전 단계에서 구해진 가중치와의 선형 조합으로 다음 단계를 위한 가중치가 된다. 이렇게 함으로써 가중치는 단조 증가 혹은 단조 감소할 수 있게 되고, 보다 정확한 검색 결과를 보여 줄 수 있게 된다.

$$w_i^{t+1} = (1-\alpha) \cdot w_i^t + \alpha \cdot w_i^{-1} \quad (9)$$

t : 반복 단계, α : 적합도 파라메타($0 \leq \alpha \leq 1$)

수식 (9)에 의해 결정된 가중치는 수식 (4)의 W_i 에 적용되고 수식(5)는 수식(10)에 적용되어 최종적인 거리 값을 계산 하게 된다.

if $y \in H_p \{h_1, \dots, h_n\}$

$$D(x, y) = P(f_1 \vee \dots \vee f_s | I) \cdot \hat{K} \quad (10)$$

else if $y \in H_n \{h_1, \dots, h_n\}$

$$D(x, y) = P(f_1 \vee \dots \vee f_s | I) / \hat{K}$$

$H_p = \{h_1(I, P_1), h_2(I, P_2), \dots, h_n(I, P_n)\}$: positive history

$H_n = \{h_1(I, N_1), h_2(I, N_2), \dots, h_n(I, N_n)\}$: negative history

\hat{K} : 결정 파라메타($\hat{K} > 1$), $P(f_i | I)$: 특징 거리 값 확률

i: color, texture, NArea, shape, location

5. 실험 결과

본 논문에서는 실험을 위해 자연 영상, 그래픽, 그림, 도시 영상 등이 포함된 3,000개의 Corel-photo CD 영상을 사용하였다. 본 논문에서는 펜티엄 III PC 450을 사용하여, 평균 1,000개의 영상을 검색하는데 약 2초의 시간이 필요하였다. 이것은 한 개 영상이 30개의 영역으로 나뉘어 있다고 가정하고 5개의 특징을 사용한다면, 일반 선형 검색을 사용할 경우 평균적으로 150,000의 비교 횟수가 필요한 반면, 부울 AND 연산자를 적용할 경우 최대 12,000의 연산만이 필요함으로 결과적으로 약 12.5배의 검색 횟수를 줄일 수 있다.

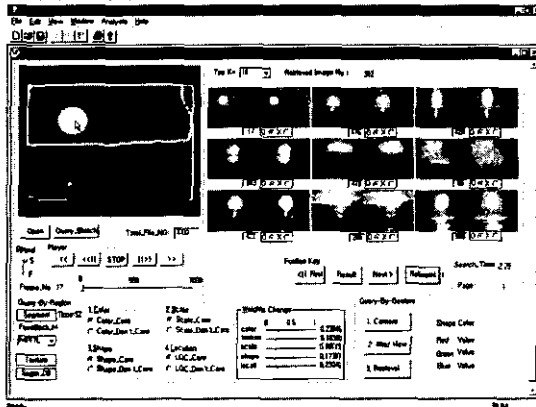


그림 2. FRIP 시스템의 사용자 인터페이스

PDB 모델과 선형 모델과의 성능 비교를 위해서, 본 논문에서는 8개의 객체(호랑이, 독수리2, 자동차, 비행기, 태양, 꽃, 흰곰)에 대해서 precision을 측정 하였다. 그림 3은 8개 객체에 대해 학습의 반복 횟수에 따른 평균 precision의 변화량을 보여준다. 그림에서 보는 것과 같이, PDB모델은 선형 조합 모델(수식 (1))이나 PCB 모델(수식 (6))과 비교하여 보다 향상되고 유연한(단조 증가) 결과를 보여줌을 알 수 있다.

6. 결론

본 논문에서는 영역 기반 영상 검색 시스템인 FRIP에 대하여 영상 분할의 한계를 극복하고, 사용자의 주관성을 영상 검색에 적용하기 위해 확률적 연관성 학습 모델(MPFRL)을 유사성 측정 단계에서 적용 하였다. 아울러 검색 모델에는 기존에 일반적으로 사용되어 오던, 선형 조합 모델을 사용하지 않고 선형 모델보다 유연한 검색 결과를 보여주는 확률적 이점 부울 모델(PDB)을 적용 하였다. 실험 결과, 본 논문에서 제안하는 방법(MPFRL+PDB)을 사용할 경우 검색 결과가 일반적인 선형 조합 모델이나 PDB모델 보다 향상되는 것을 알 수 있었다. 본 논문에서 제안하는 FRIP시스템 사용할 경우, 사용자는 특징 가중치를 일일이 조절하지 않더라도 단순한 몇 번의 클릭 동작을 통해 사용자의 주관성을 반영하고 보다 의미 있고 정확한 영상을 검색 할 수 있다.

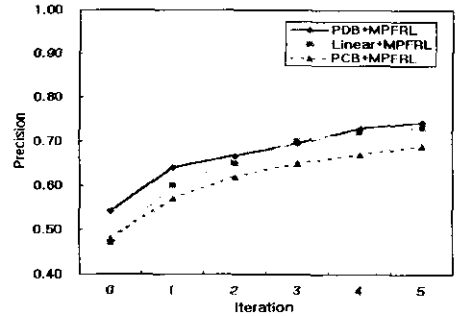


그림 3. MPFRL 방법과 세 개의 모델간의 평균 검색 성능 비교

표 1. 반복 횟수(IT)에 따른 상위 18개 영상에 대한 검색 성능 (Precision)

Query	IT-0	IT-1	IT-2	IT-3	IT-4	IT-5
Tiger	39	39	50	50	56	61
Car	33	50	50	50	61	61
Eagle1	72	89	94	94	94	94
Airplane	33	33	33	39	44	44
Sun	78	89	89	100	100	100
Eagle2	67	83	89	89	89	89
Flower	89	100	100	100	100	100
White	22	28	28	33	39	44
Average	54.1	63.9	66.6	69.3	72.9	74.1

7. 참고문헌

- [1]ByoungChul Ko, Hae-sung Lee, Hyeran Byun, Region-based Image Retrieval System Using Efficient Feature Description. Proceedings of the 15th International Conference on Pattern Recognition, vol. 4(4), pp. 283-286, 2000.
- [2]Hari Sundaram and Shih-Fu Chang, Efficient Video Sequence Retrieval in Large Repositories. Proceedings of SPIE Storage and Retrieval for Image and Video Database VII, Vol. 3656, pp. 108-119, 1999.
- [3]Michael Ortega, et al., Supporting Ranked Boolean Similarity Queries in MARS. IEEE Trans. on Knowledge and Data Engineering, Vol. 10, No. 6, Nov., 1998.
- [4]ByoungChul Ko, Jine Peng and Heyran Byun, Region-Based Image Retrieval Using Probabilistic Feature Relevance Learning. Pattern Analysis and Applications Journal (to be published in Summer, 2001)