

유사한 색상과 질감영역을 이용한 객체기반 영상검색

Object-Based Image Search Using Color and Texture Homogeneous Regions

유 현 우, 장 동 식, 서 광 규
(Hun-Woo Yoo, Dong-Sik Jang, Kwang-Kyu Seo)

Abstract : Object-based image retrieval method is addressed. A new image segmentation algorithm and image comparing method between segmented objects are proposed. For image segmentation, color and texture features are extracted from each pixel in the image. These features are used as inputs into VQ (Vector Quantization) clustering method, which yields homogeneous objects in terms of color and texture. In this procedure, colors are quantized into a few dominant colors for simple representation and efficient retrieval. In retrieval case, two comparing schemes are proposed. Comparing between one query object and multi objects of a database image and comparing between multi query objects and multi objects of a database image are proposed. For fast retrieval, dominant object colors are key-indexed into database.

Keywords : image segmentation, color, texture, VQ clustering, object-based image retrieval, content-based image retrieval, key-indexed

I. 서론

최근 인터넷과 통신기술의 발달로 멀티미디어 데이터가 폭발적으로 증가함에 따라 이러한 데이터의 효율적인 저장, 검색을 가능하게 하는 관리기술이 필요하게 되었다. 본 논문에서는 멀티미디어 데이터 중에서 영상에 관심을 두고 내용에 기반해서 효율적으로 검색하는 방법을 제안하고자 한다. 보통의 내용기반 영상검색시스템은 예제의 영상을 보여주고 이 영상과 유사한 영상을 데이터베이스에서 찾아라 등의 질의를 던지면 시스템은 예제영상과 데이터베이스영상들을 내용의 유사성에 의해 비교한 후 유사도가 높은 순서대로 사용자에게 보여준다. 대표적인 시스템으로는 QBIC[1], Photobook[2], Virage[3], VisualSEEK[4], Netra[5], Blob-world[6] 등이 있다.

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. 영상의 지역적인 정보를 표현하는 최상의 방법은 영상 내에 존재하는 여러 객체들을 정확히 분할(Strong Object Segmentation)해내고 각 객체에 대한 색상, 질감, 모양 등과 객체간의 위치정보를 적절히 표현하는 것이다. 그러나 영상을 사용자가 원하는 수준으로 정확히 분할하는 것은 컴퓨터비전분야에서 가장 어려운 부분중 하나이고 그에 따른 복잡한 알고리즘과 계산량의 증가로 실제 산업에 응용되기가 쉽지 않다. 따라서 영상검색에서는 축약적인 분할방법과 (Weak Segmentation) 분할 후의 객체의 위치정보를 삽입하여 검색하는데 적용하고 있다.

따라서 영상검색에서는 축약적인 분할방법과 (Weak Segmentation) 분할 후의 객체의 위치정보를 삽입하여 검색하

는데 적용하고 있다. 이 방법은 영상 내에 객체들을 정확히 분할하는 것이 아니라 영상 내에 동일한 정보를 가지는 영역들을 하나의 데이터로 묶어 표현하여 이 영역간의 유사도를 비교하여 검색해 내는 방법이다. 이러한 방법이 가능한 이유는 내용기반 영상검색은 완전한 일치를 찾는 패턴매칭방법을 이용하는 것이 아니라 유사한 영상을 찾는 유사매칭방법을 이용하는 것이기 때문이다. 관련연구로는 한 영상을 몇 개의 서브블록으로 나누고 각 서브블록당 색상특징치를 뽑아내어 검색에 사용하는 방법[7][8]이 있고, 이 방법의 변형으로써 쿼드트리(Quad-tree)에 근간한 색상매치방법도 제안되었다[9]. [10]에서는 색상의 코렐로그램(Correlogram)에 기반한 색상위치정보를 표현하는 방법을 제안하였고, [11]에서는 영역성장방법을 통해 색상이 유사한 객체들을 분할하고 확장된 2-D 스트링을 통해 객체간의 공간적 정보를 시그니처(Signature)의 형태로 데이터베이스에 색인하고 이 값을 검색에 사용하였다.

본 논문에서는 내용기반 영상검색중 객체기반 검색방법에 대해 다룬다. 먼저 색상과 질감정보가 동일한 영역을 VQ 알고리즘[12]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. 이러한 과정에서 사용되는 색상의 범위를 몇 개의 주요한 색상으로 표시하기 위해 색상테이블[13]을 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

객체기반 검색을 위해서는 영상 내의 각각의 객체를 분리해 낸 후 각 객체의 내용을 표현함으로서 객체 단위 혹은 객체들의 집합을 통한 전체단위의 통합 검색을 가능하게 할 수 있다. 본 연구에서는 영상 내에서 불필요한 잡음 등을 제거한 후 영상의 색상과 질감정보를 추출하고, 이 정보가 유사한 영역들을 군집화(Clustering)함으로써 객체를 분할한다.

1. 전처리 과정

내용기반 영상검색측면에서 영상 내에 존재하는 잡음은 객체단위로의 분할에 부적절하므로 본 연구에서는 R, G, B 각각의 채널에 5×5 중간필터(Median Filter)를 적용하여 잡음을 제거한 후 그레이 모풀로지를 통해 영상을 보다 간략히 표현함으로써 향후 축약표현을 가능하게 하는 기초를 제공한다.

2. 색상

전통적으로 RGB 색상공간은 각각의 뚜렷한 색상을 나타내는 수많은 빈들로 나누어 히스토그램을 통해 표현하는데, 각 히스토그램에 대해 양자화를 하지 않으면 영상간의 유사도 계산시 유클리디안 거리나 인터색션을 사용할 때 많은 계산량을 필요로 한다. 그러나 실제로 내용기반 영상검색에서는 이와 같은 세밀한 색상까지 표현할 필요가 없고 축약적 영상분할을 위해서도 영상내의 색상을 몇 개의 주요한(Dominant) 색상만으로 표현하는 것이 바람직하다. 본 논문에서는 색상히스토그램의 몇 개의 뚜렷한 색상만을 고려해 유사도를 계산하려고 한다. 먼저 검색의 적합성을 유지하면서 적절한 색상수를 표현하기 위해 표 1과 같은 기본 색상표[13]를 설계해서 사용한다. 표 1은 각 채널당 3개의 색상으로 양자화하여 총 27개의 색상으로 표현한 것이다. 따라서 모풀로지 수행후의 영상에 대해 각 화소의 R, G, B값을 구한 후 이 값이 기본색상표에 있는 어느 색상과 가장 유사한지를 판단해 기본색상표에 있는 색상으로 새로 할당을 하는 것이다.

그러나, 상기의 27개의 색상도 인간의 시각으로 보면 유사한 색상이 많으므로 색상의 그룹핑을 통해 보다 색상범위를 줄인다. 그룹핑은 연구실내의 독립적인 3인을 통해 각자가 비슷하다고 생각되는 색상을 그룹핑하여 결과를 종합하여 결정하였다(그림 1).

3. 질감

대부분의 자연영상은 질감정보를 가지고 있다. 잔디나 나무 등을 포함하고 있는 영상은 색상보다는 질감에 의한 변별력이 보다 높은 경우도 많기 때문에 영상을 분할하거나 내용기반 검색에도 이러한 질감정보는 유용하게 사용될 수 있다. 본 논문에서는 각 영상마다 7×7 블록을 써워서 수평과 수직의 변화의 유형을 조사하여 질감특징을 추출한다. 제안된 질감추출방법은 아래와 같다.

단계 1 :

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

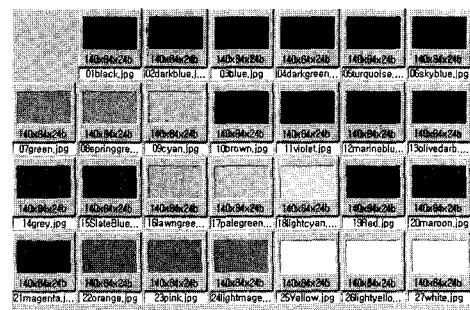
$$\begin{cases} a = (-3, -2, \dots, 2, 3) \\ b = (-3, -2, \dots, 2, 3) \end{cases}$$

7×7 블록의 각 화소에 대해 수평 방향으로 한 화소씩 이동

표 1. 기본색상표.

Table 1. Basic color table.

Color	R	G	B	Color	R	G	B
Black	0	0	0	SlateBlue	128	128	255
DarkBlue	0	0	128	LawnGreen	128	255	0
Blue	0	0	255	PaleGreen	128	255	128
DarkGreen	0	128	0	LightCyan	128	255	255
Turquoise	0	128	128	Red	255	0	0
SkyBlue	0	128	255	Maroon	255	0	128
Green	0	255	0	Magenta	255	0	255
SpringGreen	0	255	128	Orange	255	128	0
Cyan	0	255	255	Pink	255	128	128
Brown	128	0	0	LightMagenta	255	128	255
Violet	128	0	128	Yellow	255	255	0
MarineBlue	128	0	255	LightYellow	255	255	128
OliveDrab	128	128	0	White	255	255	255
Gray	128	128	128				



(a) 기본색상표

Group 1			Group 8		
Group 2			Group 9		
Group 3			Group 10		
Group 4			Group 11		
Group 5			Group 12		
Group 6			Group 13		
Group 7			Group 14		

(b) 3인에 의해 분류된 색상표.

그림 1. 색상의 그룹핑.

Fig. 1. Color grouping.

을 하면서 그 그레이 값의 변화 ∇I 가 미리 정한 임계값 σ 표의 범위를 넘는다면 수평방향의 빈도값 $E_{hor(i,j)}$ 를 하나씩 증가시킨다. 이러한 수행을 수평방향으로 블록 내에 있는 모든 화소에 수행하게 되면 한 블록내의 그레이 값의 변화정도를 알 수 있다. 여기서 i, j 는 블록에서의 각각의 화소를 나타낸다.

단계 2 :

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

수직방향(Vertical)으로 그레이 값의 변화를 측정하는 것만 다르고 단계 1과 동일하다.

단계 3 :

수직 방향과 수평방향의 E_{hor} 와 E_{ver} 의 구한 값을 모두 더한 후 0~1사이의 값으로 정규화 시키기 위하여 임계값 σ 를 넘는 수평과 수직방향의 그레이 변화 최대값인 2×49 로 나누어 재정의한다.

$$E_{total} = (E_{hor} + E_{ver}) / (2 \times 49) \quad (3)$$

4. 객체 분할

객체단위의 영상검색을 위해서는 영상내의 객체를 적절히 분할해야 한다. 본 연구에서는 VQ 기법을 통해 이전에 추출한 색상과 질감정보를 입력벡터로 받아 n차원 벡터공간상에 점으로 표현하고 이점들의 집합을 공간상 위치가 비슷한 부류끼리 군집화함으로써 영상의 객체를 분할하고자 한다.

4.1 입력벡터

한 화소당 3가지 색상 특징치(R, G, B)와 질감 특징치, 즉, 4 차원 벡터의 특징치를 추출하였는데 이것을 VQ알고리즘의 입력벡터로 사용한다. VQ는 이 입력벡터를 4차원 공간상에 근접한 정도에 따라 각 클러스터로 군집화 된다. 이 클러스터가 영상의 각 객체로 표현되고 각각의 객체에서 객체를 대표하는 특징값들을 추출함으로써 객체단위의 유사도를 판별할 수 있게 된다(그림 2). 이때 입력벡터는 요소간(R, G, B, 질감 특징치)에 값의 범위가 차이가 나므로 정규화를 실시하였다.

4.2 VQ 알고리즘

VQ[12]는 입력 데이터들을 그 분포에 따라 여러 개의 의미 있는 부류로 나눈다. VQ는 군집화 하기 전에 분류될 클러스터의 개수를 미리 정해줄 필요 없이 데이터의 분포에 적합하게 자동적으로 데이터를 분류한다. VQ를 이용하기 위해서는 군집화 과정에 사용할 클러스터간의 한계거리를 임계값으로 미리 정해주어야 하는데, 데이터의 차원과 특성, 그리고 데이터가 갖는 범위 등을 고려하여, 적절한 임계값을 실험을 통하여 결정하여 주어야 한다. 본 연구에서는 실험을 통해 가장 좋은 결과를 보이는 0.14를 임계값으로 통일하여 사용하였다. VQ에 관한 상세한 알고리즘은 [12]를 참고하기 바란다. 그림 3은 원래 영상과 VQ를 이용하여 영상을 분할한 후의 영상들을 보여주고 있다.

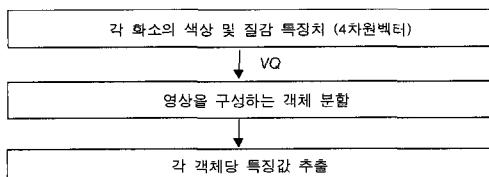


그림 2. VQ 군집화 이용시 입력과 출력.
Fig. 2. Input and output of VQ clustering.

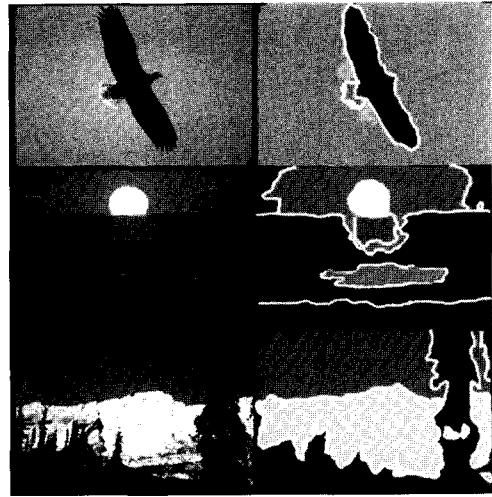


그림 3. 원영상과 분할영상.
Fig. 3. Original and segmented images.

III. 객체의 내용표현과 색인

영상을 분할한 후 각 객체를 나타내는 정보는 다음의 특징값을 이용한다. 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)등으로 나타낸다. 그림 3의 독수리 영상에 대한 각 객체의 특징정보가 그림 4에 나타나 있다.

데이터베이스 색인 시에는 주요한 색상을 키값으로 선택해 입력한다. 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 키색인화 한다. 이렇게 하면 추후 영상간의 유사도 검색 시에 이 주요한 색상이 유사한 영상만을 먼저 비교할 수 있으므로 고속검색을 가능하게 한다. 실험에서는 K를 3으로 하였다.

객체와 ID	색상 (Group)	질감	크기	위치 (좌상/우하좌표)
1	5 Group	0.15	0.82 (82%)	(2,3), (190,127)
2	1 Group	0.21	0.15 (15%)	(72,4), (113,118)
3	3 Group	0.42	0.02 (2%)	(72,73), (80,83)

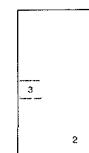


그림 4. 객체의 특징정보(영역크기순서대로 처음 3개의 주요영역만 표시)와 외곽경계사각형.
Fig. 4. Feature information of top-3 dominant objects and minimum bounding rectangles.

IV. 유사도 판별

유사도 판별을 위하여 질의 영상 내에 있는 특정한 관심 객체를 데이터베이스내의 비교대상영상에서 찾는 일대다(One Object to Multi Objects) 방법(그림 5(a))과 질의 영상내의 모든 객체와 데이터베이스내의 비교대상영상의 모든 객체와의 유사도를 판별하는 다대다(Multi-Objects to Multi-Objects) 방법(그림 5(b))을 제안한다. 일대다 방법은 질의 영상 내에 관심객체가 '새'일 경우 데이터베이스내의 영상에서 '새'가 어느 위치에 있건 상관없이 '새' 객체가 있는 영상을 찾는 것이고 다른 객체에 대해서는 관심을 갖지 않는 것이다. 그러나 다대다의 방법은 질의 영상 내에 '새'뿐만 아니라 다른 객체가 존재할 경우(예를 들면 배경인 하늘) '새'를 포함한 모든 객체의 정보를 이용하여 유사영상을 찾는 것이다.

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

단계 1 : 질의 영상에서 찾고자 하는 관심객체를 선택한다

단계 2 : 질의 영상의 관심객체에 대한 색상정보(Group 번호)를 추출하고 이 색상을 주요한 영상으로 가진 데이터베이스영상을 찾아 비교후보 영상으로 삼는다. 이 색상을 가지고 있지 않은 영상은 비교대상에서 제외된다.

단계 3 : 이제 질의 영상의 관심객체와 후보영상의 한 객체들과의 유사성을 계산할 때 두 객체간에 색상정보(그룹번호)가 같으면 아래의 식(4)를 이용하여 질감과 영역크기의 중복도(Overlap)를 계산한다.

$$OL_{o,m}(q, d) = w_t \cdot \min(TEX_q, TEX_d) + w_s \cdot \min(SIZE_q, SIZE_d) \quad (4)$$

여기서 $OL_{o,m}(q, d)$ 는 질의 객체 q 와 후보영상 객체 d 의 중복도를 나타내고, TEX_q, TEX_d 는 q 의 질감정보와 d 의 질감정보, $SIZE_q, SIZE_d$ 는 q 의 영역크기정보와 d 의 영역크기정보를 나타낸다. w_t, w_s 는 각각 질감, 영역크기에 대한 가중치를 나타내고 모두 동일한 가중치 1을 주었다(질감정보와 영역크기의 상대적인 중요도는 적용분야에 따라 달라질 수 있다. 어떤 분야에서 객체검색시 크기와 무관하게 비슷한 질감정보를 가지는 객체들이 서로 유사하다고 판단할 때는 질감정보에 보다 많은 가중치를 둘 수 있으나 본 연구에서는 일반적인 영상검색방법을 제안하는 것이므로 동일한 가중치를 주었다).

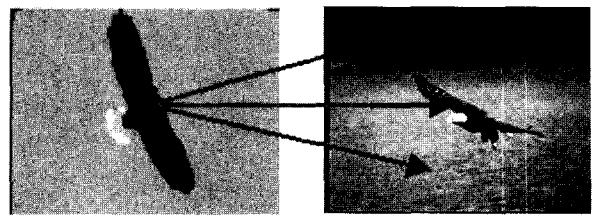
단계 4 : 단계 3을 후보영상의 동일한 색상정보(그룹번호)를 가진 객체들에 적용한 후 가장 유사도가 높은(중복도가 높은) 객체를 질의 객체와 해당영상과의 유사영상(객체)으로 판별한다.

단계 5 : 단계 3-4의 과정을 데이터베이스내의 모든 후보영상에 대해 실행한다

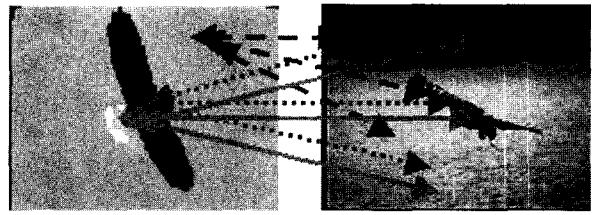
단계 6 : 최종적으로 단계 5에서 얻은 유사도에 따라 유사도가 높은 순서대로 영상을 출력한다.

2. 다대다(Multi-Objects to Multi-Objects)의 유사도 판별알고리즘

다대다 객체 매칭의 경우에는 질의 영상에 n_q 개의 객체, 데이터베이스에 있는 각각의 영상에 n_d 개의 객체가 존재하



(a)



(b)

그림 5. 영상간의 객체간 유사도 판별 (a) 질의 영상의 관심객체(독수리)와 비교대상영상내의 모든 객체(숲, 호수, 독수리)와의 유사도 비교 (b) 질의 영상내의 모든 객체(하늘, 독수리, 독수리꼬리)와 비교대상영상내의 모든 객체(숲, 호수, 독수리)와의 유사도 비교.

Fig. 5. Comparing similarity between objects (a) interest-object (eagle) of a query image and all objects(forest, lake, and eagle) of compared database image (b) all objects (sky, eagle, and eagle tail) of query image and all objects (forest, lake, and eagle) of compared database image.

고 데이터베이스에 n 개의 영상이 존재할 경우 $O(n_q \cdot n_d \cdot n)$ 의 계산과정을 거쳐야 하므로 많은 연산량이 필요하게 된다. 따라서 다대다 매칭의 경우에도 일대다와 비슷하게 유사도 계산시 필터링과정을 거친다. 먼저 1차 필터링과정으로서 질의 영상과 데이터베이스 비교 영상간에 주요한(가장 화소수가 많은) K개의 객체가 동일한 색상(동일한 색상 Group)을 가지는 영상을 찾는다(앞 절에서 K를 3으로 설정한 것을 상기할 것). 당연히 K개의 색상은 이미 색인과정에서 키값으로 정해져 있기 때문에 고속으로 가능하다. 이와 같은 과정을 거치면 n 개의 데이터베이스 영상에서 일치되는 몇 개의 영상만이 유사도 비교대상 후보가 되므로 연산시간을 줄일 수 있다. 다음으로 이 후보영상에 대해서 각 객체 단위로 질의 영상의 객체들과 유사도를 비교한 후 결과를 합산하여 검색결과를 보여준다.

단계 1 : 먼저 질의 영상과 데이터베이스의 후보영상에서 각 객체에 대해 외각사각형을 주목하여 중복되는 객체들의 쌍을 구한다.

단계 2 : Step1의 객체 쌍에서 색상이 동일한 그룹에 속한 쌍만을 선택한다. 만일 동일한 그룹에 속하지 않으면 이 두 객체간에는 유사하지 않다고 판단하여 유사도 계산에서 제외시킨다.

단계 3 : 단계 2의 객체쌍 i에 대해 (5)와 같이 질감정보와 영역정보의 중복도를 계산하여 합산한다.

$$OL_{i,i}(q, d) = w_t \cdot \min(TEX_q, TEX_d) + w_s \cdot \min(SIZE_q, SIZE_d) \quad (5)$$

여기서, $OL_{i,i}(q, d)$ 는 질의 영상의 객체 q 와 후보영상의 객체 d 인 객체쌍 i 의 중복도를 나타내고, TEX_q, TEX_d 와 $SIZE_q, SIZE_d$ 는 각각 i 에서 q 와 d 의 질감정보와 영역크기정보를 나타낸다. w_t, w_s 는 질감, 영역크기에 대한 가중치를 나타내고 일대다 판별 알고리즘과 마찬가지로 모두 동일한 가중치 1을 주었다.

단계 4 : 단계 3의 과정을 모든 객체 쌍에 적용하여 그 결과를 합산한 후 (6)을 사용하여 질의 영상과 데이터베이스 후보영상간의 유사도를 결정된다.

$$S_{m,m}(q, d) = \sum_{i=1}^p OL_{i,i}(q, d) \quad (6)$$

여기서 p 는 객체 쌍의 개수.

단계 5 : 상기의 단계 1-4과정을 질의 영상과 데이터베이스보영상에 대하여 적용한 후 유사도가 높은 순서대로 검색 결과를 보여준다.

V. 실험결과

실험을 위해 모두 2,000장의 영상을 사용하였으며 동물, 자연풍경, 꽃 등으로 구성되어 있다. 실험에서 영상분할과 검색의 적합성, 그리고 검색하는데 걸리는 시간 등을 알아보았다.

1. 영상분할

3개의 색상값과 1개의 질감 특징치를 VQ 군집화 알고리즘에 입력해 영상분할을 실행하였다. 입력매개변수로는 질감에 사용되는 임계값과 VQ 군집화에 사용되는 임계값이 있다. 본 연구에서는 현재 실험에 사용된 데이터베이스에 대하여 가장 좋은 성능을 보이는 질감 임계값 20과 VQ 군집화 임계값 0.14를 사용하였다. 물론 이러한 임계값은 분할에 사용되는 영상에 따라 다르게 설정함으로써 보다 바람직한 분할결과를 얻을 수 있다. 질감 임계값을 낮추면 보다 세밀한 그레이값의 차이까지도 질감표현에 사용할 것이고, 반대로 높이면 세밀한 부분을 무시하고 뚜렷한 밝기 변화만을 고려한다는 것이다. VQ 군집화의 경우는 임계값을 낮추면 한 객체에 대해 여러 개로 군집화 되는 경향이 있고, 반대로 임계값을 높이면 여러 객체의 특징이 한 클러스터로 군집화 될 가능성이 있다. 그림 6은 제안된 알고리즘에 의해 분할된 몇 개의 예제 영상을 보여준다.

2. 매칭방법의 결과

일대다 매칭방법은 영상의 관심객체들을 선정하여 해당



그림 6. 영상분할결과.

Fig. 6. Image segmentation results.

객체가 들어있는 영상을 검색해 내는 것이다. 그림 7을 치타 객체에 대해 검색한 검색결과를 보여준다(왼쪽의 그림은 질의영상과 관심객체를 표현하고, 오른쪽 8개의 영상은 질의에 대해 검색한 결과로 처음 8개의 가장 유사한 순서대로 좌에서 우로 위에서 아래순으로 나열하였다). 그림에서 보듯이 검색된 영상은 색상과 질감정보가 모두 유사하고 배경의 형태와 관계없이 모두 검색되었다(6번째와 8번째 영상은 다른 배경임을 주목할 것). 2번째와 4번째는 호랑이 영상이지만 색상과 질감이 모두 질의 객체인 치타와 비슷하므로 검색되었다. 현재 컴퓨터비전의 기술로는 호랑이와 치타를 구분하는 것이 쉽지 않다. 데이터베이스 내에서는 각각 10개의 호랑이와 치타영상이 있는데 처음 8개의 검색결과에 포함되지 않은 나머지 4개의 치타 영상은 배경과 치타의 구분이 모호해서 제대로 분할이 되지 않거나 객체의 색깔이 틀릴 경우, 또 크기가 배경에 비해 너무 작아서 주요한 색상으로 표현되지 못한 경우였다. 따라서 치타질의에 대한 검색의 정확도는 검색된 영상중 질의와 유사한 영상수를 검색된 영상수로 나눈 $6/8 = 0.75$ 이고, 회수율은 검색된 영상중 질의와 유사한 영상수를 질의와 유사한 데이터베이스에 있는 총 영상갯수로 나눈 $6/10 = 0.6$ 이다. 그 외 조랑말 등 다른 객체의 질의에서도 비슷한 경우의 결과(정확도 0.7이상, 회수율 0.6이상)들을 얻을 수 있었다.

다대다 매칭방법은 영상의 색상과 질감, 영역크기와 특별히 위치정보까지 유사한 영상을 검색해 내는 것이다. 그림 8과 그림 9에서 보듯이 검색된 영상은 모두 처음 3개의 주요한 색상이 유사하고 질감과 위치정보도 유사함을 볼 수 있다(각각, 정확도가 1.0, 1.0이고 회수율이 0.65, 0.53).



그림 7. 일대다 매칭의 검색결과 (a) 치타 객체 (b) 치타객체에 대한 질의 결과(유사도가 높은 순서대로 좌에서 우로 위에서 아래로 처음 8개를 보여준다).

Fig. 7. Retrieval results in comparing of One Object to Multi-Objects (a) 'cheetah' query object (b) top-8 results for 'cheetah' query in the similarity order from top to bottom and from left to right.

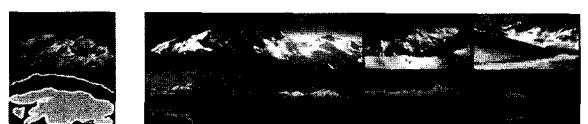


그림 8. 다대다 매칭의 검색결과 (a) 눈덮힌 산 영상의 질의 (b) 눈덮힌 산 영상의 질의 결과(유사도가 높은 순서대로 좌에서 우로 위에서 아래로 처음 8개를 보여줌).

Fig. 8. Retrieval results in comparing Multi-Objects to Multi-Objects (a) 'snow-covered mountain' query image (b) top-8 results for 'snow-covered mountain' query in the similarity order from top to bottom and from left to right.

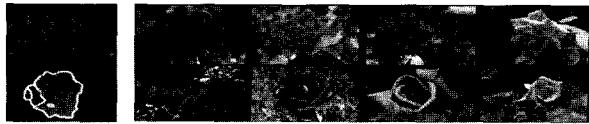


그림 9. 다대다 매칭의 검색결과 (a) 장미꽃 영상의 질의 (b) 장미꽃 영상의 질의 결과(유사도가 높은 순서대로 좌에서 우로 위에서 아래로 처음 8개를 보여줌).

Fig. 9. Retrieval results in comparing of Multi-Objects to Multi-Objects (a) 'red rose' query image (b) top-8 results for 'red rose' query in the similarity order from top to bottom and from left to right.

3. 검색시간

본 논문에서는 검색시간을 줄이기 위해서 주요한 색상을 데이터베이스에 키로 색인하였다. 주요한 색상은 그림 1에서 설명한 것과 같이 14개의 색상 그룹에서 선택되어 진다. 먼저 질의 방법으로 영상 자체가 선택되면 다대다 매칭방법이 영상의 주요 객체의 색상값이 자동 추출되고 데이터베이스에 색인되어 있는 영상 중에서 이 질의 영상과 유사한 색상을 주요 색상으로 가지는 것들만이 다음 유사도 비교에 사용된다. 그림 8의 눈덮힌 산 영상의 주요 색상은 Group 1, 2, 5이고 이러한 색상을 주요 색상으로 가지는 영상은 총 2,000 개의 영상중 25개가 존재한다. 따라서 1차 필터링과정에서 $\frac{25}{2000} \times 100 = 1.25\%$ 영상만이 다음 유사도 계산에 사용된다. 그림 9의 장미 영상의 경우는 주요 색상으로 Group 3, 6, 11이고 1차 필터링과정에서 $\frac{35}{2000} \times 100 = 1.75\%$ 의 영상만이 다음계산에 사용되었다. 본 연구에서는 이러한 관계형 데이터베이스와 유사한 방법을 영상데이터베이스에도 적용함으로써 현저하게 검색시간을 줄일 수 있었다.

질의 방법으로 영상내의 관심객체를 선택할 때에도 일대다 매칭방법과 동일한 방법으로 필터링과정을 거칠 수 있다. 그림 7의 치타객체는 색상으로 Group 9를 가지고 있고 이 Group 9를 주요 색상으로 가지는 영상은 59개가 존재한다. 따라서 $\frac{59}{2000} \times 100 = 2.95\%$ 만의 영상들이 다음 유사도 계산에 사용된다. 이 일대다 매칭방식은 검색시간의 효율면에서 다대다 방식에 비해 떨어지는데 그 이유는 다대다의 방식은 질의 영상의 3개의 주요색상모두가 일치하는 영상만이 다음 유사도 계산의 후보가 될 수 있지만 일대다의 경우는 질의 객체의 색상 1개와 일치하는 영상만이 다음 유사도 계산에 후보가 되므로 보다 많은 영상들이 후보에 오를 수 있기 때문이다.

VI. 결론

본 논문에서는 영상 객체에 기반한 검색방법을 제안하였다. 객체기반 검색을 위해서 먼저 색상과 질감정보를 이용해 영상의 축약적 분할을 하였는데 색상의 수가 증가함에 따른 계산량과 저장공간의 효율성을 위해 27개의 색상을 표시하는 색상테이블과 이를 다시 사람의 인지의 유사성에 판단해 그룹화해서 기본색상을 얻었고 질감정보로는 영상을 일정

한 수의 블록으로 나누어 각 블록당 가로방향과 세로방향의 인접하는 화소 간의 그레이 값의 차이를 누적하여 질감정보를 표현하였다. 이러한 색상과 질감을 이용해 VQ방법을 통해 군집화 함으로써 영상내의 유사한 색상과 질감정보를 가지는 객체를 추출하였다. 이후 객체기반 검색은 질의 영상의 관심객체와 데이터베이스의 비교대상영상의 객체들간의 유사도를 판별하여 가장 유사도가 높은 것을 두 영상 간의 유사도로 판단하는 일대다 매칭방법을 제안하였고, 또한 질의 영상의 모든 객체와 데이터베이스의 비교대상영상의 모든 객체간의 다대다 매칭방법을 제안하여 영상간의 유사도를 계산하였다. 제안된 방법은 고속검색을 실현하기 위해 주요한 색상값을 키색인화 해서 일치가능성이 없는 영상들을 1차적으로 제거함으로써 검색시간을 줄일 수 있도록 했다. 현재는 제안된 알고리즘을 보다 많은 영상에 적용하여 실험/분석 중에 있으며 이러한 영상검색이 보다 효과적이기 위해 견실한 영상분할 알고리즘이 필수이므로 보다 강인한 객체 분할기법에 대한 연구를 계속할 예정이다.

참고문헌

- [1] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query byimage content: The QBIC system," *IEEE Computer*, vol. 28, no. 9, pp. 23-31, September, 1995.
- [2] A. Pentland, R. W. Picard, and S. Sclaroff, "Photobook: Content-based manipulation of imagedatabases," *International Journal of Computer Vision*, vol 18, no. 3, pp. 233-254, 1996.
- [3] J. R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. C. Jain, and C. Shu, "The virage image search engine: An openframework for image management," In Proc. SPIE Vol. 2670: Storage and Retrieval for Images and Video Databases IV, pp. 76-86, February, 1996.
- [4] J. R. Smith and S. E. Chang, "VisualSEEK : A fully automated content-based image query system," in Proc. ACM Multimedia, pp.87-98, November, 1996.
- [5] W. Y. Ma and B. S. Manjunath, "Netra : A toolboxfor navigating large image databases," *Multimedia Systems*, vol. 7, no. 3, pp. 184-198, 1999.
- [6] S. Belongie, C. Carson, H. Greenspan, and J. Malik, "Color and texture-based image segmentation using EM and its application to content-based image retrieval," in Proc. International Conference on Computer Vision, pp. 675-682, January, 1998.
- [7] C. Faloutsos, M. Flickner, W. Niblack, D. Petkovic, W. Equitz, and R. Barber, "Efficient and effective querying by image content," *Technical Report, IBM Research Report*, 1993.
- [8] T. S. Chua, K. L. Tan, and B. C. Ooi, "Fastsignature-based color-spatial image retrieval," in Proc. IEEE Conf. on Multimedia Computing and Systems, pp. 362-369, 1997.
- [9] H. Lu, B. Ooi, and K. Tan, "Efficient image retrieval by

- color contents," in *Proc. of the 1994 International Conference on Applications of Databases*, pp. 95-108, 1994.
- [10] J. Huang, S. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlogram," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 762-768, 1997.
- [11] 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.
- [12] A. S. Pandya and R. B. Macy, "Pattern recognition with neural networks in C++," *IEEE Press*, 1995.
- [13] 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.



서 광 규

1969년 5월 25일생, 1997년 동국대학교 산업공학과 졸업. 고려대학교 산업공학과 석사(1999), 고려대학교 산업공학과 박사수료(2001). 현재 한국과학기술연구원 시스템연구부 CAD/CAM연구센터 연구원. 관심분야는 정보시스템, CAD/CAM, 인공지능, 멀티미디어.

유 현 우

제어·자동화·시스템공학 논문지 제7권, 제4호, 참조.

장 동 식

제어·자동화·시스템공학 논문지 제7권, 제4호, 참조