

# 내용기반 검색을 위한 분할된 영상객체간 유사도 판별

유현우<sup>0</sup>, 장동식  
 고려대학교 산업시스템정보공학과  
 (paulyhwh, jang)@korea.ac.kr

## Computing Similarities between Segmented Objects in the image for Content-Based Retrieval

Hun-Woo Yoo<sup>0</sup> and Dong-Sik Jang  
 Dept. of Industrial Systems and Information Engineering, Korea University

### 요약

본 논문에서는 내용기반 영상검색중 객체기반검색 방법에 대해 다룬다. 먼저 색상과 질감정보가 동일한 영역을 VQ알고리즘을 이용해 군집화 함으로써 동일한 영역을 추출하는 새로운 영상분할기법을 제안하고, 분할 후에 분할에 사용된 색상과 질감정보, 객체간의 위치정보와 영역크기정보를 가지고 객체간 유사도를 판별하여 영상을 검색한다. 이 때 사용되는 색상의 범위를 몇 개의 주요한 색상으로 표시하기 위해 색상테이블을 사용하고 인간의 인지도에 의해 다시 그룹화 함으로써 계산량과 데이터저장의 효율성을 높인다. 영상검색시에는 질의 영상의 관심객체와 비교대상이 되는 데이터베이스 영상의 여러 객체와의 유사성을 판단하여 영상간의 유사도를 계산하는 일대다 매칭 방법(One Object to Multi Objects Matching)과 질의 영상의 여러 객체와 데이터베이스영상의 여러 객체간의 유사도를 판단하는 다대다 매칭 방법(Multi Objects to Multi Objects Matching)을 제안한다. 또한, 제안된 시스템은 고속검색을 실현하기 위해 주요한 색상값을 키(Key)색인화 해서 일치가능성이 없는 영상들은 1차적으로 제거함으로써 검색시간을 줄일 수 있도록 했다.

### I. 서론

컴퓨터와 통신기술의 발달에 따른 많은 양의 영상데이터가 생성되고 저장됨에 따라 이러한 데이터를 효율적으로 저장, 검색하는 틀의 개발이 필요하게 되었다. 본 논문에서는 내용기반영상검색방법을 다룬다.

초기에 개발된 영상검색시스템은 주로 색상, 질감, 모양정보의 전역적 사용을 통해 검색이 이루어졌다. 그러나 이러한 방법은 영상내의 중요한 객체에 대한 정보와 객체간의 상호관계를 고려하지 않으므로 해서 보다 정확한 검색을 지원하지 못하고 있다. 따라서 영상의 지역적정보도 표현할 수 있도록한 연구가 주목받고 있다[1,2,3]. 본 논문에서는 이러한 객체기반검색방법에 대해 다룬다. 먼저 색상과 질감정보가 동일한 영역을 VQ(Vector Quantization) 알고리즘[4]을 이용해 군집화 함으로써 동일한 영역을 추출하는 새로운 영상분할기법을 제안하고, 분할 후에 분할에 사용된 색상과 질감정보를 각 객체에서 추출하여 이 정보를 가지고 객체간 유사도를 판별하여 영상을 검색한다. 이 때 사용되는 색상의 범위를 몇 개의 주요한 색상으로 표시하기 위해 색상테이블을 사용하고 인간의 인지도에 의해 다시 그룹화 함으로써 영상분할시의 계산량을 줄이고 색인할 때 데이터저장의 효율성을 높인다. 영상검색시에는 질의 영상의 관심객체와 비교대상이 되는 데이터베이스 영상의 여러 객체와의 유사성을 판단하여 영상간의 유사도를 계산하는 일대다 매칭 방법과 질의 영상의 여러 객체와 데이터베이스영상의 여러 객체간의 유사도를 판단하는 다대다 매칭 방법을 제안한다. 또한, 제안된 시스템은 고속검색을 실현하기 위해 주요한 색상값을 키(Key)화 해서 유사성이 없는 영상들은 1차적으로 제거함으로써 검색시간을 줄일 수 있도록 했다.

### II. 영상 분할

영상을 표현하는데 있어서 첫 번째 단계는 영상을 기본적인 요소, 즉 객체 구성 요소들의 집합으로 간단하게 나타내는 것이다.

### 1. 전처리 과정

내용기반 영상검색측면에서도 영상 내에 존재하는 잡음은 영상을 적절히 표현하는데 방해가 되므로 본 연구에서는 5×5의 블록을 원영상에 적용하여 R, G, B 각각의 채널에 대해 중간필터(Median Filter)를 적용하여 제거한후 모폴로지 기법을 통해 영상의 세밀한 부분을 축약적으로 표현하였다.

### 2. 색상

축약적 영상분할을 위해서 영상내의 색상을 몇 개의 주요한(Dominant)한 색상만으로 표현하는 것이 중요하다. 본 논문에서는 검색의 적합성을 유지하면서 적절한 색상수를 표현하기 위해 표 1과 같은 기본 색상표를 설계한후 이 27개의 색상에 대해 다시 인간의 인지도에 의해 그룹핑을 통해 유사한 색상을 모은다(그림 1).

표1 기본색상표

Color	R	G	B	Color	R	G	B
Black	0	0	0	White	128	128	255
DarkBlue	0	0	128	LightGreen	128	255	0
Blue	0	0	255	PaleGreen	128	255	128
DarkGreen	0	128	0	LightCyan	128	255	255
Teal	0	128	128	Red	255	0	0
LightBlue	0	128	255	Magenta	255	0	128
Green	0	255	0	Orange	255	128	0
SpringGreen	0	255	128	Pink	255	128	255
Cyan	0	255	255	LightMagenta	255	128	255
Brown	128	0	0	Yellow	255	255	0
Violet	128	0	128	LightYellow	255	255	128
Magenta	128	0	255	White	255	255	255
BlueViolet	128	128	0	Grey	128	128	128



그림 1. 기본색상표(a)와 그룹핑된 색상표(b)

3. 질감표현(Texture Representation)

질감(Texture)정보를 각 영상마다 7×7 블록을 씌워서 그 변화의 유형을 조사함으로써 추출하였다.

STEP 1:

$$E_{hor(i,j)} = E_{hor(i,j)} + 1, \quad \text{if } \nabla I = |I_{(i+a+1,j+b)} - I_{(i+a,j+b)}| > \sigma \quad (1)$$

$$E_{ver(i,j)} = E_{ver(i,j)} + 1, \quad \text{if } \nabla I = |I_{(i+a+1,j+b)} - I_{(i+a,j+b)}| > \sigma \quad (2)$$

$$\text{단, } \begin{cases} a = (-3, -2, \dots, 2, 3) \\ b = (-3, -2, \dots, 2, 3) \end{cases}$$

수평 방향과 수직방향으로 한 화소씩 이동을 하면서 그 그레이 값의 변화를 측정한다. 만약 7×7 블록에서 사용자가 임의로 정한 임계값(Threshold σ)의 범위를 넘는다면 그 빈도값을 하나씩 증가시킨다( $E_{hor(i,j)}, E_{ver(i,j)}$ ).

STEP 2:

$E_{hor}$ 와  $E_{ver}$ 를 정규화시킨다.

$$E_{total} = (E_{hor} + E_{ver}) / 49 \quad (3)$$

4. 객체 분할(Object Segmentation)

영상내의 객체를 적절히 분할하기 위해서 VQ기법을 통해 이전에 추출한 색상과 질감의 4차원 정보를 입력 벡터로 받아 벡터공간상에 점으로 표현하고 이점들의 집합을 거리상 비슷한 부류끼리 군집화(Clustering)함으로써 영상의 객체를 분할한다(그림 2참조).

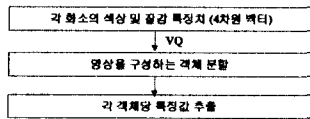


그림 2. VQ 이용시 입력 데이터와 출력

이때 사용되는 입력벡터는 요소간(R, G, B, 질감특징치)에 값의 범위가 차이가 나므로 정규화를 실시하였다. 그림 3은 원래 영상과 VQ를 이용하여 영상을 분할한 후의 영상들을 보여주고 있다.



그림 3. 원영상과 분할영상

5. 객체의 내용표현

영상을 분할한 후 각 객체를 나타내는 정보는 객체의 구분표(ID Number), 객체의 색상(Group Number), 질감정보, 객체의 크기(객체 내에 존재하는 화소의 개수가 전체영상에 비해 차지하는 비율로 나타냄), 그리고 영역의 위치정보(영역을 둘러싸는 사각형의 좌표를 사용한다)등을 이용한다. 그림 3의 독수리 영상에 대한 각 객체의 특징정보는 그림 4와 5에 나타나 있다.

객체의 ID	색상 (Group)	질감	크기 (영상/우하좌표)	위치 (영상/우하좌표)
1	5 Group	0.15	0.82 (82%)	(2,3), (196,137)
2	1 Group	0.21	0.15 (15%)	(72,6), (113,116)
3	3 Group	0.42	0.03 (2%)	(72,73), (89,83)

그림 4. 객체의 특징정보(영역크기순서대로 처음 3개의 주요영역만 표시)



그림 5. 그림 4의 각 객체의 외각경계사각형

6. 색인

주요한 색상을 키(Key)값으로 선택해 색인한다. 키 값을 선택하는 방법은 먼저 해당영상을 객체단위로 분할한 후 가장 주요한(가장 화소개수가 많은)객체 K개를 선택한다. 이 K개의 객체에서 색상값(색상 Group의 번호로 표현됨)을 추출해내고 이 색상값을 데이터베이스에 키색인(Key index)화한다. 실험에서는 K를 3으로 하였다.

III. 유사도 판별

본 연구에서는 유사도 판별에 질의 영상 내에 있는 특정한 관심객체를 데이터베이스내의 비교대상영상에서 찾는 일대다(One Object to Multi Objects) 유사도 판별과(그림 6(a)) 질의 영상내의 모든 객체와 데이터베이스내의 비교대상영상의 모든 객체와의 유사도를 판별하는 다대다(Multi Objects to Multi Objects) 판별(그림 6(b)) 알고리즘을 제안하려 한다.



(a) (b)

그림 6. 영상간의 객체간 유사도 판별( (a) 질의 영상의 관심객체(독수리)와 비교대상영상내의 모든 객체(숲, 호수, 독수리)와의 유사도 비교), (b) 질의 영상내의 모든 객체(창공, 독수리, 독수리꼬리)와 비교대상영상내의 모든 객체(숲, 호수, 독수리)와의 유사도 비교)

1. 일대다(One Object to Multi Objects)의 유사도 판별방법

STEP 1: 질의 영상에서 찾고자 하는 관심객체를 선택한다

STEP 2: 질의 영상의 관심객체에 대한 색상정보(Group 번호)를 추출하고 이 색상을 주요한 영상으로 가진 데이터베이스영상을 찾아 비교후보 영상으로 간주해서 식(4)을 이용하여 질감과 영역크기의 중복도

(Overlap)을 계산한다.

$$OL_{o,m}(a, d) = w_i \cdot \min(TEX_a, TEX_d) + w_s \cdot \min(SIZE_a, SIZE_d) \quad (4)$$

STEP 3: 가장 유사도가 높은(중복도가 높은) 객체를 질의 객체와 해당영상과의 유사도로 판별한다.  
 STEP 4: STEP 2-3의 과정을 데이터베이스내의 후보 영상의 모든 객체에 대해 실행한후 유사도가 높은 순서대로 영상을 출력한다.

**2. 다대다(Multi Objects to Multi Objects)의 유사도 판별방법**

다대다 매칭의 경우에도 일대다와 비슷한 필터링(Filtering)과정을 거친후 유사도를 비교한다.

STEP 1: 질의영상객체들과 데이터베이스영상객체들과 외각사각형이 중복되는 객체들의 쌍을 구한다.  
 STEP 2: 이 객체 쌍에서 색상이 동일한 구름에 속한 쌍만을 선택한다.  
 STEP 3: STEP 2의 객체쌍  $i$ 에 대해 식(5)과 같이 질감정보와 영역정보의 중복도(Overlap)를 계산하여 합산한다.

$$OL_{i,p}(a, d) = w_i \cdot \min(TEX_a, TEX_d) + w_s \cdot \min(SIZE_a, SIZE_d) \quad (5)$$

STEP 4: STEP 4의 과정을 모든 객체 쌍에 대해서 실시하여 결과를 합산한 후 식(6)을 사용하여 질의 영상과 데이터베이스 후보영상간의 유사도를 결정된다.

$$S_{m,m}(a, d) = \sum_p OL_{i,p}(a, d) \quad (6)$$

여기서  $p$ 는 객체 쌍의 개수.

STEP 5: 상기의 STEP 1-4과정을 질의 영상과 데이터베이스후보영상에 대해서 모두 구한 후 유사도가 높은 순서대로 내림차순으로 정렬하여 검색 결과를 보여준다.

**IV. Experimental results**

실험을 위해 사용된 영상은 모두 2,000장이며 동물, 자연풍경, 꽃등으로 구성되었다. 유사도 판단은 검색결과 유사한지의 여부를 사람이 의미적으로 파악하여 판단하였다.

**1. 일대다 매칭방법의 결과**

영상의 관심객체들을 선정하여 해당 객체가 들어있는 영상을 검색해 내도록 실험하였다. 그림 7은 치타 객체에 대해 검색한 검색결과를 보여준다

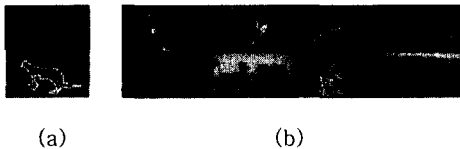


그림 7. 일대다 매칭의 검색결과: 치타 객체(a)의 질의

결과(b) (유사도가 높은 순서대로 좌에서 우로 위에서 아래로 처음 8개를 보여준다)

**2. 다대다 매칭방법의 결과**

다대다 매칭은 영상의 색상과 질감, 영역크기와 특별히 위치정보까지 유사한 영상을 검색해 내는 것이다. 그림 8에서 보듯이 검색된 영상은 모두 처음 3개의 주요한 색상이 유사하고 질감과 위치정보도 유사함을 볼 수 있다.

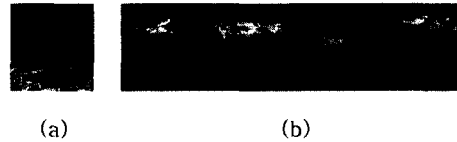


그림 8. 다대다 매칭의 검색결과 : 눈덮힌산영상 질의에 유사도가 높은 순서대로 좌에서 우로 위에서 아래로 처음 8개를 보여줌

**V. Conclusions**

본 논문에서는 영상의 객체에 기반한 검색방법을 제안하였다. 현재는 제안된 알고리즘을 보다 많은 영상을 가지고 테스트/분석중에 있으며 이러한 영상검색이 보다 효과적이기 위해 건설한 영상분할 알고리즘이 필수이므로 보다 강인한 객체 분할기법에 대한 연구가 진행중이다. 또한 객체 분할 후 객체의 모양정보를 효과적으로 표현하고 색인 하는 방법도 추후 연구과제로 생각하고 있다.

**References**

[1]Y. Chahir and L. Chen, "Searching Images on the Basis of Color Homogeneous Objects and their Spatial Relationship," *Journal of Visual Communication and Image Representation*, vol. 11, no. 3, pp. 302-326, 2000  
 [2]W. Hsu, T.S. Chua, and H.K. Pung, "Approximating Content-Based Object-Level Image Retrieval," *Multimedia Tools and Applications*, vol. 12, no. 1, pp. 59-79, 2000  
 [3]K.L. Tan, B.C. Ooi, and C.Y. Yee, "An Evaluation of Color-Spatial Retrieval Techniques for Large Image Databases," *Multimedia Tools and Applications*, vol. 14, no. 1, pp. 55-78, 2001  
 [4] A.S. Pandy and R.B. Macy, *Pattern Recognition With Neural Networks in C++*, IEEE Press, 1995  
 [5] B.M. Mehtre, M.S. Kankanhalli, A.D. Narasimhalu, and G.C. Man, "Color matching for image retrieval", *Pattern Recognition Letters*, vol. 16, no. 3, pp. 325-331, 1995