

博士學位論文

칼라영상의 내용을 기반으로 한 영상검색

嶺南大學校 大學院

電子工學科

情報通信專攻

張 正 東

指導教授 李 泰 鴻

2001年 12月

博士學位論文

칼라영상의 내용을 기반으로 한 영상검색

指導教授 李 泰 鴻

이 論文을 博士學位 論文으로 提出함

2001年 12月

嶺南大學校 大學院

電子工學科 情報通信專攻

張 正 東

張正東의 博士 學位論文을 認准함

審査委員_____ (印)

審査委員_____ (印)

審査委員_____ (印)

審査委員_____ (印)

審査委員_____ (印)

2001年 12月

嶺南大學校 大學院

感謝의 글

망설임과 설레는 마음을 멀리하고 不惑의 나이에 大學院生活을 시작하여 이제는 하늘의 命을 안다는 知命의 50대가 되었습니다. 늦게 시작한 學文이었기에 더욱 더 苦悶하고 아파했던 순간들과 술한 事緣들이 走馬燈처럼 지나갑니다.

어려운 삶 속에서도 恒常 孝道해야 한다는 마음을 간직하면서도 여전히 不孝로 남아있는 小生에게 오늘의 健康한 삶이 있게 해주신 年老하신 어머님과 世上을 떠나 계신 아버님께 이 論文을 바칩니다.

苦難과 逆境이 있을 때마다 恒常 곁에서 따뜻한 마음과 밝은 微笑로 뒷바라지를 아끼지 않으며 어려움을 함께 한 사랑하는 아내 朴 雲子에게 이 조그마한 영광을 바칩니다.

大學校 卒業 후 進路 문제로 가슴 아파하더니 이제는 어엿한 警察官이 된 귀여운 딸 志映이와 머리만 믿고 努力이 不足하다고 엄마한테 매번 혼나더니만 지금은 前方에서 國防의 義務를 다하고 있는 사랑하는 우리 玟赫이에게 앞날에 榮光이 있기를 祈願하면서 이 기쁨을 함께 나누고자 합니다.

挫折할 때면 언제나 옆에서 激勵와 指導를 해주신 따뜻한 마음을 가지신 李 泰鴻 指導教授님께 머리 숙여 感謝드립니다. 그리고 審査過程에서 많은 助言과 指導를 해주신 南 孝惠 教授님, 南 泰喆 教授님, 朴 楨日 教授님과 慶北大學校 컴퓨터工學科 金 恒俊 教授님께 깊은 感謝를 드립니다.

大學院生活동안 함께 同苦同樂하며 어려운 環境에서도 恒常 힘이 되어준 映像通信 研究室 여러분의 고마움을 잊지 않겠습니다. 또한 바쁜 業務중에서도 勉學의 雰圍氣를 이끌어 준 職場 先後輩 同僚들에게도 感謝드립니다.

삶이 곧 배움의 연속이라는 敎訓을 항상 가슴에 지니며 오늘도 생각하며, 깨어 있는 姿勢로, 나에게 주어진 일에 最善을 다 할 것이며, 謙遜하며, 貪慾스럽지 않은 삶을 살아가도록 努力하겠습니다.

2001년 12월

張 正 東

목 차

목 차.....	i
그림목차.....	iii
표 목 차.....	vi
I. 서 론.....	1
1.1 연구배경 및 목적.....	1
1.2 논문의 구성.....	3
II. 내용기반 영상검색.....	5
2.1 관련 연구.....	5
2.2 내용기반 영상검색시스템 개요 및 구성.....	10
2.3 질의 · 검색방법.....	13
2.4 대표적인 영상검색시스템.....	19
III. 영상의 특성표현.....	24
3.1 색 상.....	24
3.2 질 감.....	34
3.3 형 태.....	36
IV. 색상-공간정보를 고려한 영상검색.....	38
4.1 순차영역분할.....	40
4.2 거리 히스토그램과 특성벡터.....	47
4.3 유사도 측정과 순위결정.....	49

4.4 성능평가.....	55
V. 실험결과 및 분석.....	61
VI. 결 론.....	87
참고문헌.....	89
영문초록.....	100

List of Figures

Fig. 2.1	Schematic configuration of typical image retrieval system.....	11
Fig. 2.2	Internet connection of image retrieval system for networking.....	12
Fig. 2.3	Illustration of thumbnail browsing.....	17
Fig. 2.4	Representation of 2-D and 3-D display for retrieved results.....	18
Fig. 3.1	Grey level histogram for RGB channel. (a) Original image (b) R channel (c) G channel (d) B channel.....	25
Fig. 3.2	The RGB cube for additive devices, such as CRT displays.....	26
Fig. 3.3	The computation of the connected components for CCV method.....	28
Fig. 3.4	The connected components and CCV table.....	29
Fig. 3.5	Illustration of three scales of Munsell system.....	31
Fig. 3.6	Munsell color system, (a) organization the colors of the Munsell Value and (b) organization the colors of the Munsell Hue.....	31
Fig. 3.7	The RGB derived color space (a) HSV, (b) HLS and (c) the most commonly used HSV space.....	32
Fig. 3.8	Illustration of a boundary-based representation.....	37
Fig. 3.9	Illustration of a region-based representation.....	37
Fig. 4.1	Schematic of Proposed System.....	39
Fig. 4.2	Schematic of Sequential-Clustering.....	43
Fig. 4.3	Sample landscape image.....	44
Fig. 4.4	1 st clustered region using Sequential-Clustering method.....	45
Fig. 4.5	2 nd clustered region using Sequential-Clustering method.....	45
Fig. 4.6	3 rd clustered region using Sequential-Clustering method.....	46

Fig. 4.7	Composite clusters of each clustered regions.....	46
Fig. 4.8	Distance Histogram of a clustered region.....	48
Fig. 4.9	Matching for information retrieval.....	51
Fig. 4.10	Window for displaying of retrieved result images.....	54
Fig. 4.11	Recall vs. Precision.....	56
Fig. 4.12	Illustration of Recall vs. Precision.....	59
Fig. 5.1	Initial menu display of image retrieval system considered.....	69
Fig. 5.2	Clustered image samples using the Sequential-Clustering. (a) Original (b) Composited Clusters (c) 1 st cluster (d) 2 nd cluster (e) 3 rd cluster.....	70
Fig. 5.3	Results of Precision vs. Recall. (a) "y0021" (b) "y0174" (c) "y0175" (d) "y0229" (e) "y0237" (f) "y0407".....	71
Fig. 5.4	Retrieved results for "Starfish (y0096)" image. (a) Proposed System (1/2/3) (b) Swain (1/29/152) (c) CCV (1/44/70) (d) Correlogram (1/3/5).....	73
Fig. 5.5	Retrieved results for "Fish (y0089)" image. (a) Proposed System (1/2/3/4/6/10) (b) Swain (1/2/7/8/3/66) (c) CCV (1/2/15/8/3/80) (d) Correlogram (1/2/3/16/4/22).....	75
Fig. 5.6	Retrieved results for "Butterfly (y0154)" image. (a) Proposed System (1/2/3/4/5/6/7/9/10) (b) Swain (1/2/3/6/8/10...) (c) CCV (1/2/4/9..) (d) Correlogram (1/2/3/4/6/..).....	77
Fig. 5.7	Retrieved results for "Flower (y0158)" image. (a) Proposed System (1/2/3/4/9) (b) Swain (1/3/5/11/235) (c) CCV (1/2/4/17/229) (d) Correlogram (1/2/3/11/142).....	79

Fig. 5.8 Retrieved results for "Butterfly (y0177)" image.

(a) Proposed System (1/2/3/4) (b) Swain (1/2/59/91)	
(c) CCV (1/2/3/147) (d) Correlogram (1/2/13/30).....	81

Fig. 5.9 Retrieved results for "Landscape (y0197)" image.

(a) Proposed System (1/2/3) (b) Swain (1/38/67)	
(c) CCV (1/3/7) (d) Correlogram (1/2/4).....	83

Fig. 5.10 Retrieved results for "Flower (y0373)" image.

(a) Proposed System (1/2/3/4/5/6/7/8/9/10) (b) Swain (1/2/4/5/7/..)	
(c) CCV (1/2/3/4/5..) (d) Correlogram (1/2/3/4/5/6/7/8/..).....	85

List of Tables

Table 2.1	Typical academic Image Retrieval System.....	23
Table 4.1	Sequential-Clustering algorithm.....	44
Table 4.2	Algorithm to compute a Distance Histogram.....	49
Table 5.1	Classification of image database.....	62
Table 5.2	Methods used in the evaluation.....	64
Table 5.3	Performance comparison by <i>Eff.</i>	65
Table 5.4	Retrieved results for each query image.....	67

칼라영상의 내용을 기반으로 한 영상검색

張 正 東

嶺南大學校 大學院

電子工學科 情報通信專攻

(指導教授 李 泰 鴻)

요 약

영상의 이름이나 주석을 기억하지 않아도 필요한 영상을 검색할 수 있으며, 주제어로 표현할 수 없는 영상의 고유한 특성까지 검색에 이용할 수 있는 영상의 내용에 기반 한 접근방법은 영상의 특징을 자동으로 추출할 수 있으며, 영상의 속성을 근거로 유사한 영상을 검색할 수 있다. 그리하여 본 논문은 칼라영상의 내용을 기반으로 한 영상검색 방법에 대하여 연구하였다.

내용기반 영상검색 기법의 연구 중에서 가장 많이 사용되고 있는 히스토그램 방법은 영상이 다르더라도 유사한 히스토그램을 가질 수 있으므로 검색되는 영상의 개수가 많아지면 질의영상과는 전혀 유사하지 않는 영상을 검색하였다. 이러한 히스토그램 방법에서 검색성능을 저하시키고 있는 요인은 히스토그램 특성상 영상 자체에서 지니고 있는 개체의 위치 또는 개체간의 공간적인 정보를 포함할 수 없다는 문제점이었다. 본 논문은 이와 같이 기존의 히스토그램 방법에서 공간정보를 포함할 수 없다는 문제점을 해결하기 위하여 영상의 개체를 분할하며, 분할된 영역에서 색상과 공간정보를 함께 얻을 수 있는 방법을 제안하였다.

제안한 내용 중에서 영상의 개체를 분할하는 방법으로는 영상 내의 유사한 색상을 가지는 화소들을 같은 그룹으로 군집하는 패턴인식분야에서 사용하고 있는 순차영역분할 기법을 영상검색에 도입하였으며, 분할된 영역에서 특성벡터를 추출하는 방법으로는 영역의 중심점으로부터 화소들의 떨어진 정도

를 구분하여 거리히스토그램을 이용함으로써 공간정보를 확보하였다. 그리고 색상정보는 각 각의 영역에서 포함하고 있는 색상의 평균값을 특성벡터로 색인함으로써 시스템의 구현이 용이하고, 계산이 단순한 색상-공간정보를 함께 고려할 수 있는 영상검색 방법을 제안하였다.

제안한 방법의 유효성을 증명하기 위하여 주관적인 방법과 객관적인 평가 도구를 병행 사용하였으며, 비교 대상으로는 기존의 연구 중에서 색상정보만을 고려한 Swain의 히스토그램 인터섹션 방법과 색상과 공간정보를 함께 포함한 Pass 등이 제안한 CCV 및 Huang의 Correlogram 방법을 이용하였다.

제안한 방법에서 색인하는 특성벡터는 색상평균값 9개와 거리히스토그램 9개를 합하여 총 18개의 요소를 사용하고 있는 반면, 다른 방법인 Swain에서는 2048, correlogram은 256, 그리고 CCV는 128개의 특성벡터를 사용하고 있음에도 불구하고, 제안한 방법이 다른 방법보다 작은 저장공간을 가지면서도 *Eff*에 의한 성능평가에서 5위, 10위, 15위 및 20위의 순위에서 모두 성능이 우수하였다. 특히 순위 20위의 검색결과에서 Swain은 69.16 %, CCV는 75.49 %, Correlogram은 83.14 %로 나타나고 있으나, 제안한 방법은 91.65%의 성능을 보였다. 또한 Precision vs. Recall 평가에서도 각 질의영상에서 대부분의 Recall 값에서 Precision의 값이 상위를 점하고 있어 제안한 방법이 모두 우수함을 알 수 있었으며, 주관적인 평가에서도 제안한 방법의 우수성을 영상표시 윈도우 창을 통해 시각적으로 확인할 수 있었다.

I. 서론

1.1 연구배경 및 목적

정보화사회는 산업사회와는 달리 정보가 사회를 지배하는 도구로서 사용되며, 정보를 여하한 방법으로 이용하는가에 따라 사회의 지배구도를 바꾸게 할 것이다. 정보화사회에서 필연적으로 따르는 정보통신기술의 발달은 인터넷, 인트라넷 등의 대중화를 가져왔으며, 인터넷의 대중화는 문자뿐 아니라 음성, 영상 등 멀티미디어 정보를 급속하게 증가시키고 있다. 이러한 멀티미디어 정보는 정지영상과 컴퓨터에서 생성하는 합성영상, 애니메이션, 카메라를 이용해서 얻을 수 있는 비디오 및 음성, 음향 등과 같은 오디오 자료를 포함한다.

멀티미디어 정보 중에서 영상은 특정자료를 검색하고자 할 경우 영상자체가 문자에 비하여 용량이 큰 특성을 가지며, 또한 비정형적이어서 신속하고 효율적인 검색이 용이하지 않은 점이 있다. 즉, 영상은 포함하고 있는 내용이 복잡하여 기존의 질의방법인 주석에 의한 검색방식은 원하는 자료를 찾아내기가 쉽지 않으며, 주제어 및 주석을 효과적으로 관리하기에도 커다란 어려움이 있다. 이러한 관점에서 영상을 효과적으로 관리하기 위한 새로운 검색기법이 필요하게 되었으며, 더불어 영상을 효과적으로 전송·저장·관리하는 기술이 세계적으로 중요한 핵심 연구과제로 다양하게 연구되고 있다.

영상검색에 관한 연구는 사용자의 간단한 명령으로 원하는 영상자료를 신속하게 찾아 볼 수 있도록 하는데 초점을 두고 있다. 영상자료의 효율적인 검색을 수행하기 위한 접근방법들은 문자기반과 의미론기반 및 내용기반의 세가지로 나누고 있다[PASS96a, TAO99].

문자에 기반을 둔 접근방법은 주제어를 이용하여 영상을 찾는 방법으로서 각각의 영상에 대하여 주제어를 수동으로 생성하여 문자로 기술한 뒤, 이 문자를 중심으로 영상을 검색한다. 하지만, 각 영상의 적절한 주제어 또는 주석을 사용자가 부여해야 하므로 사용자에게 따라 한 영상에 대한 주제어 또는 주석이 다를 수 있으며, 주제어를 설정할 때 개인의 주관이 개입될 수 있다. 또

한 사용자가 주어진 영상에 대한 고유의 주제어 또는 주석을 기억하고 있어야 하며, 수작업에 의존하므로 방대한 영상자료들을 취급하기가 불가능하고, 사용자의 관점과도 상이하여 검색이 비효율적으로 된다.

의미론기반 검색은 영상을 추상화된 특징들로 표현하는 방법으로 질의어의 구문이나 의미를 분석하여 그 분석된 내용을 특징으로 사용하여 검색을 시도하는 고차원적인 방법이다. 그러나 특징을 추출하기 위한 방법개발이 적용 분야에 따라 달라질 수 있어 제한된 환경에서 사용하고 있다.

내용기반 영상검색은 영상의 내용요소들인 색상[SWAIN91], 질감[GUIBAS95], 물체의 모양[SCASSELLATI94] 및 공간적 관계[CHANG87, HOU92] 등과 같은 영상의 특징과 속성을 이용하여 영상자료간의 유사성을 계산하고, 계산된 유사도(similarity distance)를 바탕으로 영상을 검색하는 방법이다. 이 방법은 영상의 특징과 속성을 자동으로 추출한 특성벡터를 색인정보로 저장하는 시스템을 구현할 수 있다는 장점이 있다[PASS96b]. 이와 같이 영상검색을 수행하는 접근방법 중에서 최근에는 자동적으로 색인정보를 추출할 수 있고, 유사도에 의한 질의를 수행할 수 있으며, 그 결과로 유사한 영상을 찾을 수 있는 내용기반 영상검색에 관한 연구가 활발히 진행되고 있다. 이러한 내용기반 영상검색기법은 과학 데이터베이스, 영상·화상·음향 데이터베이스 및 DNA·유전자(genome) 데이터베이스와 같은 다양한 응용분야에서 효율적이고도 중요한 기술로 사용될 수 있다.

내용기반 영상검색을 수행하기 위해 영상이 지니고 있는 다양한 특징 중에서 영상의 전체적인 색상 히스토그램과 모멘트들을 이용하면 다른 정보와 비교하여 상대적으로 특성을 쉽게 추출할 수 있을 뿐 아니라, 영상의 크기나 방향의 변화에 강한 성질을 지니고 있어 영상검색에 매우 유용한 정보를 제공한다. 하지만, 이러한 방법은 영상내의 공간정보를 충분하게 고려하지 못하고 있어, 유사한 히스토그램을 가진 다른 영상을 검색하게 되므로 영상의 수량이 많아지면 검색결과가 좋지 않게 되는 경향이 있었다[STRICKER94b, 95].

본 논문에서는 히스토그램 방법의 단점인 공간정보가 부재하여 발생하는 문제를 해결하기 위해 영상을 유사한 색상을 가진 영역으로 분할하여 분할된 영역에서 특성을 계산하는 방법으로 계산량이 매우 적을 뿐 아니라 색인정보

의 수가 적어 저장공간이 작은 장점을 지닌 색상-공간정보를 함께 고려하기 위한 영상검색방법을 제안하였다. 여기에서 영상의 영역분할방법은 패턴인식 분야에서 사용되던 방법으로 영역의 개수를 사용자가 임의로 지정할 수 있으며, 처리속도가 빠른 장점을 지닌 순차영역분할방법을 영상검색에 처음으로 도입하였고, 분할된 영역에서 영상의 특성을 추출하는 방법으로는 해당영역의 중심점으로부터 계산한 거리 히스토그램을 특성벡터로 제안하였다.

1.2 논문의 구성

본 논문은 연구배경 및 목적에 이어 다음과 같이 기술하였다.

2장에서는 기존의 내용기반 영상검색에 관한 연구에 대하여 분석하였다. 본 연구분야의 핵심기술인 특성추출방법에 대한 연구와 색상 히스토그램기법의 단점인 공간정보 부재문제를 해결하기 위해 지금까지 연구되어온 논문을 중점적으로 분석하였다. 또한 내용기반 영상검색시스템의 구성 및 각 구성요소에 대한 기능, 그리고 사용자 인터페이스를 통해 시스템에 접근하는 방법인 질의·검색 방법에 대하여 서술하였다.

3장에서는 영상의 특징 또는 속성을 표현하기 위한 색상과 질감, 그리고 물체의 모양 등을 영상검색문제에 초점을 맞추어 연구하였다.

4장에서는 색상-공간정보를 고려하기 위한 순차영역분할기법과 분할된 영역의 중심점으로부터 거리 히스토그램을 제안하였다.

그리고 본 논문에서 적용하고 있는 색인된 특성벡터를 사용하여 벡터간의 거리를 계산하는 유사도 측정방법과 시스템의 효율성을 평가하기 위한 성능평가방법에 대한 연구내용을 기술하였다.

5장에서는 본 논문에서 제안하고 있는 방법에 의해 제작된 시스템을 기존의 방법들과 비교실험·분석하였으며, 끝으로 6장에서 결론으로 마무리하였다.

II. 내용기반 영상검색

내용기반 영상검색시스템은 무수히 많은 영상자료에서 효과적이면서도 효율적으로 정보를 이용할 수 있어야 하며, 사용자의 이용측면에서 데이터베이스에 친숙하지 않더라도 영상의 내용에 근거하여 관련 있는 영상을 검색할 수 있어야 한다[GUVIVADA95].

본 장에서는 이러한 내용기반 영상검색시스템에 관하여 기존에 연구되어 왔던 방법들을 분석하였으며, 또한 이러한 시스템은 어떻게 구성되어 사용되고 있는지에 대하여 기능별로 연구한 내용을 기술하였다. 그리고 사용자 인터페이스에서 사용자가 질의를 수행하는 방법과 현재 개발되어 사용되고 있는 내용기반 영상검색시스템에 대하여 용도별(상업 및 학술적인 용도)로 나누어서 기술하였다.

2.1 관련 연구

멀티미디어 응용분야의 확장과 더불어 영상이나 비디오정보는 기하급수적으로 증가하며, 이러한 정보는 미래의 필요성에 의해 적절하게 분류되고 검색될 수 있는 환경이 되어야 한다. 이와 같이 영상의 내용기반 검색기법의 연구는 매우 중요하므로 많은 사람들이 이에 대한 연구를 활발히 진행하고 있다.

영상검색분야와는 달리 패턴인식분야에서는 여러 종류의 영상 중에서 어떤 하나의 영상과 동일한 영상을 찾아내는 작업을 연구하여 왔다. 다시 말하면, 패턴인식기술은 몇 가지 표준화 된 패턴(원, 사각형 등)을 미리 정해 놓고 최근접 분류방식(Nearest Neighbor Classification)에 의해 영상의 특징을 추출하는 기법을 사용하고 있다. 그러나 일반적으로 영상자료는 매우 다양한 정보를 담고 있어 표준패턴을 정하는 것이 거의 불가능하며, 패턴인식의 대상이 되는 영상정보가 한정되어 있으므로 영상검색에 적용하기에는 많은 문제점이 있다.

Grosky 및 Mehrotra 등은 영상에 포함하고 있는 객체들의 윤곽선을 벡터화하여 이를 영상의 특징값으로 사용하였다. 영상에서 추출된 특징값들은 거리, 방향, 인접요소와 이루고 있는 각도 등을 계산하여 영상간의 유사성을 평가하였다. 검색에 소요되는 시간을 단축하기 위해 특징값들을 트리형태로 색인한다[GROSKY92, MEHROTRA95]. 이 방법은 영상내의 객체의 크기와 위치에 관계없이 적용할 수 있다는 장점은 있으나, 객체의 형태가 제한된 환경에서만 사용할 수 있다는 단점이 있다.

Hirata 및 Kato에 의한 QVE는 검색자가 그린 개략적인 스케치 또는 복사본을 기반으로 하여 영상을 검색하는 방법을 제안하였다. 이 방법은 영상을 작은 크기로 축소한 아이콘으로 특징값을 사용한다. 영상정보 저장시스템에 저장된 각 영상자료는 모양과 색채에 대한 아이콘 영상을 가지고 있으며 이를 기반으로 검색작업을 수행한다[HIRATA92]. 이들의 방법은 다양한 형태의 영상자료를 다룰 수 있다는 장점을 가지고 있으나, 영상을 검색하기 위해서 모든 영상자료를 화소단위로 순차적으로 비교해야 하는 단점을 가지고 있다.

Jacobs 등은 특성벡터의 추출과정에서 처음으로 웨이브릿 변환을 사용하였다. 그들의 연구에서는 색상영상에 대하여 직접적으로 웨이브릿 변환을 적용하여 얻은 계수들을 영상의 R, G, B 각 채널마다 따로 특성벡터를 구성하였다[JACOBS95]. 이는 특성벡터의 차원이 너무 크므로 처리시간이 많이 걸려 기존의 색인기술을 이용하는데 많은 어려움이 있었다. 또한, Milanese는 영상을 직각 좌표계에서 로그-극 좌표계로 변환한 후 이를 푸리에 변환하는 방법으로, 변환한 좌표계가 회전·이동·크기의 변화에 강한 성질을 이용하였다[MILANESE99]. 그러나 이 방법은 회전 등의 변화를 계산하기 위해 많은 시간이 필요하여 대용량의 영상시스템에는 적용하기가 곤란하다는 단점이 있다.

버클리 대학에서 개발한 Chabot은 주석정보와 영상의 특성벡터를 함께 고려한 시스템으로 특성벡터로는 색상과 히스토그램을 이용하였다[OGLE95]. 예를 들어 사용자는 『꽃』 또는 『풍경』과 같은 주제어로 질의할 수 있으며, 부가적인 기능으로 색상을 선택하여 영상을 검색할 수 있도록 하였다. 이 시스템은 내용기반 영상검색에 접근하기 위한 색상·질감·물체의 모양 등과 같은 특징들을 사용하였으나, 특성정보를 추출하기 위한 과정이 복잡하며, 계산량이

많아 처리시간이 많이 소요되고 있다.

이용할 수 있는 영상의 범위가 광범위하게 확대됨에 따라 영상 내에 있는 객체를 자동으로 처리할 뿐 아니라, 이를 정확히 추출하는 문제는 컴퓨터비전 분야에서 아직도 해결하지 못한 숙제로 남아 있다. 더구나 각 영상의 객체는 보는 이의 관점, 물체의 흡수, 영상의 변형에 따라 다르게 나타날 수 있으므로 이러한 복잡한 변수들을 고려하여야 하는 문제가 있다. 따라서 객체나 영역을 근사한 모양으로 표현하기 위해 객체 또는 영역을 사각형으로 표현하기도 한다. Hsu는 색상 히스토그램은 영상내의 색상 빈의 빈도수만으로 유사도를 측정하므로 모양이 전혀 다른 영상을 검색하는 오류를 범할 수 있음을 주장한다. 이들은 이를 개선하기 위해 영상의 엔트로피를 계산하고 추출된 영역을 사각형 폴로 표현한 후, 영역의 대표색상과 공간정보를 분석하여, 통합된 색상-공간정보에 기반 한 검색방법을 제안하였다[HSU95]. Ooi 등은 영역의 분할개수에 대한 연구를 진행하여 4개의 영역이 적당함을 재현율(recall)과 정확율(precision)의 개념을 사용하여 유효성을 증명하였다[OOI97]. Smith 등은 HSV(Hue, Saturation, Value)공간에서 색상에 의해 공간을 분할하고, 히스토그램 백-프로젝션방법을 사용하여 관심영역으로 추출된 각 영역을 이진화하여 특성을 색인하는 것으로 VisualSeek에 적용하고 있다[SMITH96a, 96c].

영상의 영역이나 객체들을 기반으로 하는 색상정보가 공간적인 분포를 고려하기 위한 방법으로 매우 유용하지만, 영상의 영역을 분할하는데 어려움이 있으므로 간단한 고정크기의 분할방법이 다양하게 이용된다. Stricker 등은 영상을 중앙의 타원형과 겹치는 영역을 포함하여 5영역으로 구분하고, 각 영역에서 히스토그램 및 1, 2, 3차 모멘트를 추출하여 비교하였다[STRICKER96]. 이들은 HSV공간에서 히스토그램과 평균색상, 분산, 왜곡을 사용하여 계산량을 줄이는 효과를 가져온 반면, 영상에 나타나는 색상정보를 상세하게 반영하지 못하는 단점이 있다. Dimai는 영상을 8×8의 블록으로 계층적으로 분할하고 분할된 영역에서 색도와 채도의 값들로 색인한다[DIMAI97].

Zhang 등은 영상의 내용상 전체적으로 색상이 비슷한 영상을 검색하기 위해 고정된 크기의 국부적인 색상 히스토그램을 사용함으로써 계산이 간단하여 다양한 응용분야에서 효과적임을 보여주고 있다[ZHANG94]. Belongie 등은 영상

을 『BlobWorld』라는 형식으로 표현하여 영상검색에 적용하고 있다. 이는 자연영상을 점(blob)으로 나누고 각 영역에서 중앙과 주변에서 나타나는 대표색상을 특성벡터로 이용하였다[BELONGIE98]. Mehtre는 클러스터링 방법에 의해 영역을 분할하고 이 영역에서 7가지의 색상 모멘트를 계산하여 복합적인 특성추출방법을 제안하였다[MEHTRE98].

영상을 고정 또는 가변블록으로 분할하지 않고 공간적인 색상분포를 표현하기 위한 다른 접근방법이 있다. Pass 등은 영상내에 포함하고 있는 같은 색상의 응집여부를 이용한 CCV(Color Coherence Vectors)를 제안하여 기존의 히스토그램 방법을 개선하였으나[PASS96a, 96b], 공간정보가 극히 적다는 단점이 있다. 또한 이들은 색상 히스토그램의 장점을 살리고 국부적인 화소들의 특성들으로써 다차원의 히스토그램을 구성하는 결합 히스토그램을 제안하기도 하였다[PASS98]. Huang 등도 공간정보를 고려하기 위한 방법으로, 주어진 색상의 화소가 임의의 거리에 존재할 확률을 계산한 테이블을 작성하여 색인하는 Color Correlogram을 제안하였다[HUANG97]. Mandal 등은 색상정보를 반영하기 위해서 색상 히스토그램에 대한 모멘트를 이용하였고, 공간정보를 고려하는 방법으로는 웨이블릿 부 영상을 가우시안 분포로 근사화한 뒤, 그 계수를 이용하였다[MANDAL96]. 이 방법은 특징의 수가 적어 색인에는 효율적이지만, 물체나 배경의 큰 변화 등에 민감하게 되는 단점이 있다. 또한 Vertan 등은 기존의 히스토그램에서 윤곽성분을 부가하기 위해 라플라시안 연산자를 사용함으로써 공간적으로 구축된 색상분포를 나타내고 있으며, 확률적으로 계산하는 히스토그램 표현방식과는 달리 멤버십 함수로 그룹별 대표성을 표현하는 퍼지 히스토그램으로 영상의 내용을 표현하고 있다[VERTAN00].

이와 같이 영상이 지니고 있는 다양한 특징 중에서 영상의 전체적인 색상 히스토그램과 모멘트들을 이용하면 다른 정보와 비교하여 상대적으로 비교적 쉽게 추출할 수 있을 뿐 아니라, 영상의 크기나 방향의 변화에 강한 성질을 지니고 있어 영상검색에 매우 유용한 정보를 제공하고 있지만, 영상내의 공간 정보를 충분하게 고려하지 못하고 있어 영상의 개수가 많아지면 검색결과가 좋지 않게 되는 경향이 있다[STRICKER94b, 95]. 내용기반 영상검색에서 공간적인 구조를 함께 고려하기 위한 최근의 연구에서 색상의 분포를 더욱 상세히 표현

하기 위한 노력들이 시도되고 있다. 이러한 것들을 통틀어 Tao 등은 공간을 포함한 색상색인(spatial color indexing)기법이라 부르고 있다[TAO99b].

지금까지 내용기반 영상검색에 관련된 기존의 연구를 분석하였다. 앞서 지적한 바와 같이 색상 히스토그램은 전혀 다른 영상이지만 유사한 히스토그램을 가질 수 있으므로 이를 영상검색에 적용할 경우 사용자의 의도와는 전혀 틀리는 영상을 검색하는 결과를 가져온다. 이러한 히스토그램의 단점인 공간 정보 부재문제를 해결하기 위해 영역을 고정 또는 가변 분할하는 방법을 시도하고 있으나, 유사한 영역의 분할이 용이하지 않아 검색결과가 좋지 않았다.

본 논문에서는 이러한 문제점을 해결하기 위해 영역분할방법으로 순차영역분할기법을 도입하며, 분할된 영역에서 거리 히스토그램을 이용하는 색상과 공간정보를 함께 고려하기 위한 방법을 제안한다.

2.2 내용기반 영상검색시스템 개요 및 구성

컴퓨터에 저장되는 영상이 증가하면서 문자기반 영상검색은 한계를 나타내게 되었다. 예를 들어 풍경화가 담긴 영상을 검색할 경우, 『영상을 제작한 사람의 이름』, 또는 『고유 번호』, 『제작 년도』 등의 주석보다는 『수평선 위에 저무는 저녁노을』과 같은 식으로 영상의 내용을 직접적으로 언급하여 검색하는 것이 훨씬 더 편리할 것이다. 이러한 내용에 의한 방법을 영상검색에 적용하기 위해서는 색상과 물체의 모양과 같은 시각적인 정보를 데이터베이스에 미리 저장해 놓아야 한다.¹⁾

내용기반 영상검색시스템이란 영상 데이터베이스에 저장되어 있는 영상들로부터 영상에 포함된 내용²⁾을 기반으로 하여 검색하는 시스템을 말한다. 이러한 내용기반 영상검색시스템을 구성하기 위한 요소들을 나타내면 그림 2.1

1) 내용기반 영상검색시스템에서 질의에 대한 검색을 원활히 수행하기 위해서는 영상으로부터 정보를 추출하는 과정이 필수적이다.

2) 여기에서 포함된 내용이라 함은 영상의 전체적인 색상, 색상의 비율, 질감, 물체의 모양 등 영상의 다양한 시각적 특성을 의미한다.

과 같다. 그림에서 보는 바와 같이 내용기반 영상검색시스템은 일반적으로 전처리과정과 질의과정, 그리고 검색·출력과정으로 나누어 구성한다.

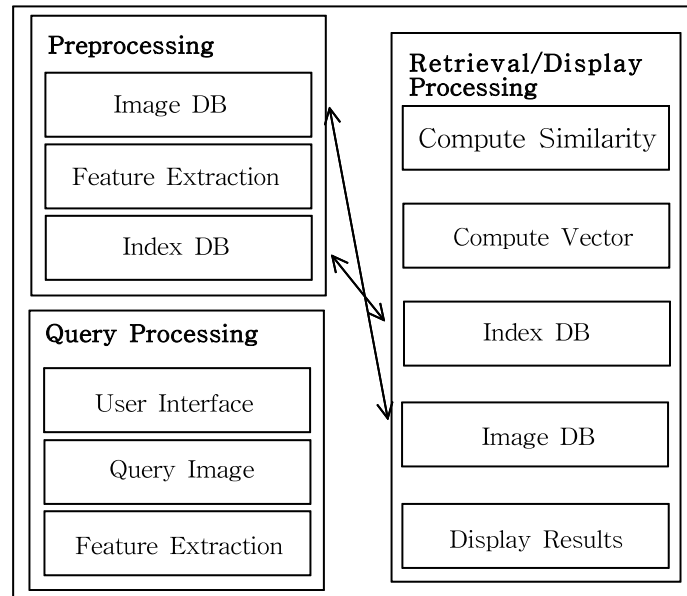


Fig. 2.1 Schematic configuration of typical image retrieval system.

전처리과정은 데이터베이스에 저장된 영상을 컴퓨터에 의한 자동적인 방법으로 영상자료의 특성을 추출하는 과정과 추출된 특성요소를 색인DB에 저장하는 과정으로 나눈다. 여기에서 가장 핵심기술인 특성추출은 영상내의 색상, 질감, 물체의 모양과 물체의 공간적인 관계들을 종합하여 벡터로 기술한다. 그리고 이러한 벡터들을 DBMS에 의해 색인정보로 관리하도록 함으로써 향후 사용자의 질의에 응답하도록 준비한다.

질의과정에서는 사용자가 사용자 인터페이스인 시스템 입력도구³⁾를 통하여 예제영상에 의한 질의 또는 사용자가 직접적으로 스케치한 영상을 입력하면 전처리과정에서 사용했던 동일한 특징추출방법에 의해 질의영상의 특성벡터가 생성된다.

3) 예, 키보드, 마우스, 그래픽 도구 등

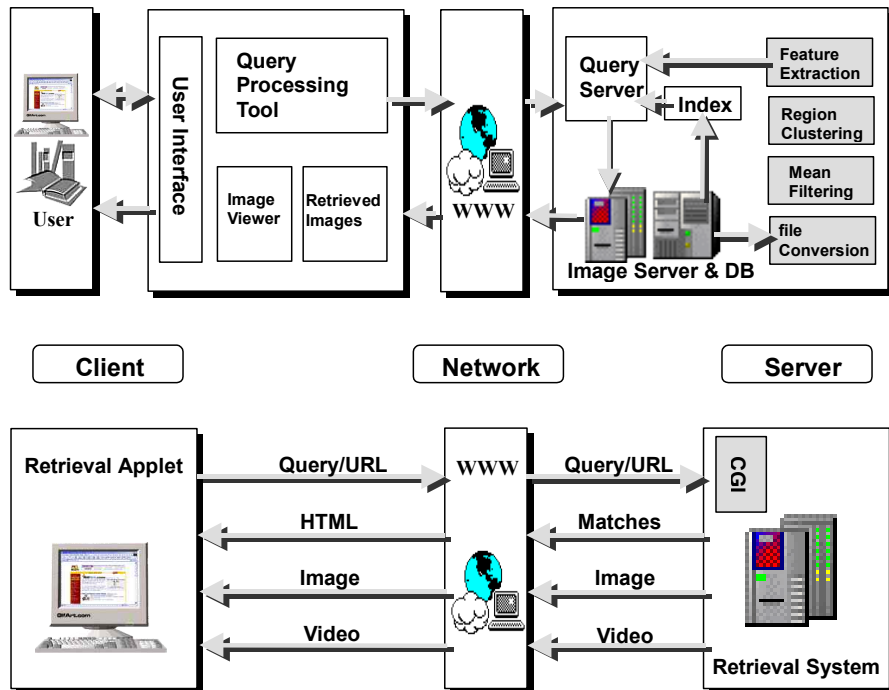


Fig. 2.2 Internet connection of image retrieval system for networking.

검색·출력과정은 데이터베이스 내에 있는 색인정보와 사용자 인터페이스에서 생성한 특성벡터를 비교·분석하는 과정으로 벡터간의 거리를 계산하고, 유사도를 측정하여 사용자가 질의한 영상과 가장 유사한 영상들을 순위를 부여하여 결과영상으로 출력하게 된다. 이러한 영상검색시스템은 네트워크를 통하지 않고 단독으로 사용할 수도 있지만 그림 2.2와 같이 인터넷을 통하여 웹상에서 원격으로 시스템에 접근하여 영상을 검색할 수 있도록 구성하기도 한다.

2.3 질의·검색방법

사용자가 요구하는 영상을 원활히 검색하기 위해서는 영상자료의 접근이

용이하여야 한다. 영상자료에 접근하기 위한 질의·검색하는 형식에는 다양한 방법이 있다. Gudivada 및 Jain 등은 이러한 여러 종류의 질의·검색방법을 잘 구분하고 있다. 이들은 "영상검색 전략"이라는 주제로 접근하고 있으며, 질의, 브라우징 및 내비게이션과 같은 용어를 사용하고 있다[GUDIVADA95, JAIN96].

2.3.1 질의방법

2.3.1.1 주제어에 의한 질의

주제어에 의한 질의는 영상검색시스템에서 가장 많이 사용하고 있는 방법이다. 디지털 카메라를 제작·판매하는 제작자 또는 디지털 사진앨범과 같은 영상을 취급하는 제작자들은 사진을 관리하기 위한 도구로서 주제어에 의한 문자기반의 검색도구를 제공하고 있다. 이러한 도구들은 미리 주제를 선정해 놓고 사용자가 사진을 분류할 때 그 주제어에 맞추어 저장되도록 한다. 주제어 분류기법은 통상적으로 다층구조를 형성하고 있어 사용자가 "풍경화"라는 메뉴를 선택하면 『사막』, 『숲』, 『호수』, 『강』 및 『바다』와 같은 부 메뉴를 보여 주어 그룹별 계층적 관리가 가능하도록 하고 있다.

2.3.1.2 예제영상에 의한 질의

사용자 인터페이스에서 다양한 예제영상을 준비해 두고 이를 선택하면 데이터베이스내에 있는 후보영상들과 서로 비교한다. 사용자가 사용하기에 편리하도록 다양하게 유사도 측정방법을 선택할 수 있도록 질의방법을 개선하고 있는 시스템들도 있다. 예제영상에 의한 질의방식은 전반적으로 다음과 같은 3가지 형식을 취하고 있다. 즉,

- i) 영상의 특성(색상, 질감, 모양 등)을 선택할 수 있는 기능 (예, QBIC)
- ii) 특성 간에 가중치를 조절할 수 있는 기능 (예, Excalibur)
- iii) 특성 및 가중치를 시스템에서 전부 결정 (예, Image Surfer)하는 방법으로, 예시로 주어지는 영상들은 데이터베이스에 저장되어 있는 영상 중 임의로 선정되거나 대표적인 특성을 갖는 영상들로 미리 정해져 있어 사용자가 질의 영상을 선택하기만 하면 별도의 특성추출과정을 거치지 않아도 되므로, 시스

템의 입장에서 보면 매우 효율적이라고 할 수 있다. 그러나 사용자가 검색하려는 목표영상과 유사한 영상이 예시로 주어진 영상 중에서 찾을 수 없는 경우에는 계속 다른 예제영상들을 제공받아 유사한 영상을 선택하거나 목표영상과 얼마간 차이가 있는 영상을 선택할 수밖에 없으므로 한 번에 만족할 만한 결과를 얻기가 어려운 단점이 있다.

2.3.1.3 영상의 특징에 의한 질의

사용자가 직접적으로 영상의 특징 또는 영상의 속성을 설정하여 질의를 수행하는 방법으로, 이 방법은 영상에 대한 사전지식이 필요하고 구성된 특성들을 이해하고 있어야 한다. 이러한 기법을 채택하고 있는 QBIC시스템을 예를 들면, 히스토그램에 의한 질의 중에서 사용자가 『17 %의 푸른색상과 15 %의 노란 색상을 지닌 영상을 찾아라.』 하는 식으로 질의할 수 있다. 현재 QBIC에서는 히스토그램방법에서 사용자의 편리성을 도모하기 위해 색상도표와 색상의 점유율을 도식적으로 나타내 주는 방법을 사용한다[FALOUTSOS95, NIBLACK93].

2.3.1.4 사용자 스케치에 의한 질의

사용자가 예제영상 대신 개괄적으로 영상의 내용을 스케치한 영상을 질의 영상으로 사용한다. 사용자 인터페이스에는 그리기 도구를 제공하여 손쉽게 스케치할 수 있도록 한다[BIMBO97, JACOBS95, KATO92, VAIL96]. 스케치하는 방법으로는 사용자가 직접적으로 윤곽을 스케치하고, 스케치한 윤곽의 내부를 색상으로 채우는 방법과, 그리드를 제공하여 원하는 형태의 그리드를 선택함으로써 스케치가 완성되는 기법을 사용한다.

사용자가 작성하는 질의영상은 각 물체의 색상, 모양, 윤곽선 정보, 물체의 배치 등 다양한 특성들을 표현할 수 있으므로 이러한 특성들을 검색에 사용할 수 있으며, 사용자가 원하는 특성을 자유롭게 표현할 수 있다는 장점이 있다. 그러나 이러한 방식의 문제점은 사용자가 작성한 영상에서 표현한 특성이 사용자의 기억이 잘못되었거나, 작성과정의 잘못으로 인하여 목표영상과 많은 차이를 보일 수 있다는 점이다. 이 방식을 이용한 시스템에서는 사용자가 작성하는 질의의 형태를 제한하거나[BIMBO97], 사용자가 작성하는 질의가 어느 정

도 목표영상에 근접한 형태로 입력된다는 가정을 전제로 개발을 진행하고 있으며, 질의에 오류가 있는 경우를 대비한 처리문제는 향후의 연구과제로 남겨 놓고 있다[JACOBS95].

2.3.1.5 메타-검색 (meta-search)

인터넷의 성장으로 웹에서 정보, 특히 영상정보를 검색하기 위한 요구가 급증하고 있으므로 전 세계에 분포되어 있는 웹사이트로부터 영상을 검색하기 위한 메타-검색엔진이 필요해진다. 콜롬비아 대학에서 개발된 MetaSeek⁴⁾는 이러한 메타-검색엔진의 대표적인 시스템으로 VisualSEEK, WebSEEK, QBIC 및 Virage의 내용기반 영상검색시스템에 의해 지원을 받으며 운용되고 있다.

2.3.2 브라우징

영상자료를 화면에 나타내기 위한 방법으로 가장 쉽게 접근할 수 있는 방법은 영상을 한 장씩 표시해 주는 방법이 있으나, 이는 비효율적이므로 잘 사용되지 않는다. 현재 일반적으로 사용되고 있는 브라우징방법은 두 가지 형태인 썸네일(thumbnail) 브라우징과 슬라이드 쇼로 영상을 출력하고 있다.

썸네일 브라우징은 한 화면에 다수의 영상을 표시하기 위해 작은 크기로 축소된 영상인 썸네일 영상을 사용한다. 이것은 화면에 나타난 영상을 클릭 하면 실제크기의 영상 또는 화면 크기로 재조정된 영상을 표시해 준다.

슬라이드 쇼는 실제크기의 영상이나 화면크기로 조정된 영상을 한 장씩 화면에 보여 주는 방법으로 사용자는 화면이 바뀌는 영상간의 슬라이드 되는 시간을 조절할 수 있도록 하고 있다. 일반적으로 영상검색시스템에서 채용하고 있는 브라우징방법은 대부분 썸네일방법을 사용하고 있는데 반해, 슬라이드 쇼방법은 특별한 용도 외에는 거의 사용하지 않고 있다.

내용기반 영상검색시스템에서 많이 사용하고 있는 썸네일 브라우징방법을 표시하면 그림 2.3과 같다. 이에는 통상적으로 화면의 이동을 용이하게 하기 위하여 『이전』 과 『다음』 과 같은 마우스 버튼을 제공한다. 또한 축소된 영

4) <http://ctr.columbia.edu/metaseek/>

상 중에서 관심이 있는 영상을 확대해 볼 필요가 있는 경우에는 해당화면을 클릭 하면 확대된 영상(원 영상)을 보여준다[LA100].

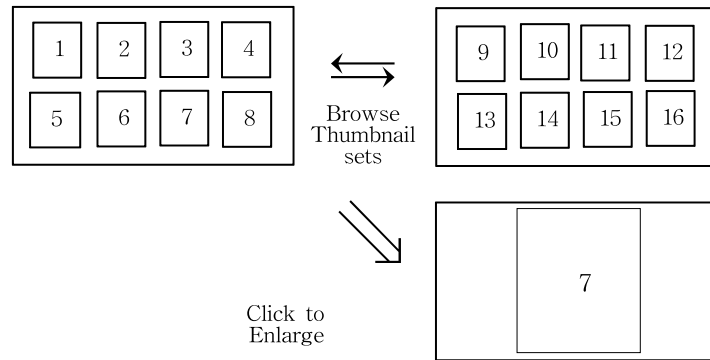


Fig. 2.3 Illustration of thumbnail browsing.

2.3.3 네비게이션

인터넷을 『정보의 보고(寶庫)』 또는 『정보의 바다』로 비유하는 것을 우리들은 자주 접하게 된다. 인터넷에서 정보를 탐색하기 위해서는 여기저기를 떠돌아다녀야 하는데, 이러한 의미에서 네비게이션이라는 용어를 사용한다. 즉, 사용자가 영상을 검색하기 위해 특정시스템을 통할 경우 인터넷의 영상정보공간을 항해할 수 있도록 고안된 것이다.

현재 사용되고 있는 전통적인 방식의 인터넷 브라우저는 직관에 의한 인식공간인 기하학적인 성질들을 자연스럽게 이끌어 내지 못하고 있으며, 유사한 영상과의 관계도 수치에 의해 순위를 결정하고 있다. 이러한 점을 보완하기 위해 네비게이션이라는 새로운 아이디어가 탄생되었으며, 이에 대한 새로운 이론이 많이 제안되고 있다. 그림 2.4에서 보는 바와 같이 데이터베이스내의 자료 또는 검색된 영상을 나타내기 위해 2-D 또는 3-D 방법을 사용하여 표현한다[LA100].

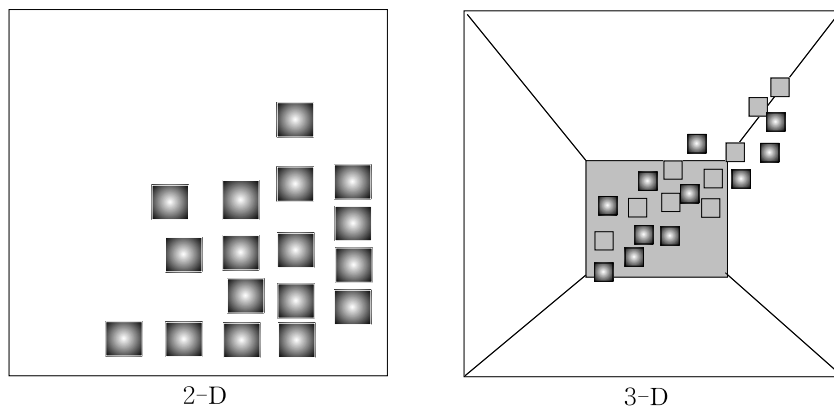


Fig. 2.4 Representation of 2-D and 3-D display for retrieved results.

2.4 대표적인 영상검색시스템

앞장에서 살펴 본 바와 같이 대체적으로 초기의 시스템들은 『예제에 의한 질의』 기법을 많이 채택하였으며, 최근에는 『사용자가 작성하는 질의』에 대한 관심이 높아져 연구가 활발히 진행되고 있다. 하지만 각 질의기법은 사용하는 특성에 따라 장단점이 있으므로 실제 시스템에서는 목적에 맞게 병행해서 사용해야 한다.

내용기반 검색기술을 이용한 다양한 종류의 검색시스템이 개발·운용되고 있다. 이러한 시스템은 상업용으로 개발된 것이 있는 반면, 순수하게 학술적인 연구목적으로 개발된 것도 있다. 본 장에서는 이러한 시스템 중 대표적으로 알려진 시스템에 대하여 분석하였다.

2.4.1 상업용 시스템

2.4.1.1 QBIC(Query by Image Content) 시스템⁵⁾

QBIC은 내용기반 영상검색시스템의 대표 격으로 상업적으로 가장 성공한 시스템 중 하나인 시스템이다. 이는 IBM Almaden 연구소에서 개발되었으며, 영상의 히스토그램이나 레이아웃, 질감 및 색상에 의한 검색이 가능하도록 기본적인 특징에 기반 한 다양한 질의를 제공한다. 현재 데모형식으로 이용할 수 있는 영상자료는 우표와 상표영상을 준비하고 있다. 또한, 데이터베이스에 저장된 영상에 대하여 사람이 의미정보를 부여할 수 있도록 주제어에 의한 검색을 허용함으로써 제한적이거나 의미정보에 의한 검색도 가능하다. 그러나 기본적인 특징과 의미정보를 표현하는 논리적인 특징이 체계적이지 못하며, 다양하고 복잡한 질의방법들이 상호 인터페이스가 통합되어 있지 않아 오히려 사용자측면에서는 이용하기가 어려운 면이 있다.

2.4.1.2 Virage : VIR Image Engine⁶⁾

Virage사에서 개발한 VIR(Visual Information Retrieval) 영상검색엔진은 영상

5) <http://www.qbic.almaden.ibm.com/cgi-bin/stamps-demo>

6) <http://www.virage.com/index.html>

의 속성을 지역적인 색상, 전체적인 색상, 영상의 구조인 모양특성과 질감으로 나누어 생각한다[BACH96]. 이 시스템은 『위성사진』과 같이 특정한 분야의 영상들을 속성에 따라 구분하여 검색할 수 있도록 하고 있다. 또한 이 시스템은 비디오검색 또는 관리에 사용할 수 있도록 기능을 확장하고 있다.

2.4.1.3 Excalibur: Visual Retrieval Ware⁷⁾

Excalibur회사에서 개발한 Visual Retrieval Ware는 디지털 영상의 특성을 관리하며, 영상검색을 위해 색상, 모양, 질감, 밝기 구조, 색상구조, 화면 비율 등을 각 항목마다 0에서 5까지 범위의 값을 주어 사용자가 선택하여 사용할 수 있도록 하고 있다. 그러나 의미정보에 의한 검색기능은 제공하지 않고 있다. 또한 이 시스템은 온라인상에서 사용자가 실험할 수 있도록 데모 사이트를 운영하고 있다. 이 기술은 야후의 Yahoo! Image Surfer에 기술을 제공하고 있으나 여기서는 각 속성들의 가중치를 조절할 수 없도록 되어 있다.

2.4.2 학술용 시스템

2.4.2.1 VisualSEEK⁸⁾

콜럼비아 대학에서 개발된 VisualSEEK[SMITH96c]는 웹을 이용한 영상검색시스템으로, 객체 또는 영역간의 공간적인 배열관계를 이용하는 Safe⁹⁾시스템과 색상, 질감과 색상의 조합으로 영상을 검색하는 CBVQ¹⁰⁾, 그리고 내용기반 영상 및 비디오 검색엔진인 WebSEEK로 구성되어 있다. 이 시스템의 가장 큰 특징은 시각적 특징에 의한 내용기반 질의와 공간정보에 의한 공간질의가 동시에 가능하다. 영상에서 색상영역에 관한 정보를 추출한 후에, 이러한 정보를 이용하여 각 색상영역간의 공간적인 관계를 고려하여 검색하는 방법으로 영역의 공간관계를 고려할 수 있다는 장점이 있다. 또한 이 시스템은 주석에 의한 검색기능도 함께 제공하고 있다.

7) <http://www.informix.com/informix/products/options/udo/datablade/dbmodule/excalibur2.htm>

8) <http://disney.ctr.columbia.edu/VisualSEEK/>

9) Spatial and Feature query system

10) Content-Based Visual Query system

2.4.2.2 MARS

MARS¹¹⁾는 일리노이 대학에서 개발한 시스템으로서, 영상의 저수준 특성인 색상, 질감, 모양 및 레이아웃 특성을 이용하였으며, 『노란 꽃』, 『호수의 일몰』과 같은 고수준의 의미정보와 논리적인 조합을 하여 검색이 가능하도록 하였다. 또한 사용자에게 편의를 주기 위해 주제어에 의한 검색기능도 함께 제공한다. 이 시스템의 주요한 특징은 검색결과로 나타난 요소들을 이용하여 더 나은 검색결과를 유도하기 위해 온-라인 상에서 결과내 질의¹²⁾기능을 제공한다[RUI97, ORTEGA97].

2.4.2.3 NETRA¹³⁾

NETRA는 미국의 산타 바바라에 있는 캘리포니아 대학에서 개발한 시스템으로 알렉산드리아 전자도서관 프로젝트에서 핵심적으로 사용된 기술이다 [MANJUNATH95, GOLDSTEIN99]. 이 시스템은 유사한 영역을 찾기 위해서 분할된 영역에서 색상과 질감, 모양, 그리고 공간적인 위치정보를 사용하여 검색한다. NETRA의 주요한 특징은 객체나 영역을 기반으로 한 검색에서 영상분할기법을 사용하고 있으며, 문자에 의한 검색이 가능하도록 주제어를 제공하고 있다. 즉, 『푸른 하늘』, 『꽃』, 『산』과 같은 주제어로 영상을 분류하고 이에 대한 검색기능을 함께 제공하고 있다[DENG99, MA97a, b, c,].

2.4.2.4 Photobook 및 FourEyes¹⁴⁾

MIT의 Media연구소에서 개발한 Photobook은 영상에 대한 주석을 붙이기 위한 보조도구인 FourEyes를 포함하고 있다[PENTLAND94]. 이 시스템은 컴퓨터가 사용자의 도움을 받아 영상의 주석을 붙일 수 있는가를 실험하기 위해 개발되었다. 즉, 영상의 영역 중에서 인간이 『빌딩』, 『거리』 등과 같이 이름을 붙이면 이 시스템은 자동적으로 영상의 나머지 부분을 추론하여 주석을 처리

11) Multimedia Analysis and Retrieval System

12) Relevance Feedback. 검색된 결과에서 질의를 계속하여 사용자가 원하는 결과에 좀더 접근하기 위한 방법이며, 요즈음 각 웹사이트에서 결과내에서 다시 검색하도록 하는 방법과 유사하다.

13) <http://maya.ece.ucsb.edu/Netra/>

14) <http://www-white.media.mit.edu/~tpminka/photobook/>

한다. 웹에서 운용하고 하고 있는 데모 사이트의 데이터베이스는 질감영상, 열
굴영상 등이 있다[PICARD95, MINKA96].

2.4.2.5 기타

앞서 언급한 영상검색시스템 외에도 버클리 대학에서 개발한 『BlobWorld』
및 시몬 프레이저 대학에서 개발된 『C-Bird』 등 다른 주제들을 가지고 다양
한 방법으로 학술적인 관점에서 많은 연구가 진행되고 있다. 현재 운용 중인
대표적인 영상검색시스템을 웹 주소를 포함하여 표 2.1에 표시하였다.

Table 2.1 Typical academic Image Retrieval System.

Name	Demo	Organization	Address
<i>C-BIRD</i>	○	Simon Fraser University (Canada)	http://jupiter.cs.sfu.ca/cbird/
<i>BlobWorld</i>	○	UC Berkeley university (USA)	http://elib.cs.berkeley.edu/photos/blobworld/
<i>El Nino</i>	×	California University (USA)	http://vision.ucsd.edu/~ssantini/papers.htm
<i>ImageRover</i>	○	Boston University (USA)	http://www.cs.bu.edu/groups/ivc/ImageRover/
<i>PICSearch</i>	○	Helsinki University (Finland)	http://www.cs.helsinki.fi/~klemstro/picsearch/
<i>PicToSeek</i>	○	Amsterdam University (Netherlands)	http://zomax.wins.uva.nl:5345/zomax/HTML/zomax.htm
<i>Surfimage</i>	○	INRIA (France)	http://www-rocq.inria.fr/imedia/
	○	Purdue University (USA)	http://min.ecn.purdue.edu/~jauyuen/
	○	RMIT University (Australia)	http://goanna.cs.rmit.edu.au/~nepal/publications.html
	○	Stanford University (USA)	http://bergman.stanford.edu/~zwang/imsearch/imsearchdemo.html
	○	Southern California University (USA)	http://viola.usc.edu/extranet/Projects/database-interactiveimage/default.htm

III. 영상의 특성표현

앞장에서는 내용기반 영상검색시스템의 개요 및 구성과 사용자가 시스템에 접근하는 방법인 질의·검색기법에 대해서 연구하였으며, 덧붙여 현재 개발 운용되고 있는 시스템 중 대표적인 사례를 연구함으로써 본 분야에 대한 세계적인 동향이 어떠한지를 살펴보았다. 여기에서 특이한 점은 상업적인 용도로 개발된 시스템보다 학술적인 목적으로 개발된 시스템이 매우 다양하며, 우수한 성능을 보이는 시스템도 있었다. 앞서 연구한 바와 같이 이러한 시스템에서 공통적으로 사용하고 있는 영상의 특징이나 속성을 표현하기 위한 방법은 주로 "색상과 질감 및 영상의 모양을 어떻게 다루어야 하는가?" 하는 문제였다. 그리하여 본 장에서는 이러한 영상의 특성을 표현하기 위한 방법에 관하여 색상, 질감 및 형태의 순서로 기술하였다.

3.1 색 상

색상은 영상의 특징을 시각적으로 가장 잘 표현하는 특성이며, 외견상 보이는 모든 물체는 색상과 밝기에 의해 구별되어 질 수 있다. 영상의 내용을 표현하기 위하여 전체적인 색상특성과, 색상과 물체의 모양을 조합하여 표현하는 국부적인 색상특성을 사용하여 표시할 수 있다. 특정한 영상을 색상으로 표현하기 위한 가장 일반적인 방법은 색상의 밝기를 히스토그램으로 변환하는 것이다. 예를 들면, 그레이-스케일 영상은 그림 3.1에서 보는 바와 같이 특정한 그레이-레벨을 갖는 화소들의 수를 누적한 히스토그램으로 변환할 수 있다. 즉, 원래의 색상영상을 R, G, B 각 채널의 그레이-레벨 영상으로 분리한 후, 채널별로 0에서 255까지의 밝기를 가진 화소들을 계수하여 히스토그램으로 표시한다.

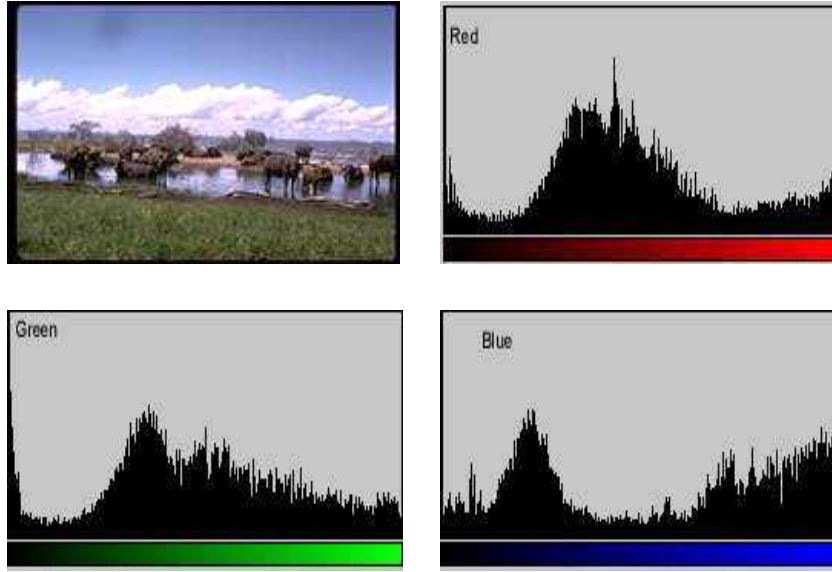


Fig. 3.1 Grey level histogram for RGB channel.

a	b	(a) Original image	(b) R channel
c	d	(c) G channel	(d) B channel

히스토그램을 영상검색기법으로 최초로 도입한 Swain에 의하면, "히스토그램은 영상의 이동과 회전에 불변하며, 영상의 크기, 카메라 각도의 변화, 물체의 흡수(occlusion) 등에 덜 민감하다. 그러므로 히스토그램은 영상검색을 위한 기법으로 영상간의 유사도를 표현하는데 사용하는 것이 가능하다."고 주장하며, opponent RGB공간에서 색상 히스토그램에 관한 다양한 실험을 하고 있다 [SWAIN91]. 여기에서 Swain이 영상검색에서 이용한 opponent RGB공간은 인간의 시각체계와 유사한 rg , by , wb 축을 사용한다. 이러한 공간을 식으로 표현하면,

$$rg = r - g \quad (3-1)$$

$$by = 2 \times b - r - g$$

$$wb = r + g + b$$

이 된다. 식에서 r , g , b 는 RGB공간¹⁵⁾의 각 채널을 표시한다. 여기에서 RGB

15) RGB공간과 반대개념을 가진 공간을 CMY(K)공간이라 한다. 이 공간은 감산혼색에 의해 생성된 색상으로 프린터와 같은 장치에 사용된다. RGB 공간의 단점은 색상이 시각적으로 일정한 분포를 보

색상공간이란 매우 간단한 구조인 정육면체를 이루고 있으며, 정육면체의 각 축은 색상의 밝기를 나타내는 공간으로서 그림 3.2와 같다.

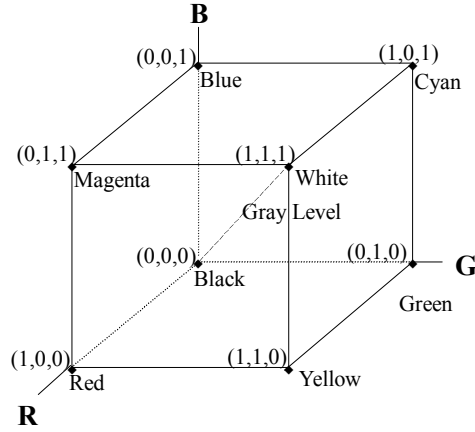


Fig. 3.2 The RGB cube for additive devices, such as CRT displays.

영상간의 유사도를 표현하기 위하여 히스토그램을 사용할 경우 조명조건
의 변화나, 해상도가 달라지면 히스토그램에 영향을 미치게 되므로, 이러한 영
향을 줄이고자 두 히스토그램간의 비교는 식(3-2)과 같은 히스토그램 인터섹션
을 사용한다. 즉,

$$H(I, M) = \sum \min(I_j, M_j) \quad (3-2)$$

와 같다. 여기에서 I , M 은 각각 질의영상과 목표영상, j 는 히스토그램 빈이다.
위의 식에서 알 수 있는 바와 같이 히스토그램 인터섹션의 결과는 두 영상간
의 색상 빈에서 같은 색상의 화소수를 나타낸다. 영상의 크기에 관계없이 결
과가 0과 1 사이의 값을 가지도록 정규화하기 위한 식을 나타내면,

이고 있지 않을 뿐 아니라, 직관적이지 못한 점이다.

$$H(I, M) = \frac{\sum \min(I_j, M_j)}{\sum_{j=1}^n M_j} \quad (3-3)$$

이 된다. 히스토그램 인터섹션방법은 질의영상의 색상 히스토그램을 데이터베이스내의 모든 영상의 히스토그램과 비교해 보는 검색방법으로 계산이 간단하다는 장점을 지니고 있다. 색상 히스토그램에 대한 연구로는 Stricker에 의한 히스토그램의 용량에 관한 연구가 있으며, 이 논문에서 누적 히스토그램을 제안하고 있다[STRICKER94]. 이러한 히스토그램은 색상의 밝기가 히스토그램으로 변환될 때 영상자체에서 포함하고 있는 공간정보는 상실하게 되어 히스토그램 비교방법을 영상검색에 이용할 경우 공간정보의 부재로 검색결과가 좋지 않게 된다.

공간정보 부재문제를 해결하는 방법으로 영상의 공간적인 밀집과 분산관계를 고려하여 화소들을 분리하는 CCV방법을 Pass 등은 제안하였다. 그들에 의하면, "유사하지 않는 두 개의 영상일 지라도 히스토그램 상으로 비슷한 결과가 나타날 수 있다"고 주장하며, CCV기법을 활용하여 이를 보완하고자 하였다[PASS96b]. CCV에서는 양자화된 색상별로 영상내의 색상이 응집되어 있는지 아니면 분산되어 있는지를 구분하여 색상의 응집·분산상태를 구분한다. 예를 들면, 그림 3.3에서 보는 바와 같이 임의의 화소값 10과 19사이를 1, 20에서 29까지를 2 등과 같이 진행하면 그림 3.3(b)가 되며, (b)에서 임의의 화소의 상하좌우에 있는 화소들 간의 연결여부를 (c)와 같이 문자로 표시하면, 화소들의 응집여부를 그림 3.4과 같이 작성할 수 있다.

22 10 21 22 15 16	2 1 2 2 1 1	B C B B A A
24 21 13 20 14 17	2 2 1 2 1 1	B B C B A A
23 17 38 23 17 16	2 1 3 2 1 1	B C D B A A
25 25 22 14 15 21	2 2 2 1 1 2	B B B A A E
27 22 12 11 21 20	2 2 1 1 2 2	B B A A E E
24 21 10 12 22 23	2 2 1 1 2 2	B B A A E E
(a)	(b)	(c)

Fig. 3.3 The computation of the connected components for CCV method.

그림 3.4에서 색상 (1, 2, 3)에 대한 화소들의 분포는 [(12, 3),(20, 0),(0, 1)]과 같이 되며, 이를 해당되는 영상의 특성으로 이용할 수 있다. 여기서 집합요소인 숫자조합 중 앞의 수는 응집, 뒤의 수는 분산을 나타내며, 응집여부를 판단하기 위한 값은 임의의 값인 4로 한 경우를 예로 하였다. 즉, 1로 구분된 자료 중 A 항목에 속하는 자료의 수는 4보다 크므로 응집으로 분류하고, C 항목에 분류된 자료는 4보다 적은 3의 값을 가지므로 분산으로 분류하였다.

Label	A	B	C	D	E		Color	1	2	3
Color	1	2	1	3	2	⇒	Coherence	12	20	0
Size	12	15	3	1	5		Incoherence	3	0	1

Fig. 3.4 The connected components and CCV table.

공간정보를 고려하기 위한 다른 방법으로 Huang은 Color Correlogram을 제안하고 있다. 이는 임의의 한 화소 i 가 k 거리만큼 떨어진 화소 j 를 찾을 확률을 상관관계로 작성한 테이블을 특성벡터로 이용한다. 그들에 의하면, Color Correlogram은

- i) 색상의 공간 상관관계를 포함하며,
 - ii) 계산하기 쉬우며,
 - iii) 특성벡터의 크기가 매우 작은 장점이 있음을 주장하고 있다[HUANG97, 98].
- 또한 Stricker는 영상을 다섯 개의 영역¹⁶⁾으로 분할하여 각 영역들이 최소한의 공간정보를 포함할 수 있도록 하고 있다. 이러한 영역들로부터 색상의 모멘트인 평균, 분산, 왜곡을 계산하여 특성벡터로 사용함으로써 영상의 이동과 회전에 강한 알고리즘을 제안하고 있다[STRICKER96]. 이들이 모멘트를 계산하는 식은 다음과 같다. 즉,

16) 중앙의 타원형(fuzzy regions)을 중심으로 영상을 4등분하여 분할하였음.

$$E_{l,i} = \frac{1}{N R_I^{tot}} \sum_p R_I(P) C_i(P) \quad (3-3)$$

$$\sigma_{l,i} = \left(\frac{1}{N R_I^{tot}} \sum_p R_I(P) (C_i(P) - E_{l,i})^2 \right)^{\frac{1}{2}} \quad (3-4)$$

$$s_{l,i} = \left(\frac{1}{N R_I^{tot}} \sum_p R_I(P) (C_i(P) - E_{l,i})^3 \right)^{\frac{1}{3}} \quad (3-5)$$

이 된다. 여기에서 N 은 전체영상의 화소수이며, R_I 은 fuzzy 영역의 멤버십 함수, $C_i(P)$ 는 임의의 화소 P 에서 i 번째 채널, $R_I^{tot} = \sum_{j=1}^N R_I(j)$ 이며, $E_{l,i}$, $\sigma_{l,i}$, $s_{l,i}$ 는 각 영역별 평균색상, 분산 및 왜곡을 나타내고 있다. 영역의 경계에 가까운 화소들은 중앙에 있는 화소들 보다 가중치를 더 적게 주어 멤버십 함수를 조절하여 영상의 회전이나 이동에 각 모멘트가 완만하게 변하도록 하였다.

Gong은 RGB색상공간이 가지는 좌표계가 인간에게 직관적인 의미정보를 전달하여 주지 못하므로 인간의 시각인식체계와 비슷한 Munsell색상을 사용할 것을 제안하며, 11개의 색상으로 특성벡터를 구성한 히스토그램에서 조명의 변화나 잡음의 영향을 최소화하는데 노력하고 있었다[GONG96]. 비슷한 방법으로 Cox는 HSV로 변환한 영상에서 12개의 색상으로 특성벡터를 구성하여 영상검색에 적용하고 있었다.[COX96a, 96b]. 이러한 Munsell¹⁷⁾ 및 HS계열의 공간은 인간의 시각인식체계와 유사하므로 영상처리분야에서 많이 사용되고 있다. Munsell의 색상표현방법은 색도(hue), 명도(value), 채도(chroma)의 세 가지 색상속성을

17) Munsell공간은 1905년경 미국의 화가이자 색채 연구가에 의해 개발된 색상공간[MUNSELL69]을 1940년 미국의 광학회에서 수정·개량한 것이다. 수정색상은 처음에 생각한 기초 구조 개념과 기호를 표시하는 방법은 같으나, 시각적인 색채의 단계수정과 CIE표준색상으로 변환하는 것 등의 개선이 있었다. 우리나라에서도 한국공업규격(KS A00620)으로 채택하고 있다. 색상을 표시하는 방법은 HV/C의 순서로 기록하며, 5Y8/10 인 경우에는 색도는 5Y, 명도 8, 그리고 채도가 10인 색상을 표현하고 있다. 이러한 표현은 시각인식체계의 크기와 동등하지 않아서 다른 채도레벨의 같은 색상을 갖는 것은 다른 색상을 갖는 것으로 인식되고 있으며, 또한 채도단위와 명도단위도 시각적인 크기와 일치하지 않는 단점을 지니고 있다.

그림 3.5와 그림 3.6와 같이 삼차원의 좌표계에 표시한다.

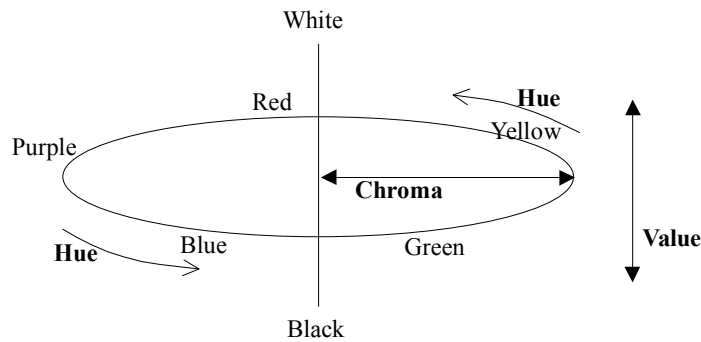


Fig. 3.5 Illustration of three scales of Munsell system.

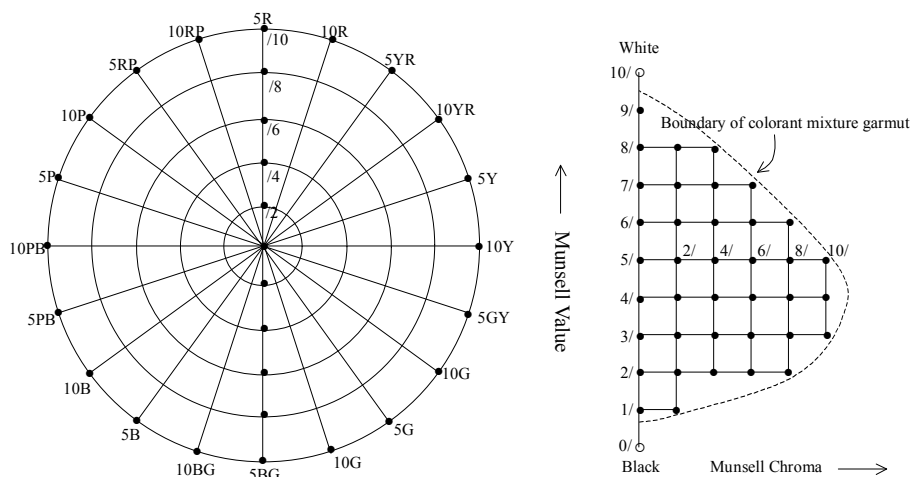


Fig. 3.6 Munsell color system, (a) organization the colors of the Munsell Value and (b) organization the colors of the Munsell Hue.

그리고 HS계열¹⁸⁾의 색상공간을 그림 3.7에 나타내었다. 그림에서 보는 바와 같이 중앙의 선은 밝기의 정도를 나타내고, 채도(saturation)는 중앙 축과의

18) HSV, HLS 및 HSI공간은 약간씩 차이는 있지만, 동일 계열로써 RGB공간을 변형하여 개발하였다. HS계열의 색상공간간 차이는 채도와 밝기의 강도를 표현하는 방법에 따라 약간씩 다르게 표현한다. 즉, HSV는 밝기가 "0" 일 때 채도는 "0", "1"일 때 채도가 최대가 되므로 직관적이지 못한 색상공간을 제공한다. HLS는 이 보다는 좀 더 직관적으로 밝기에서 0.5되는 위치에 채도가 최대가 되도록 하였다. HSI와 HLS 공간은 이름만 다를 뿐 실제로 거의 동등하다. 이 중에서 가장 일반적으로 사용되고 있는 것은 HSV와 HLS를 변형, 조합한 단순한 원통형인 그림 3.7 (c)이다. 이는 흰색과 검정색을 포함한 모든 색에 대한 채도가 일정한 분포를 갖도록 한다.

거리, 그리고 색도는 중앙을 둘러싸고 있는 원 또는 육각형으로 표시한다.

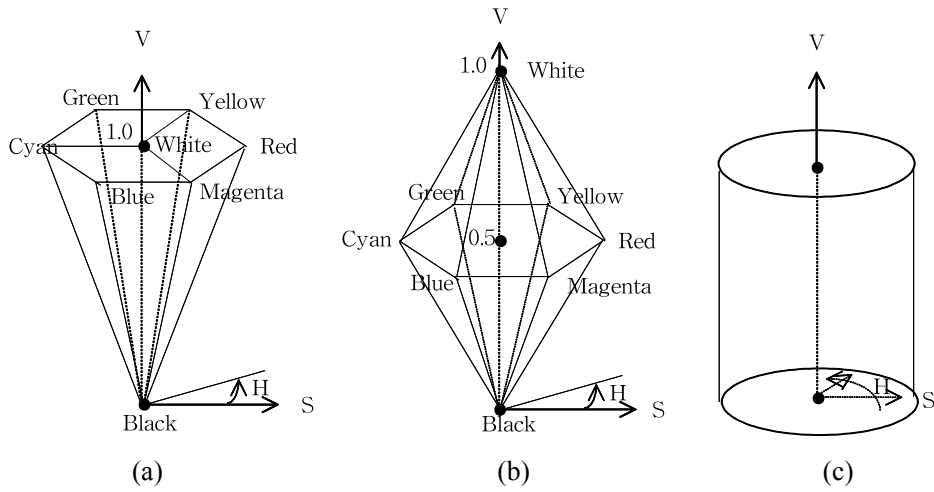


Fig. 3.7 The RGB derived color space (a) HSV, (b) HLS and (c) the most commonly used HSV space.

HS계열의 색상들은 RGB공간을 변형하여 사용상의 간편성을 도모함으로써 직관적인 면은 개선되었으나, 인식상 균등한 색상분포를 제공하지 않는 단점(비직선성)이 있다. 다시 말하면, RGB가 R, G, B의 비율에 의해 색상을 결정하는 수학적 모델이라면, HS 모델은 직관적인 모델이다.¹⁹⁾

앞에서 살펴 본 바와 같이 색상은 영상에서 보이고 있는 가장 대표적인 속성으로서 영상검색에서 특징벡터로 가장 많이 사용하고 있으며, 현재까지 이용되고 있는 영상검색시스템에서도 다른 특징요소들 보다 우선적으로 사용되고 있다. 본 논문에서도 이러한 점을 감안하여 RGB공간의 색상을 이용하여 영상내의 특징을 추출하는 색상-공간을 기반으로 한 영상검색에 대하여 연구를 진행하였다.

19) 직관적인 모델과 수학적 모델의 차이는, 색상을 정의할 때 색상 이름을 알고 하느냐, 또는 모르고 하느냐하는 것이다. 직관적인 모델에서는 색상의 이름을 알고 색상을 지정하며, 수학적 모델은 색의 구성 요소의 비율을 정하지만 그것이 어떤 색상인지는 쉽게 알 수 없음을 의미한다.

3.2 질 감

질감에 대한 연구는 상당히 많은 성과와 발전을 거듭해온 분야로서, 인간의 질감인식을 기반으로 한 일반적인 수학적 모델을 가지면서 매우 복잡한 양상을 지니고 있다. 그러나 대부분 통계적기법과 주파수에 기초를 둔 방법으로 질감을 분할하는 기법이 무수하게 연구되어 오고 있지만 아직도 이에 대한 완전한 일반적인 모델을 가지고 있지 못하다.

물체는 질감특성을 가질 때 거칠음과 부드러움, 방향성과 임의성, 규칙과 불규칙 등 양면성을 가지는 경우가 허다하다. 예를 들면, 영상에서 『풀』이나, 『구름』, 『호랑이』의 무늬 등은 이러한 좋은 예가 된다. 영상검색에서 질감에 관한 관심이 높아지고 있긴 하지만 이용하고 있는 질감특성의 대부분은 기본적인 질감모델을 사용하고 있는 게 현실이다. 영상검색에서 질감특성을 사용함에 있어 문제점으로 지적할 수 있는 것은 영상의 해상도에 따라 질감의 특성도 변화한다는 것이다. 다시 말하면, 영상의 크기가 커짐에 따라 질감특성의 주파수 성분도 높은 주파수 쪽으로 변하게 된다. 『호랑이』 영상의 예에서 영상의 크기가 작으면 전체적인 무늬는 질감특성을 유지하지만 머리 부분에는 질감특성이 사라지게 된다[SEABORN98].

질감에 대한 용어는 다양한 방법으로 정의되고 있다. Parker는 "질감의 주요한 특성은 영상의 영역에서 나타나고 있는 패턴 또는 반복하는 형태의 패턴으로 정의"하고 있다[PARKER97]. 비록 질감의 패턴은 색상에 의해 나타나고 있지만, 일반적으로 질감을 분석하는 데는 주로 그레이-레벨을 사용하고 있다[PENTLAND94]. 질감을 분석하기 위한 방법들은 영상처리의 문헌에서 광범위하게 찾아 볼 수 있으며, 가장 일반적으로 사용되고 있는 방법은 그레이-레벨에 의한 통계적인 기법이 주류를 이루고 있다[JAIN89, CASTLEMAN96, PARKER97]. 질감인식과 통계적인 접근방법을 이론화한 Julesz 등은 텍스톤(textons)이라는 용어를 사용하여 질감특성을 표현하기 위한 개선된 방안을 제시하고 있으며[JULESZ73], Tamura 등은 질감의 특성을 조악함, 대비성, 방향성, 직선성, 균일성 및 거칠음의 6가지²⁰⁾로 구분하고, 이들과 인간의 시각적인 인식과의 상관관계를 16개의

20) coarseness, contrast, directionality, line-likeness, regularity, roughness

패턴을 설정하여 연구·실험하였다[TAMURA78].

Freeman 등은 조절 가능한(steerable) 필터를 설계하고, 위상분석, 윤곽검출, 모양형성 등 여러 분야에 이러한 필터를 적용하였으며[FREEMAN91], Picard 등은 자연 상태의 천연적인 질감에서 대표적인 방향성분을 추출하기 위하여 앞서와 같은 조절 가능한 필터를 사용하였다[PICARD94]. Liu 등은 주기성, 방향성, 임의성을 고려한 2-D Wold이론을 기반으로 계산한 질감특성을 이용하여 영상모형을 표현하였다[LIU96, 99]. 또한 Ma 등은 항공사진을 질감특성을 이용한 영상검색시스템으로 구현하였다. 이들은 항공사진을 동일한 성분을 가진 영역으로 분할하고, Gobor필터에 의해 구해진 질감특성을 가진 영상을 특성벡터로 사용하며, 혼합신경망 알고리즘을 사용하여 질감의 유사도를 근거로 계층적인 구조에 의한 질감분류를 시도하고 있다[MA97b].

질감특성을 이용하여 검색시스템에 적용하기 위한 최근의 연구로는 Smith 등에 의한 영상의 공간-주파수 분석으로 추출된 이진 질감특성이 있으며[SMITH96a], Belongie 등은 EM기법을 사용하여 영상을 분할한 후, 색상과 질감특성을 조합하여 영상검색에 이용하고 있다[BELONGIE98].

3.3 형 태

영상에서 형태특성은 색상이나 질감과 비교하여 상대적으로 특성추출이 어렵다. 영상내의 객체의 형태를 정의하기 위해 윤곽정보를 많이 사용하고 있으나, 이러한 윤곽은 상대적으로 밝기의 차이가 뚜렷한 경우에는 쉽게 처리할 수 있지만 실제영상에는 이러한 경우는 많지 않다. 인간이 영상을 인식하는 것은 잡음이나 관련성이 없는 형태를 무시하고 중요한 부분만 인식하는 것과 같이 객체의 형태특성을 논리적으로 구성하기는 매우 어려운 문제이다. 형태 특성추출에 관한 연구가 오랫동안 지속되어 왔음에도 불구하고 아직도 많은 연구가 필요한 것은 이러한 점을 잘 반영하고 있다. 그리하여 대부분의 영상 검색에서 채택하고 있는 형태특성은 적당한 범위에서 타협적으로 이용하고 있으며, 이와는 달리 형태에 대한 처리를 사용자에게 전적으로 의존하기도 한다.

형태 또는 객체를 분석하기 위한 기술은 일반적으로 경계선과 영역기반 방식으로 접근하는 것이 대표적이다[JAIN89, WATT98a, CASTLEMAN96]. 이러한 방법들은 영상의 내용을 표현하기 위해 영상코딩기법을 사용하는데, 이를 표현하면 그림 3.8 및 그림 3.9와 같이 나타낼 수 있다[LA100]. 그림에서 보는 바와 같이, 경계선 접근방법은 영상 내에 존재하는 윤곽 또는 객체의 경계선을 추출한 후, 연결된 윤곽의 방향을 고려하여 객체의 모양대로 순서에 의해 번호를 부여하는 방법으로 코딩하며, 경계선 접근방법은 동일한 특성을 가진 화소들로 그룹을 형성한 영역을 임의크기의 사각형으로 표현한 후, 트리구조로 계층적인 형태에 의해 영상을 표현한다.

사용자가 사용자 인터페이스를 통해 개략적인 스케치에 의해 입력한 형태를 바탕으로 검색하는 시스템을 Bimbo 등은 제안하고 있으며[BIMBO96], Mehrotra 등은 경계선을 기반으로, 정규화된 점들로 구성된 특성벡터를 색인한 FIBSSR(Feature Index-Based Similar-Shape Retrieval)시스템과, 2-D 및 3-D 형상의 검색기능을 가지고 있는 MUSEUM(Multi-SEmantic Unstructured data Model)을 구현하였다[MEHROTRA95].

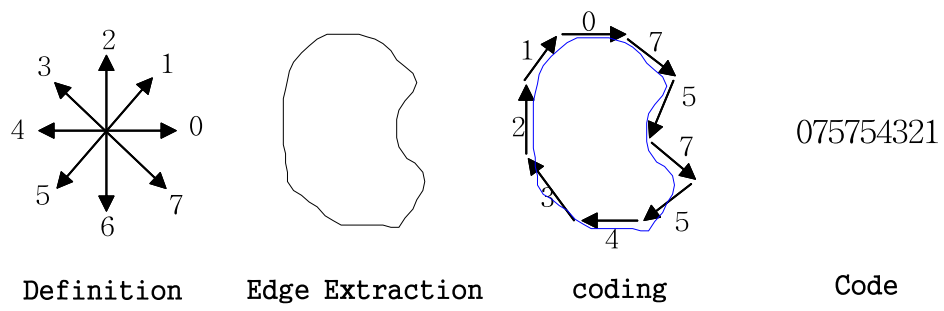


Fig. 3.8 Illustration of a boundary-based representation.

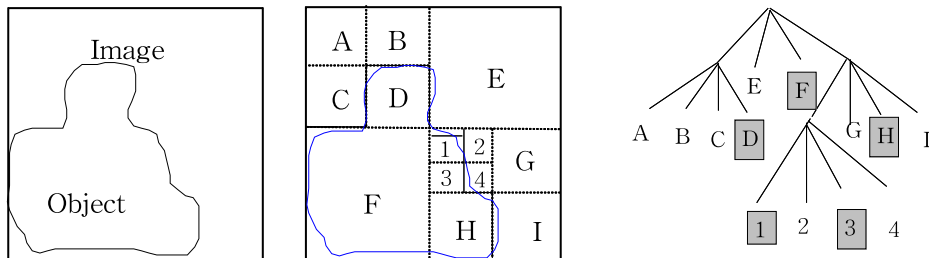


Fig. 3.9 Illustration of a region-based representation.

IV. 색상-공간정보를 고려한 영상검색

내용기반 영상검색에 대한 연구는 영상정보의 중요성을 인식하고 있는 많은 사람들에 의해 활발히 진행되어 왔으며, 이러한 연구 중에서 가장 큰 관심은 색상이었다. 이러한 색상을 영상검색에 고려한 가장 대표적인 방법은 Swain에 의한 히스토그램 인터섹션 방법이다. 이 방법은 같은 색상을 가진 화소들의 개수를 색상 빈의 빈도수에 따라 계수하여 색상 히스토그램으로 표현한다. 또한, Hafner 등은 영상의 색상분포 뿐 아니라 히스토그램의 각 구간사이의 거리와 최대거리에 대한 비율을 유사행렬로 이용하여, 히스토그램 사이의 유사성을 구분하였다[HAFNER95].

이러한 방법은 영상이 지니고 있는 다양한 특징 중에서 영상의 전체적인 색상 히스토그램과 모멘트들을 이용하여 비교하므로 다른 정보와 비교하여 상대적으로 쉽게 추출할 수 있으며, 계산이 복잡하지 않고, 영상의 크기나 방향의 변화에 강한 성질을 지니고 있어, 영상검색에 매우 유용한 정보를 제공한다. 그러나 포함된 물체의 위치나 공간적인 관계를 나타내지 못하고 색상분포만을 고려함으로써 유사한 히스토그램을 가지는 서로 다른 영상을 구별하지 못하고 있다. 그리하여 영상내의 공간적인 정보를 고려하기 위한 방법으로 영상의 엔트로피를 계산하고, 추출된 영역을 사각형 꼴로 표현하였으며, 또한 영상을 점들로 표현하여 각 영역에서 중앙과 주변에서 나타나는 대표색상을 특성벡터로 이용하기도 하였다.

영상의 영역이나 객체들을 기반으로 하는 색상정보가 공간적인 분포를 고려하기 위한 방법으로 매우 유용하지만, 영역을 분할하는데 어려움이 있으므로 간단한 고정크기의 분할방법이 다양하게 이용되고 있다. 그리하여 Stricker 등은 영상을 중앙의 타원형과 겹치는 영역을 포함하여 5영역으로 구분하였으나, 영상에 나타나는 색상정보를 상세하게 반영하지 못하였다[STRICKER96]. 이와는 달리 영상을 고정 또는 가변블록으로 분할하지 않고 공간적인 관계를 표현하기 위하여 Pass 등은 영상 내에 포함하고 있는 같은 색상들이 응집되어 있는지를 파악하여 색상별로 응집과 분산표를 작성한 CCV를 제안하여 기존의 히스토그램 방법을 개선하였다[PASS96b]. 공간정보를 고려하기 위한 다른 방법으

로는 주어진 색상의 화소가 임의의 거리에 존재할 확률을 계산한 테이블을 작성하여 영상의 특성벡터로 사용하는 Color Correlogram이라는 방법을 제안하였다[HUANG97]. 그러나 이러한 방법들은 색상 히스토그램의 단점을 개선은 하고 있으나, 여전히 공간정보의 부재문제를 충분히 해결하지 못하고 있었다.

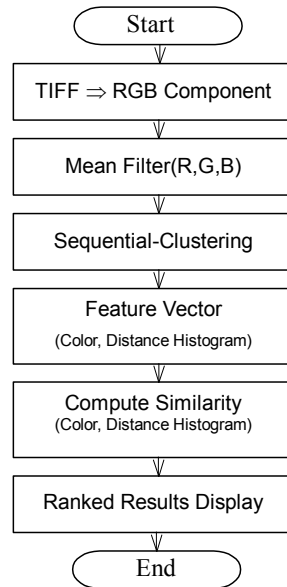


Fig. 4.1 Schematic of Proposed System.

그리하여 본 논문에서는 색상-공간정보를 고려하기 위한 방법으로 영상을 분할하는데 있어 고정분할방법을 사용하지 않고 영상의 색상이 동일한 영역을 적응적으로 처리하여 영역을 분할하는 기법을 사용한다. 또한 분할된 영역에서 포함하고 있는 평균색상과 해당영역의 화소들이 중심점으로부터 떨어진 거리를 3가지로 구분하여 계수한 거리 히스토그램으로 영상을 표현함으로써 색상과 공간을 함께 고려할 수 있는 알고리즘을 그림 4.1과 같이 제안한다.

그림 4.1에서와 같이 본 논문에서 제안하고 있는 방법은 RGB공간에서 영상을 처리한다. RGB공간에서 영상을 처리하기 위해 먼저 TIFF형식의 칼라 영상을 RGB요소로 나누게 된다. 원 영상에 포함된 잡음의 영향을 억제할 뿐 아니라 다음 처리과정의 순차영역분할단계에서 분할된 영상이 군집되는 효과가 커지도록 할 뿐 아니라, 작은 영역으로 분할되는 것을 최소화하기 위해 평균 필터를 사용한다.

이러한 과정을 거쳐 처리된 영상을 기초로 하여 본 논문에서 제안한 순차

영역분할 방법을 도입하여 영상을 3영역으로 분할하며, 분할된 영역에서 거리 히스토그램 및 유사도를 계산하고, 계산된 유사도에 의해 순위별로 영상을 출력하는 시스템을 구성한다.

4.1 순차영역분할

영상 내에서 특성이 동일한 영역으로 영상을 분할하기 위해 패턴인식분야에서 사용되고 있는 순차영역분할기법[THEODORIDOS99]을 영상검색에 이용할 수 있도록 제안한다[장정동2001]. 이 방법은 연속되는 일련의 자료를 군집화하는 기법으로 처리속도가 매우 빠른 장점을 지니고 있다. 이의 군집과정을 살펴보면 먼저 입력되는 최초의 자료를 첫번째 군집영역의 대표값으로 초기 설정한다. 그리고 이 대표값과 뒤이은 자료와의 거리를 계산하여 임의의 설정치와 비교하게 된다. 이때 계산한 거리가 설정한 값에 의해 정해진 범위를 벗어나면 두번째 생성되는 새로운 군집영역을 형성하게 되고, 정해진 범위 안에 있으면 해당군집영역에 할당하게 함으로서 일련의 연속되는 자료를 군집화한다. 여기에서 군집영역을 생성할 때 해당영역의 대표값은 다음과 같은 방법으로 계산한다. 즉, 입력되는 영상의 화소값인 f 와 군집영역 C 간의 거리를 $d(f, C)$ 와 같이 나타내면,

$$d(f, C) = d(f, \overline{C}) \quad (4-1)$$

가된다. 이때 \overline{C} 는 군집영역 C 의 대표값(평균값)을 나타내며, 생성된 군집영역 C 에 입력영상의 화소의 값 f 가 할당될 때마다 해당군집영역의 대표값을 식 (4-2)과 같이 적응적으로 변화시키게 된다. 즉,

$$\overline{C}_k^{new} = \frac{(nC_k^{new} - 1) \overline{C}_k^{old} + f}{nC_k^{new}} \quad (4-2)$$

이며, 여기서 nC_k^{new} 는 입력화소 f 가 k 번째 군집영역에 할당된 상태에서 계산된 화소의 수, \overline{C}_k^{new} 및 \overline{C}_k^{old} 는 각각 입력화소 f 가 군집영역 C_k 에 할당되기 전과 후의 대표값을 나타낸다.

본 논문에서는 순차영역분할 알고리즘을 RGB색상공간에서 적용한다. 군집 영역의 구분을 더욱 정확히 하려면 군집영역결정루틴과 영역분할루틴으로 나누어 최초루틴에서는 각 영역의 대표값 즉, 색상평균값을 결정하고, 다음 루틴에서 영역을 분할하는 기법을 사용한다. 생성되는 영역의 대표값인 색상평균값은 다음과 같이 계산한다. 즉,

$$\overline{R}_k = \frac{\sum r_k}{n_k}, \quad \overline{G}_k = \frac{\sum g_k}{n_k}, \quad \overline{B}_k = \frac{\sum b_k}{n_k} \quad (4-3)$$

$$R_k = |r - \overline{R}_k|, \quad G_k = |g - \overline{G}_k|, \quad B_k = |b - \overline{B}_k| \quad (4-4)$$

$$D = \sqrt{R_k^2 + G_k^2 + B_k^2} \quad (4-5)$$

$$Min(TH, D) = \begin{cases} C_k, & TH \leq D \\ C_{k+1}, & TH > D \end{cases} \quad (4-6)$$

과 같다. 여기에서 k 는 제안한 알고리즘에 의해 생성된 군집영역의 수, n_k 는 k 번째 영역의 화소수를 나타낸다. 또한 \overline{R}_k , \overline{G}_k , \overline{B}_k 및 r_k , g_k , b_k 는 k 번째 영역의 RGB 3채널의 평균치 및 화소의 값이며, r , g , b 는 처리되기 위해 입력

되는 RGB 각 채널의 화소값이다. R_k, G_k, B_k 는 k 번째 영역에서 입력되는 화소값과 해당군집영역의 평균값과의 거리이며, D 는 식(4-5)에서 계산된 거리를 RGB 3채널에 대하여 고려한 유클리디안 거리이다. 그리고 TH 는 분할되는 영역간의 구분을 짓는 임의의 설정치를 표시한다.

앞서 기술한 바와 같이 입력되는 영상의 화소값인 f 와 군집영역 C 간의 거리를 $d(f, C)$ 로 나타내고, 분할영역의 개수를 q , 같은 영역으로 구분하기 위한 임의의 설정치를 TH , 그리고 m 을 생성된 군집영역의 수라고 하면, 본 논문에서 사용한 순차영역분할 알고리즘을 그림 4.2와 나타내었다. 그림에서 보는 바와 같이 $Min(TH, D)$ 을 판단하여 입력되는 화소들을 해당되는 군집영역으로 포함시키고 있다. 그림 4.2에서 표시된 방법을 구현하기 위한 알고리즘을 표 4.1에 나타내었다. 앞서 언급한 내용과 같이 군집영역결정루틴과 영역분할루틴에서 각각의 처리가 간단히 이루어지고 있다.

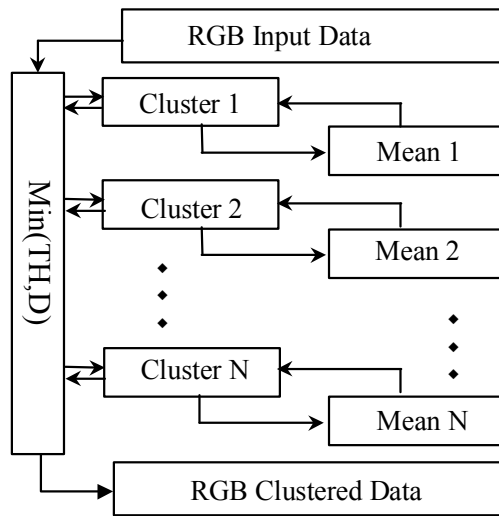


Fig. 4.2 Schematic of Sequential-Clustering.

Table 4.1 Sequential-Clustering algorithm.

```
Cluster Determination
•  $m = 1$ 
•  $C_m = \{f\}$ 
• For  $i = 2$  to  $N$ 
    Find  $C_k: d(f_i, C_k) = \min_{1 \leq j \leq m} d(f_i, C_j)$ 
    If  $(d(f_i, C_k) > TH)$  AND  $(m < q)$  then
         $m = m + 1$ 
         $C_m = \{f\}$ 
    End {If}
• End {For}
Pattern Classification
• For  $i = 1$  to  $N$ 
    If  $f_i$  has not been assigned to a cluster, then
        Find  $C_k: d(f_i, C_k) = \min_{1 \leq j \leq m} d(f_i, C_j)$ 
         $C_k = C_k \cup \{f_i\}$ 
        where necessary, update representatives.
    End {If}
• End {For}
```

제안한 알고리즘에서 군집영역의 수를 3개, TH 를 55로 설정하여 그림 4.3의 "꽃" 그림을 분할하면, 첫 번째 분할되는 영역은 그림 4.4와 같다.



Fig 4.3 Sample flower image.

그림에서 보는 바와 같이 노란 꽃을 제외한 하단의 잎들이 개체로서 분할되고 있음을 알 수 있다. 그림의 좌측 잎의 왼쪽에 있는 직선은 원 영상에서 나타나고 있는 영상의 내용과는 무관한 검은색 테두리이며, 이는 색인벡터 계산시 무시되도록 처리한다.



Fig. 4.4 1st clustered region using Sequential-Clustering method.



Fig. 4.5 2nd clustered region using Sequential-Clustering method.

그림 4.5는 두 번째 분할된 영역이며, 배경부분이 부분이 분할되고 있다. 그림 4.4 및 4.5에서 작은 점들이 분산되어 나타나는 것은 그 부분의 색상이 해당영역의 색상과 비슷하므로 분할된 영역에 남게 된다. 이러한 작은 점들에 의한 영향을 줄이기 위하여 영역을 분할하기 전에 평균필터를 사용하여 작은 영역으로 군집하는 것을 억제한다.



Fig. 4.6 3rd clustered region using Sequential-Clustering method.

그림 4.6은 세 번째 분할된 영역으로서 노란색상을 가진 꽃의 영역을 분할하였다. 그리고 그림 4.7은 분할된 영역을 한 눈에 알아 볼 수 있도록 각 영역을 합성하였다. 합성영상에서 보는 바와 같이 제안하는 알고리즘이 개체의 구분을 매우 우수하게 처리하고 있음을 시각적으로 확인할 수 있다.



Fig. 4.7 Composite clusters of each clustered regions.

4.2 거리 히스토그램과 특성벡터

공간정보를 고려하기 위해 분할된 영역에서 영상검색에 필요한 특성을 추출하기 위해서는 해당영역의 분산값과 영역의 크기, 각 영역의 중심점과의 각도 등, 여러 가지 특성요소를 계산하여 비교할 수 있으나, 본 논문에서는 영상의 공간정보를 충분히 고려할 수 있을 뿐 아니라, 계산이 매우 간단한 거리 히스토그램을 제안한다[장정동2001a].

거리 히스토그램은 색상 히스토그램에서 색상을 기준으로 빈을 구성하는 것처럼, 특정한 위치로부터 거리를 계산하여 히스토그램을 구성하게 한다. 즉, 제안한 순차영역분할 알고리즘으로부터 분할된 영역에 속하는 화소들의 위치를 해당영역의 중심점으로부터 계산된 거리를 3등분으로 구분한다. 이와 같이 계산된 거리의 각 부분에 속하는 화소의 개수를 계수하여 히스토그램으로 표현한다. 다시 말하면, 거리 히스토그램의 빈을 각 영역의 중심좌표로부터 가까운 위치, 중간 위치 및 먼 위치로 히스토그램화하여 영상의 특성벡터로 이용한다. 이러한 거리 히스토그램을 구분하는 방법을 그림으로 표현하면 그림 4.8과 같고, 이를 계산하기 위한 알고리즘은 표 4.2에 나타내었다.

이와 같이 본 논문에서 제안하는 거리 히스토그램은 하나의 분할된 영역에서 단지 3개의 정보만으로 위치정보를 고려할 수 있는 매우 단순한 방법으로 색상 히스토그램의 단점인 공간정보 부재문제를 해결할 수 있다.

본 논문에서는 색상-공간정보를 함께 고려하기 위한 방법을 제안하고 있으므로, 거리 히스토그램과 더불어 분할된 영역으로부터 영상이 지니고 있는 색상정보를 계산하여야 한다. 색상정보를 계산하기 위해 각각의 영역에서 RGB 3채널의 색상평균값을 특성벡터로 사용한다. 이러한 색상평균값은 식(4-7)에 의해 계산할 수 있으며, 거리 히스토그램은 식(4-8)과 같다. 즉,

$$Mean_{(r,g,b)}^k = \frac{\sum f_{(r,g,b)}}{255 \cdot N_k} \quad (k=1,2,3) \quad (4-7)$$

$$H(i) = n_i / N \quad (i=1,2 \cdots ,9) \quad (4-8)$$

이 된다. 여기에서 k 는 분할영역의 개수, N_k 는 해당영역의 화소수, 255는 그레이 레벨, 그리고 n_j 는 해당거리에 포함되는 화소수, N 은 전체 화소수를 나타낸다. 그리하여 $Mean_{(r,g,b)}^k$ 는 r, g, b 채널별로 분할된 영역의 색상평균값을 표시하게 되며, 거리 히스토그램 H 는 9개의 빈을 가지게 된다. 이와 같이 계산된 색상평균값 9개와 9개의 빈을 가진 거리 히스토그램을 합하여 총 18개의 특성값을 특성벡터로 하여 본 논문에서 제안하고 있는 색상-공간정보를 고려하는 색인정보로 저장한다.

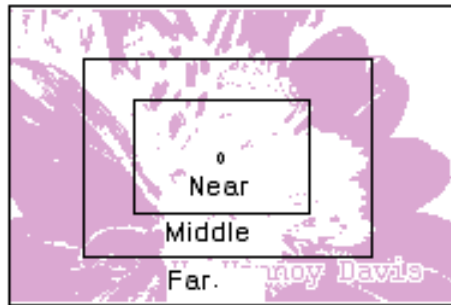


Fig. 4.8 Distance Histogram of a clustered region.

Table 4.2 Algorithm to compute a Distance Histogram.

<ul style="list-style-type: none"> • Compute the center coordinate of clustered region, C_x, C_y. • For Distance Histogram, compute <i>Near, Middle, Far</i> according to the following rules:
<p>If $C_x + spx1 > \hat{f}_x > C_x - spx1$ and $C_y + spy1 > \hat{f}_y > C_y - spy1$,</p> <p style="padding-left: 40px;"><i>Near++</i></p> <p>else If $(C_x + spx2 > \hat{f}_x > C_x + spx1$ or $C_x - spx1 > \hat{f}_x > C_x - spx2)$ and $(C_y + spy2 > \hat{f}_y > C_y + spy1$ or $C_y - spy1 > \hat{f}_y > C_y - spy2)$,</p> <p style="padding-left: 40px;"><i>Middle++</i></p> <p>else</p> <p style="padding-left: 40px;"><i>Far++</i></p> <p style="padding-left: 80px;">End {If}</p> <ul style="list-style-type: none"> • spx, y: Distance interval of the coordinate x, y, respectively. • \hat{f}_x, y: Inut datal of the coordinate x, y.

4.3 유사도 측정과 순위결정

앞장에서도 같이 분할된 영상에서 특징을 계산한 특성벡터를 사용하여 질의영상과 검색대상영상을 비교하여 유사한 영상을 선별하게 되는데, 이러한 과정을 유사도 측정과 순위결정과정이라 한다. 본 장에서는 유사도 측정이란 무엇이며, 유사도는 왜 측정해야 하는지? 에 대해서 연구하고 본 논문에서 사용하고 있는 유사도 측정방법을 기술한다.

4.3.1 유사도 측정

질의영상에 의해 추출된 특성벡터는 데이터베이스 내에 있는 영상들의 색인벡터와 비교함으로써 영상검색이 시작된다. 일반적인 검색시스템과는 달리 내용기반 영상검색은 정확한 검색을 요구하는 것이 아니고, 질의영상과의 유

사성의 순서에 따라 데이터베이스 내에 있는 영상을 순위별로 검색하여야 한다. 질의벡터와 색인벡터간의 유사한 정도는 유사도 측정 또는 유사도 계산과 정으로 수행한다[GUVIVADA95]. 즉, 일반적인 검색시스템에서는 정확히 일치하는 자료를 검색하지만, 내용기반 영상검색에서는 질의영상과 특성값이 정확히 일치하는 하나의 영상을 찾는 것이 아니라, 특성값이 유사한 다수의 영상들을 검색결과로 제시하여 사용자가 최종적으로 선택할 수 있도록 한다.

이와 같이 검색결과를 계산하는 방법은 지난 수십년간 정보검색 분야에서 문헌검색에 사용된 다양한 방법을 찾아 볼 수 있다. 이러한 다양한 방법 중에서 대표적인 『단순한 정합』, 『Dice 계수』, 그리고 『Jaccard 계수』에 의한 방법은 다음과 같이 나타낸다[RIJSBERGEN79]. 즉,

$$S_{sm}(Q, T) = |Q \cap T| \quad \text{단순한 정합} \quad (4-9)$$

$$S_{Dice}(Q, T) = 2 \frac{|Q \cap T|}{|Q| + |T|} \quad \text{Dice의 방법} \quad (4-10)$$

$$S_{Jaccard}(Q, T) = \frac{|Q \cap T|}{|Q \cup T|} \quad \text{Jaccard의 방법} \quad (4-11)$$

과 같으며, 여기서 Q 는 질의벡터, T 는 데이터베이스내의 색인벡터를 말하며, 식(4-9)은 질의벡터와 인덱스벡터에서 같은 특성을 가지는 인수들의 개수를 나타낸다. 다시 말하면, 영상에서 시각적으로 같은 색상을 가지는 화소들의 개수를 표시하는 것과 같다. 식(4-10)과 (4-11)는 식(4-9)에서 표현한 정합의 개념을 포함하고는 있지만 다른 방법임을 알 수 있다. 그림 4.9에서 이러한 방법들을 5×5의 질의벡터와 색인벡터간의 비교를 예로 들었다. 여기에서 단순한 정합은 같은 자료만 계수하여 18이며, Dice 및 Jaccard 방법은 0.72로서 비교대상간의 정합계수는 최대 1이하가 되도록 정규화하고 있다.

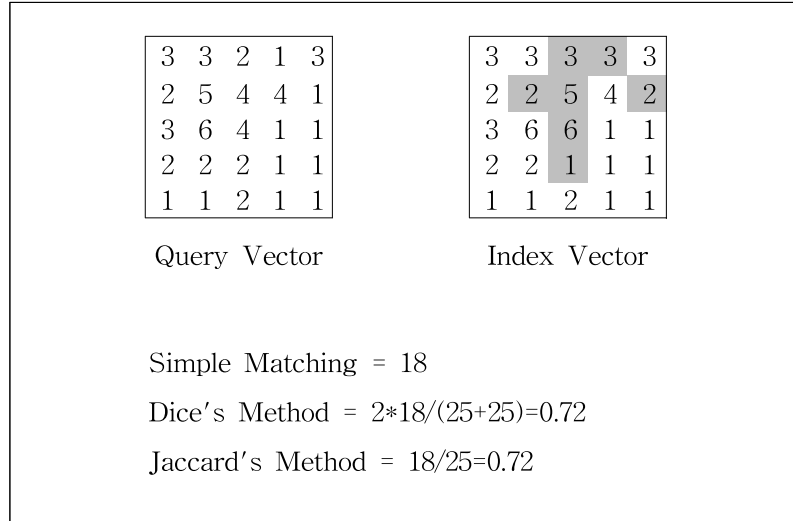


Fig. 4.9 Matching for information retrieval.

영상검색시스템에서 가장 많이 사용되는 유사도 측정방법은 식 (4-12)에서와 같이 벡터간의 거리를 계산하는 L_p norm 계산방법이다. 즉,

$$L_p(\overline{Q}, \overline{T}) = \left| \sum_i^n (Q_i - T_i)^P \right|^{1/P} \quad (4-12)$$

와 같으며, 두 영상간의 유사성을 측정하기 위한 함수는 유클리드 거리함수를 이용한다, 유클리드 거리함수의 특징은 정규직교변환 하에서 그 거리가 항상 보존된다는 것이다. 유클리드 거리를 두 벡터간의 거리로 사용하면, 위 식에서 P 의 값은 2로 할 수 있다. 그리하여 다음 식과 같이 유사도를 계산한다. 즉,

$$Distance_2 = \left| \sum_i^n (Q_i - T_i)^2 \right|^{1/2} \quad (4-13)$$

$$Similarity_2 = 1 - Distance_2 \quad (4-14)$$

$$Distance(Q, T) = \frac{w_1 D_1 + w_2 D_2}{w_1 + w_2} \quad (4-15)$$

과 같으며, 여기서 w_1, w_2 는 특성요소들간의 가중치를 나타낸다. 위의 식에서 알 수 있는 바와 같이 질의영상의 벡터와 색인벡터간의 차이를 식 (4-13)에서 구한 후, 그 값의 역으로 식 (4-14)과 같이 유사도를 측정한다. 특성벡터의 요소가 복합적인 요소로 구성되어 있으면 벡터간의 가중치를 주어 식(4-15)과 같이 정규화하여 유사도에 의한 검색순위를 결정한다.

본 논문에서 구성하고 있는 색인정보는 색상평균값과 거리 히스토그램이므로 두 영상간의 유사도를 측정하기 위해 특성벡터로 저장된 색상평균값과 거리 히스토그램에 의한 각각의 유사도를 계산한 후 이를 통합한 전체 유사도를 계산한다. 즉, 색상평균값은 식(4-16)처럼

$$dC(Q, T) = \sum_{k=1}^3 \sqrt{Rc_k^2 + Gc_k^2 + Bc_k^2} \quad (4-16)$$

이 된다. 또한 거리 히스토그램과 전체 유사도는 각각 식(4-17) 및 식(4-18)과 같이 나타내면,

$$dH(Q, T) = \sum_{n=1}^9 \left| \frac{H_{Q,n} - H_{T,n}}{H_{Q,n} + H_{T,n}} \right| \quad (4-17)$$

$$S_D = \frac{w_1 \cdot dC + w_2 \cdot dH}{w_1 + w_2} \quad (4-18)$$

와 같이 계산한다. 여기서 w_1, w_2 는 특성벡터간의 가중치, Q 및 T 는 각각 질의 영상과 검색대상영상이다. 또한 Rc_k, Gc_k, Bc_k 는 RGB 3채널에서 색상평균값의 거리를, H_n 는 거리 히스토그램을 표시한다. 그리고 dC 및 dH 는 각각 색상 및 거리 히스토그램에 대한 유사도이며, S_D 는 전체적인 유사도를 나타낸다.

4.3.2 순위 결정

순위결정이라 함은 검색된 영상이 유사도 측정에 의해 계산된 수치가 작은 순서에서 큰 순서로 나열하는 것이며, 검색결과에는 통상적으로 검색된 영상과 함께 순위 및 유사도 수치가 표시된다. 이러한 순위는 시스템의 성능을 평가할 때 매우 중요한 자료로 사용된다. 본 논문에서는 검색결과를 쉽게 찾아보며, 순위를 시각적으로 확인할 수 있는 영상표시 윈도우창을 그림 4.10과 같이 제작하였다. 그림에서 보는 바와 같이 좌측 첫째 영상이 질의영상이며, 우측으로 검색결과를 10장씩 보여 준다. 각 영상의 하단에는 영상의 이름과 계산된 유사도, 그리고 순위를 차례대로 표시한다. 영상의 개수가 많은 경우 검색된 영상을 찾아보기 쉽게 하려면 하단에 『바로가기』, 『처음』, 『끝』, 『이전』 과 『다음』의 단추를 준비하여 이동을 원활히 하였다. 또한 다른 시스템의 결과를 확인할 수 있도록 『다른 결과』 단추를 구비하였다.



Fig. 4.10 Window for displaying of retrieved result images.

4.4 성능 평가

영상검색시스템을 개발하는 과정에서 시스템의 구성과 영상내의 특징을 추출하는 방법의 개발도 중요하지만 이에 못지않게 시스템의 성능을 분석할 수 있는 성능평가과정은 매우 중요하다. 그러나 성능평가를 위해 사용하고 있는 평가도구가 문헌마다 달라서 일관성 있는 기준을 설정하지 못하고 있으며, 성능을 평가하기 위해 공통적으로 사용할 수 있는 실험영상도 존재하지 않아 일반적인 사진영상을 시험용으로 사용하고 있으므로, 이의 개선을 위한 많은 연구가 요구되고 있다. 그리하여 본 논문에서는 이러한 성능평가에 대한 개념을 정리하고 적절한 평가도구를 사용함으로써 제안된 시스템의 평가에 대한 객관성을 확보하고자 한다.

현실적으로 영상검색기술의 적용과 검색능력을 향상시키는데 장애물로 작용하는 여러 요인들이 존재하고 있어, 이 분야를 연구하는 연구자들은 다음과 같은 의문을 항상 가지게 된다. 즉,

- i) 영상검색 방법들의 성능은 어떻게 비교하는가?
- ii) 영상특징들은 어떻게 선택하며, 추출된 특성은 어떻게 평가하는가?
- iii) 새로운 알고리즘을 발표하는 논문은 어떻게 조사하는가?

iv) 영상검색기술과 시스템능력을 향상하기 위해 상호협조는 어떻게 이루어지는가? 하는 의문들은 지난 수십 년간에 걸쳐 문자검색, 데이터베이스 및 컴퓨터비전과 같은 다양한 분야에서 일어나고 있었다. 이러한 의문에 답하기 위해 문자검색용 테스트-베드를 준비하기 위한 노력을 TREC²¹⁾에서 수행하고 있으며, 또한 문자검색시스템의 성능향상과 평가를 위한 포럼을 운영하기도 한다[SMITH98]. 이러한 연구노력의 결과는 영상검색분야에도 많은 영향을 끼쳐, 표준영상과 벤치마킹할 수 있는 질의영상이 가까운 미래에 출현하리라 본다.

영상검색시스템은 영상들을 순위를 부여하여 검색하는 것이 목적이지만, 순위를 결정함에 있어 수작업으로 질의영상과의 관련성 여부를 결정해야 하는 주관적인 면이 있었다. 그럼에도 불구하고 시스템의 전체적인 검색효율성은 이러한 관련 있는 영상들을 모두 알고 있을 경우에만 측정이 가능하게 된다. 일반적으로 정보검색시스템의 테스트-베드는 다음과 같은 조건들이 구비되고

21) Text Retrieval Conference

있어야 한다. 즉,

i) 검색대상의 표준 자료 (문서, 영상 등)

ii) 표준 질의 자료

iii) 질의와 관련 있는 자료목록

iv) 평가도구 등이며, 정보검색에서 사용하고 있는 검색효율성 평가방법은 영상검색에서도 적용할 수가 있으며, 이에 대한 기본적인 개념을 그림 4.11에 표시한다.

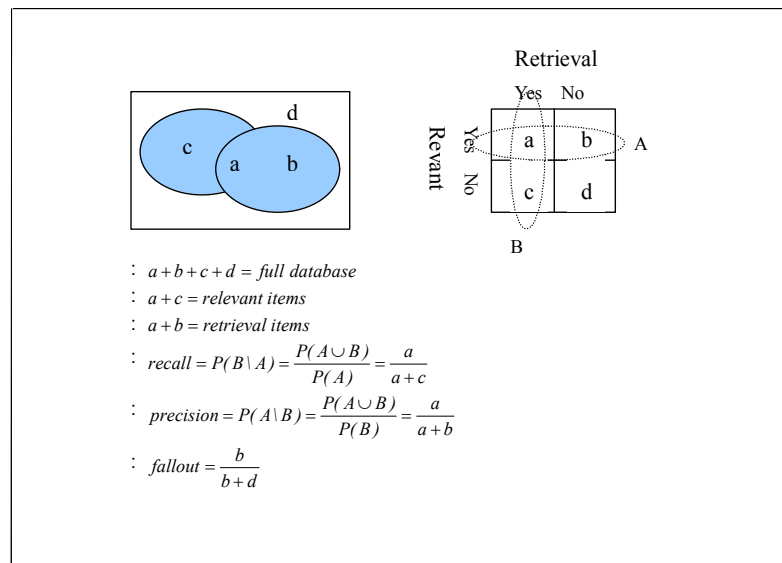


Fig. 4.11 Recall vs. Precision.

그림 4.11에서 a는 정확하게 검색된 정보의 집합, b는 검색은 되었으나 관련이 없는 것, c는 관련은 있으나 검색에 실패한 정보, 그리고 d를 검색되지도 않고 관련도 없는 정보의 집합이라고 하면, A는 검색된 정보의 집합(a+b), B는 관련이 있는 정보들의 집합(a+c)으로 볼 수 있다. 이러한 벤 다이어그램을 기본으로 재현율과 정확율 및 실패율(fallout)을 계산하면 그림과 같다.

정보검색을 위하여 시스템에 질의하면 순위를 부여하여 검색결과를 보여 주는데, 이때 k번째 되는 순위까지를 검색결과로 한다면 다음과 같은 값들을 구할 수 있다. 여기서 순위 n에 들어가는 정보의 집합을 V_n 이라면, $V_n \in \{0, 1\}$ 이며, 그림 4.11에서 a, b, c, d를 각각 A_k , B_k , C_k , D_k 로 다시 나타내면,

$A_k = \sum_{n=0}^{k-1} V_n$, $B_k = \sum_{n=0}^{k-1} (1 - V_n)$, $C_k = \sum_{n=0}^{N-1} V_n - A_k$, $D_k = \sum_{n=0}^{N-1} (1 - V_n) - B_k$ 로 표시할 수 있다. 이러한 값들로부터 표준 정보검색척도로 사용되는 재현율, 정확율 및 실패율을 계산할 수 있으므로 이를 식으로 나타내면 다음과 같다. 즉,

$$Recall : R_k = \frac{A_k}{A_k + C_k} \quad (4-19)$$

$$Precision : P_k = \frac{A_k}{A_k + B_k} \quad (4-20)$$

$$Fallout : D_k = \frac{B_k}{B_k + D_k} \quad (4-21)$$

이 된다. 위의 식에 의하여 정보검색시스템의 효율성을 측정할 수 있으며, 다른 시스템과도 비교할 수 있다.[SMITH96b, 98].

정보검색분야에서 사용하고 있는 개념을 영상검색시스템의 효율성을 평가하기 위해 도입한 다양한 기법들이 있다. 즉, 본 논문에서 다루고 있는 내용기반 영상검색은 정보검색에서와 같이 정확한 검색을 요구하는 것이 아니라 유사한 영상을 검색하게 되므로 앞에서 언급한 유사도 측정과 연관하여 평가 방법을 고려해야 한다. 재현율, 정확율 및 실패율을 사용하여 영상검색시스템의 성능을 평가하는 방법들은 다음과 같이 다양한 방법이 있다.

- i) P_k vs. R_k : 검색시스템의 효율성을 나타내며 가장 많이 사용되는 도구로써 신뢰성이 높음.
- ii) A_k vs. B_k : 검색 성공률.
- iii) A_k vs. F_k : 상대적 검색 성공률.
- iv) P_k at cut-off $k = \text{int}(\sum_{n=0}^{N-1} V_n)$: 관련되는 영상을 최하위까지 고려.
- v) average P_k at $R_k = 0.2, 0.5, 0.8$: recall 0.2, 0.5, 0.8인 경우.

vi) average P_k at eleven recall point : 11개의 recall점.

vii) $AVRR$ (average rank of all relevant items) : 유사한 영상이 검색된 순위들의 평균값으로 질의영상과 가장 유사한 영상의 순위는 “0”임.

viii) $AVRR/IAVRR$ (ideal average relevant rank) : $IAVRR$ 은 유사한 영상들이 상위에서 차례로 검색되는 경우를 말한다.

즉, $IAVRR = (0 + 1 + 2 + \dots + (N-1))/N$ 이다.

ix) $\frac{A_k}{B_k}$ vs. A_k : 응답률.

이상과 같은 평가도구들은 대부분 재현율, 정확율과 실패율을 이용하여 계산하게 된다. 또한 이러한 도구들은 공통적인 척도와 함께 상호 연관성을 가지고 평가해야만 올바른 성능평가가 이루어 질 수 있다. 그런 의미에서 정보 검색에서 표준으로 사용되고 있는 평가방법인 재현율 대 정확율(Recall vs. Precision)은 많은 장점을 지니고 있다. 그림 4.12는 제안한 알고리즘의 유효성 증명하기 위해 본 논문에서 사용했던 방법으로서, 범례에서와 같이 파란색은 제안한 방법, 붉은 색은 Swain, 노랑은 CCV, 그리고 연 파랑은 Correlogram 방법에 의한 그래프를 그린 것이다.

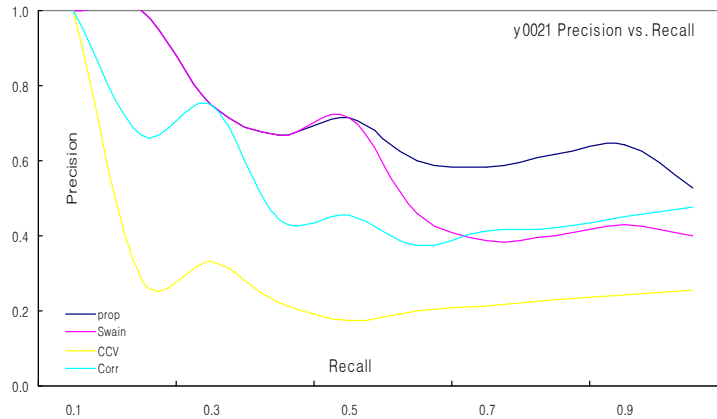


Fig. 4.12 Illustration of Recall vs. Precision.

이 그림에서 알 수 있는 바와 같이 재현율에 대한 정확율의 값이 높으면, 상대적으로 더 나은 시스템이라 평가한다. 즉, 모든 재현율이 가지는 값에 대

하여 정확율의 값이 높으면 비교하는 시스템보다 성능이 좋다고 평가한다. 이러한 평가도구에 의해 영상검색시스템을 평가할 때 고려해야할 중요한 문제는 다음과 같이 두 가지로 생각할 수 있다.

첫째는, 질의영상에 대해 유사한 영상을 판단함에 있어 개인의 주관이 개입된다는 점이며,

두 번째는, 대용량의 데이터베이스에서 시험해야할 경우에는 유사한 영상을 별도로 구분해야 한다는 것이다. 영상의 개수가 많아지면 실험을 하기 위해 유사한 영상을 일일이 구분한다는 것은 보통의 작업이 아니므로 본 논문에서는 다음과 같은 식을 사용하여 성능평가를 수행한다. 즉,

$$Eff = \begin{cases} \frac{n}{N}, & \text{if } N \leq T \\ \frac{n}{T}, & \text{otherwise} \end{cases} \quad (4-22)$$

과 같다. 여기서 n 은 검색된 유사한 영상의 개수이고, N 은 유사한 영상의 개수이며, T 는 검색결과로 고려할 순위를 나타낸다. 식에서 보는 바와 같이 $N \leq T$ 이면 재현율과 같으며, $N > T$ 이면 앞서 언급한 정확율과 같은 값을 가지게 된다[TAO99b].

V. 실험결과 및 분석

본 논문에서 제안하고 있는 색상-공간정보를 고려하기 위한 영상검색의 유효성을 실험하기 위하여 Windows 98을 운영체제로 하는 PentiumIII 650MHz를 사용하였으며, 실험용 시스템을 구현하기 위해 사용된 언어는 Microsoft Visual C++ 6.0 컴파일러이다. 또한 검색된 영상을 출력하기 위한 영상표시 윈도우창은 Visual Basic을 사용하여 제작하였으며, 제안된 방법에 의해 구현한 시스템의 초기화면은 그림 5.1에 나타내었다. 또한 그림 5.4에서 그림 5.10에 표시된 영상표시 윈도우창은 주관적인 성능평가를 하기 위해 검색결과를 시각적으로 비교하는데 이용하였다.

실험에 사용된 영상은 버클리 대학의 『BlobWorld』 시스템에서 사용한 총 418개의 JPEG영상을 192×128크기의 TIFF 영상으로 변환하여 사용하였으며, 질의방법은 본문에서 언급한 바와 같이 예제영상에 의해 질의하였다. 질의영상의 개수는 전체영상에서 53개의 영상을 임의로 선택하여 실험하였으며, 사용된 실험영상과 질의영상의 종류를 분류하면 표 5.1과 같다.

제안하고 있는 색상-공간정보를 고려하기 위한 과정으로 영상의 개체영역을 구분하기 위하여 순차영역분할방법을 사용하였다. 순차영역분할과정에서 TH값은 RGB 각 3채널 모두 실험에 의해 선정한 55이며, 개체영역을 분할하기 위한 영역의 개수는 3개로 고정하였다. 영역의 개수를 3개 이상으로 한 경우 검색성능에 별로 영향을 미치지 않았을 뿐 아니라, 분할영역이 많아지면 색인용량이 커지는 단점이 있으므로 본 실험에서는 영역의 개수를 최소화하였다. 제안한 알고리즘에 의해 분할된 영역의 색상과 합성영상 및 원 영상을 그림 5.2에 나타내었다. 그림에서 보는 바와 같이 각 영역의 색상과 개체구분이 확실히 되고 있음을 확인할 수 있다.

Table 5.1 Classification of image database.

Items	Relevant images	Query images
Bird	34	2
Butterfly	50	9
Landscape	68	17
People	38	4
Tree	48	3
Flower	56	6
Animal	83	5
Sea, Fall	16	4
Fish	12	2
The others	13	1
Total	418	53

본 논문에서는 분할된 영역의 색상평균값을 특성벡터의 첫 번째 색인정보로 이용하였으며, 두 번째 정보로는 공간정보를 확보하기 위한 방법으로 거리 히스토그램을 사용한다. 거리 히스토그램은 분할된 영역에서 영상의 중심점으로부터 가까운 구간을 0.3, 중간 위치를 0.6, 그 외를 먼 구간으로 계산하였다. 그리고 색상평균값과 거리 히스토그램 사이의 유사도의 가중치인 w_1 , w_2 는 각각 4, 2.5를 주어 실험하였다. 여기에서 거리 히스토그램을 계산하기 위한 구간과 가중치는 실험에 의해 선정한 값이다.

제안한 시스템을 평가하기 위한 방법으로, 주관적인 평가와 객관적인 평가를 병행하였다. 주관적인 평가는 검색된 결과를 영상표시 윈도우창에서 시각적으로 직접 확인하는 방법을 사용하였으며, 객관적인 평가는 식 (4-19), (4-20) 및 (4-22)을 사용하였다. 또한 평가의 객관성을 유지하기 위해 동일한 영상자

료로서 제안한 방법과 Swain에 의한 방법, CCV 및 Correlogram에 의한 방법들과 비교하였다.

제안한 방법은 RGB공간에서 특성벡터를 추출하였으며, 이러한 색인정보는 분할된 3영역에서 각각 평균색상 3개와 각 영역에서 계산한 9개의 거리 히스토그램을 합쳐 총 18개의 요소를 특성벡터로 하고 있다. 반면, Swain에 의한 히스토그램 인터섹션방법은 opponent RGB공간에서 rg축과 by축을 각각 16등분하였고 wb축을 8등분함으로서 $16 \times 16 \times 8 = 2048$ 빈이 되도록 양자화하여 이를 색인정보로 사용하고 있다. Pass 및 Huang에 의한 CCV, Correlogram은 RGB공간을 $4 \times 4 \times 4 = 64$ 빈이 되도록 양자화를 같은 방법으로 하고 있는 반면, CCV에서는 색상별로 영상내의 색상이 응집 또는 분산되어 있는지를 구분하여 128개의 요소를 특성벡터로 사용하고 있고, Correlogram에서는 화소간의 거리(distance set) $D = \{1, 3, 5, 7\}$ 을 주어 계산하므로 64×4 즉, 256개의 요소로 이루어진 특성벡터를 사용하고 있다. 제안한 방법과 Swain의 방법, CCV 및 Correlogram에 의한 방법들을 비교할 수 있도록 요약하면 표 5.2와 같다. 표에서 색인용량은 총 418개의 영상을 색인하는데 소요되는 저장공간을 말하며, 파일명 4,180바이트(10바이트 \times 418영상)를 포함한 공간이다.

시스템의 성능평가방법으로 식(4-22)의 Eff 를 사용하며, 평가의 객관성을 확보하기 위해 임의로 선택한 53개의 질의영상에 대하여 실험하였다. 각각의 질의영상에 대한 5위, 10위, 15위, 20위에서 실험한 검색결과를 전체적으로 볼 수 있도록 한 것이 표 5.4이며, 이에 대한 평균을 표 5.3에 표시하였다.

Table 5.2 Methods used in the evaluation.

Method	Color Space	Feature	Distance	Cluster	Storage (byte)
Proposeed	RGB	18	L2	3	52,672
CCV	RGB	128	L1	No	226,560
Correlogram	RGB	256	L1	No	547,584
Swiain'	opponent RGB	2048	Histogram Intersection	No	3,456,800

표 5.3에서 보는 바와 같이 본 논문에서 제안한 방법이 다른 방법보다 전체적으로 검색된 순위에 관계없이 우수한 성능을 보이고 있다. 실험결과에서 제안한 방법은 5위에서 83.02 %, 10위 79.41 %, 15위 86.46 % 및 20위에서 91.65 %까지 유사한 영상을 검색하고 있으나, 다른 방법들은 상대적으로 낮은 검색율을 보이고 있다. 특히 20위에서는 Correlogram은 83.14 %, CCV는 75.49 %, Swain에 의한 방법은 69.16 %를 보이고 있어 검색성능이 제일 우수한 Correlogram보다 제안한 방법이 8.5 % 더 성능개선을 이루고 있어, 전체적으로 검색순위에 관계없이 매우 우수함을 확인할 수 있었다.

또한 다른 평가방법으로 식(4-19) 및 (4-20)에 의한 재현율 대 정확율을 계산 하여 질의영상 중 "Buffalo(y0021)", "Butterfly(y0174, y0175)", "Landscape(y0209, y0237)" 및 "Flower(y0407)"영상에 대한 그래프를 그림 5.3에 표시하였다. 그림에서 보는 바와 같이 각 각의 재현율에 대한 정확율의 값이 다른 방법과 비교하여 높게 나타나고 있어 전반적으로 이 실험에서도 제안한 방법이 우수함을 알 수 있었다.

Table 5.3 Performance comparison by *Eff.*

Method	5th Rank	10th Rank	15th Rank	20th Rank
Proposed	83.02	79.41	86.46	91.65
Swain'	62.99	58.35	61.54	69.16
CCV	69.97	64.38	67.42	75.49
Correlogram	75.66	69.55	73.62	83.14

주관적인 평가하기 위해 영상표시 윈도우창에 표시한 그림을 그림 5.4에서 그림 5.10까지 타 방법에서 검색된 결과와 함께 나타내었다. 영상표시 윈도우창은 좌측 첫 번째 영상이 질의영상이며, 우측으로 나열된 순서에 따라 검색된 순위가 높아지게 되며, 질의영상과도 거리가 멀어진다.

그림 5.4는 "Starfish(y0096)"에 대한 10위권 내에 검색된 영상들을 보이고 있다. 그림에서 보는 바와 같이 제안한 방법은 유사한 영상 모두가 5위권 내에서 검색되었지만, Correlogram에서는 y0095는 3위, y0094는 5위에 검색되었으며, CCV에서는 44위, 70위로, 그리고 Swain의 방법은 29위, 152위로 각각 검색되었다. 따라서 제안한 방법은 "Starfish"영상을 상위순위에서 모두 검색하였지만, 다른 방법에서는 Correlogram에 의한 방법을 제외하고 상위순위의 범위를 벗어나는 영상이 많았다. 마찬가지로 그림 5.5는 "Fish(y0089)" 영상에 대한 결과를 보여 준다. 여기에서도 제안한 시스템의 우수성을 확인 할 수 있었다. 즉, 제안한 방법은 검색순위가 1위, 2위 3위, 4위, 6위 및 10위로 검색되어 유사한 영상 모두가 상위권에 출현하였으나, Swain은 검색순위가 그림에서 1/2/7/8/3/66, CCV는 1/2/15/8/3/80, Correlogram은 1/2/3/16/4/22의 순위로 검색되었다. 그림 5.6에서부터 그림 5.10까지 나타난 결과에서도 마찬가지로 제안한 방법에서 검색순위가 앞서므로, 시각적인 면에서도 제안한 방법의 성능이 우수함을 확인할 수 있었다.

Table 5.4 Retrieved results for each query image.

NAME 번호	RANK 5						RANK 10						RANK 16						RANK 20					
	Prop	SWAIN	OOD	OOD	Prop	SWAIN	OOD	OOD	Prop	SWAIN	OOD	OOD	Prop	SWAIN	OOD	OOD	Prop	SWAIN	OOD	Prop	SWAIN	OOD	OOD	OOD
1 Y0013	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80
2 Y0016	6	100	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80
3 Y0020	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80
4 Y0021	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80
5 Y0032	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80
6 Y0044	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80
7 Y0055	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80
8 Y0058	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80
9 Y0069	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80	4	80
10 Y0080	2	60	4	100	2	60	4	100	2	60	4	100	2	60	4	100	2	60	4	100	2	60	4	100
11 Y0096	3	100	1	33	1	33	1	33	1	33	1	33	1	33	1	33	1	33	1	33	1	33	1	33
12 Y0098	2	100	1	60	1	60	1	60	1	60	1	60	1	60	1	60	1	60	1	60	1	60	1	60
13 Y0112	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100
14 Y0122	3	80	2	40	3	80	2	40	3	80	2	40	3	80	2	40	3	80	2	40	3	80	2	40
15 Y0137	2	40	3	80	2	40	3	80	2	40	3	80	2	40	3	80	2	40	3	80	2	40	3	80
16 Y0147	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100
17 Y0154	6	100	3	80	3	80	4	80	3	80	4	80	3	80	4	80	3	80	4	80	3	80	4	80
18 Y0158	4	80	3	80	3	80	4	80	3	80	4	80	3	80	4	80	3	80	4	80	3	80	4	80
19 Y0163	6	100	6	100	6	100	6	100	6	100	6	100	6	100	6	100	6	100	6	100	6	100	6	100
20 Y0164	3	80	6	100	3	80	6	100	3	80	6	100	3	80	6	100	3	80	6	100	3	80	6	100
21 Y0174	3	76	1	26	3	76	1	26	3	76	1	26	3	76	1	26	3	76	1	26	3	76	1	26
22 Y0175	4	100	1	26	2	60	2	60	4	100	1	26	1	26	3	76	1	26	3	76	1	26	3	76
23 Y0177	4	100	2	60	3	76	2	60	4	100	2	60	3	76	2	60	4	100	2	60	3	76	2	60
24 Y0185	6	100	1	20	5	100	5	100	9	80	4	40	6	80	7	70	12	88	5	36	9	84	11	79
25 Y0187	4	100	1	26	1	26	4	100	4	100	1	26	3	76	4	100	4	100	1	26	3	76	4	100
26 Y0203	3	80	2	40	2	40	1	20	8	80	2	20	4	40	3	30	10	71	4	29	6	36	7	50
27 Y0215	4	80	4	80	6	100	4	80	8	80	8	80	12	76	8	60	11	69	10	63	16	100	11	69

Table 5.4 Retrieved results for each query image (continued).

번호	NAME	RANK 5						RANK 10						RANK 15						RANK 20															
		SWAIN			OOD			SWAIN			OOD			SWAIN			OOD			SWAIN			OOD												
		Prop	8WAIN	OOD	OOD	Prop	8WAIN	OOD	OOD	Prop	8WAIN	OOD	OOD	Prop	8WAIN	OOD	OOD	Prop	8WAIN	OOD	OOD	Prop	8WAIN	OOD	OOD										
28	Y0218	4	80	5	100	5	100	5	100	7	70	10	100	10	100	8	80	10	67	12	80	11	73	12	80	15	100	12	80	14	93	15	100		
29	Y0220	5	100	5	100	4	80	5	100	10	100	8	80	8	80	9	90	12	86	11	79	13	93	11	79	14	100	11	79	14	100	14	100		
30	Y0222	4	80	3	80	5	100	5	100	10	100	4	40	10	100	9	90	15	100	9	80	14	93	14	93	20	100	12	80	20	100	19	95		
31	Y0227	4	80	4	80	5	100	5	100	8	80	8	80	8	80	10	100	12	80	11	73	12	80	13	87	15	100	13	87	15	100	15	100		
32	Y0228	4	80	4	80	5	100	5	100	10	100	8	80	10	100	9	90	15	100	9	80	14	93	11	73	18	100	14	78	18	100	15	83		
33	Y0229	4	80	3	80	5	100	3	80	10	100	3	30	7	70	2	20	15	100	6	40	8	53	2	13	20	100	8	40	10	50	2	10		
34	Y0231	4	80	4	80	5	100	4	80	9	90	7	70	8	80	8	80	14	93	8	53	12	80	10	87	16	100	9	56	16	100	13	81		
35	Y0232	5	100	4	80	4	80	3	80	10	100	7	70	7	70	5	50	14	93	9	60	14	93	6	40	14	74	9	47	17	89	19	100		
36	Y0235	5	100	4	80	5	100	5	100	9	90	7	70	8	80	9	90	12	80	10	67	12	80	14	93	13	87	13	87	13	87	15	100		
37	Y0236	2	40	2	40	3	60	2	40	5	83	4	67	5	83	3	50	5	83	4	67	5	83	4	67	6	40	6	100	6	100	6	100		
38	Y0237	5	100	3	80	5	100	5	100	10	100	6	60	8	80	9	90	14	93	9	60	14	93	14	93	16	100	11	69	16	100	16	100		
39	Y0240	5	100	3	80	4	80	5	100	7	70	4	40	9	90	8	80	10	71	8	57	12	86	11	79	14	100	8	57	14	100	13	93		
40	Y0248	5	100	2	40	4	80	4	80	9	90	6	60	6	60	9	90	14	93	8	53	10	87	13	87	14	74	19	100	13	88	16	84		
41	Y0262	3	80	3	80	3	80	3	80	4	80	4	80	3	80	3	80	5	100	4	80	4	80	4	80	4	80	4	80	4	80	4	80	5	100
42	Y0276	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100	2	100
43	Y0280	4	80	3	80	4	80	4	80	7	88	3	38	6	76	4	50	8	100	3	38	6	76	6	76	7	88	4	50	6	76	7	88		
44	Y0302	4	80	4	80	2	40	4	80	7	88	5	63	5	63	8	100	8	100	6	76	6	76	8	100	8	100	6	76	6	76	8	100		
45	Y0320	3	76	2	50	2	50	3	76	3	76	3	76	2	50	3	76	3	76	3	76	2	50	4	100	3	76	3	76	3	76	4	100		
46	Y0330	3	60	3	60	2	40	2	40	3	60	2	40	2	40	2	40	2	40	4	80	3	60	4	80	4	80	4	80	5	100	3	60	3	60
47	Y0362	3	60	1	20	3	60	3	60	4	40	1	10	6	60	3	30	5	50	1	10	7	70	5	50	4	40	1	10	10	100	8	80		
48	Y0373	4	80	4	80	5	100	5	100	9	90	6	60	6	60	8	80	11	100	8	73	7	64	10	61	11	100	8	73	7	64	11	100		
49	Y0374	5	100	2	40	4	80	5	100	8	80	2	20	8	80	9	90	14	100	3	21	9	64	12	86	14	100	3	21	9	64	14	100		
50	Y0375	5	100	2	40	4	80	5	100	9	90	4	40	6	60	9	90	13	87	5	33	9	60	14	93	14	88	5	31	10	83	16	100		
51	Y0379	3	80	2	40	2	40	3	80	3	33	2	22	3	33	4	44	8	89	3	33	3	33	3	33	7	78	9	100	3	33	4	44	9	100
52	Y0407	4	80	4	80	4	80	4	80	9	90	8	80	7	70	7	70	14	93	9	60	7	47	7	47	15	94	16	100	7	44	8	50		
53	Y0408	5	100	4	80	4	80	5	100	9	90	8	80	6	60	9	90	13	87	9	60	9	60	11	73	13	87	15	100	9	60	12	80		
평균		83.02	82.99	89.97	75.86	79.41	58.35	64.38	69.55	86.46	61.54	67.42	73.62	81.65	69.16	75.49	83.14																		

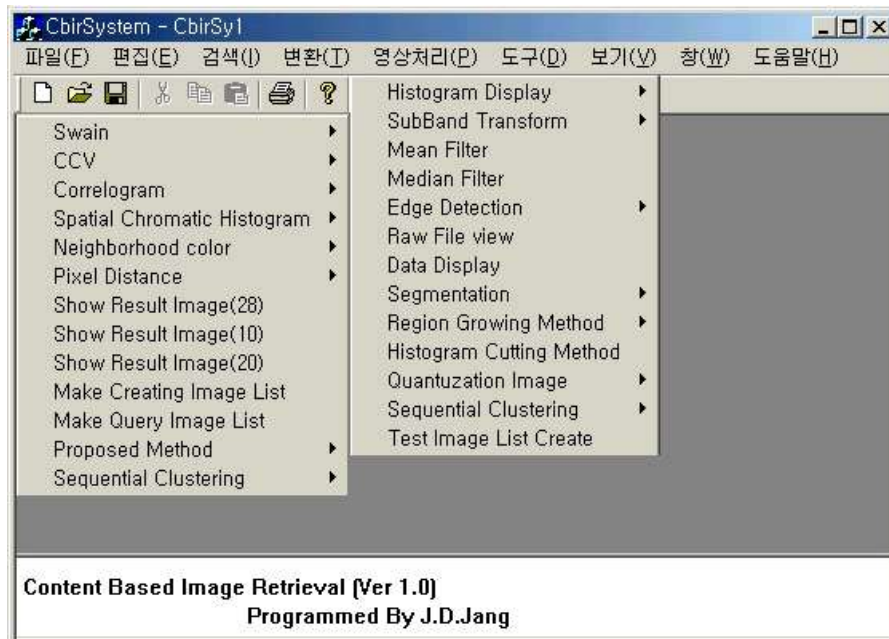


Fig. 5.1 Initial menu display of image retrieval system considered.

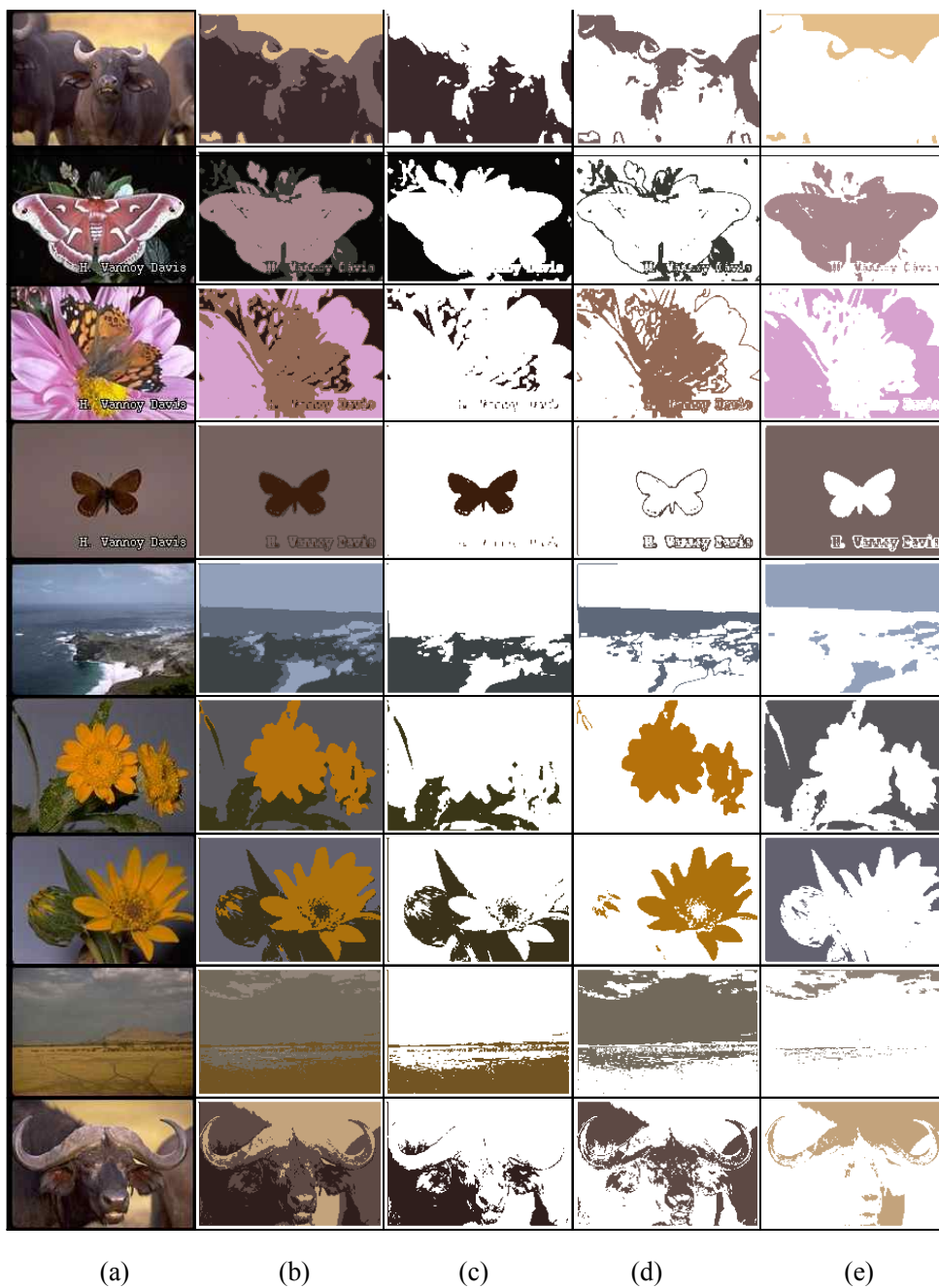
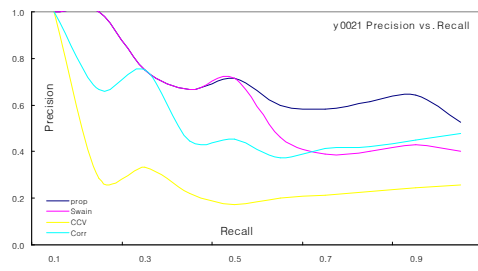
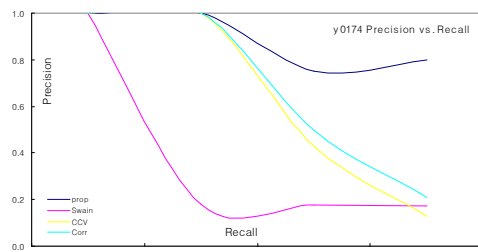


Fig. 5.2 Clustered image samples using the Sequential-Clustering.

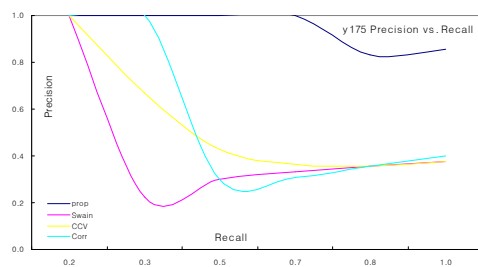
(a) Original (b) Composited clusters (c) 1st cluster (d) 2nd cluster (e) 3rd cluster



(a)



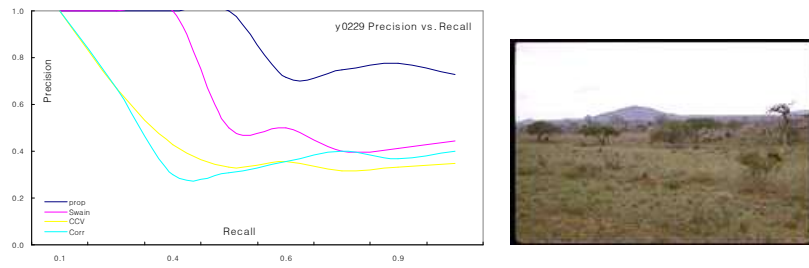
(b)



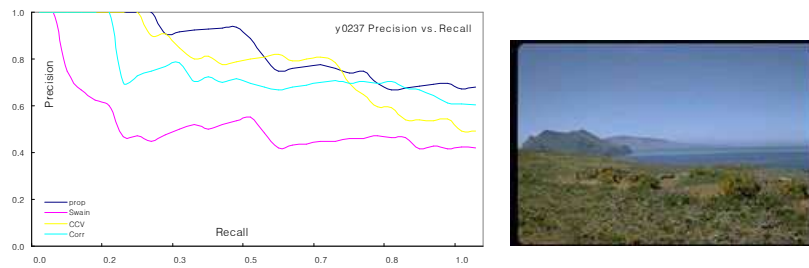
(c)

Fig. 5.3 Results of Precision vs. Recall.

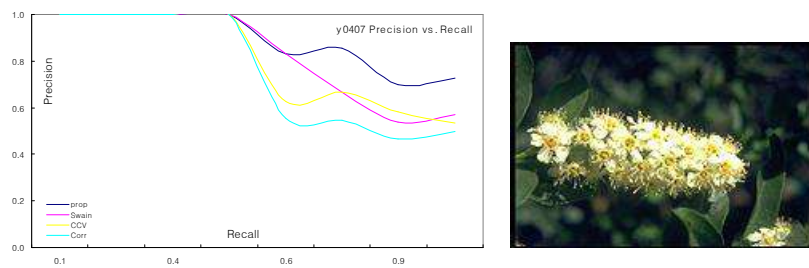
(a) "y0021" (b) "y0174" (c) "y0175"



(d)



(e)



(f)

Fig. 5.3 Results of Precision vs. Recall (continued).

(d) "y0229" (e) "y0237" (f) "y0407"



(a)



(b)

Fig. 5.4 Retrieved results for "Starfish (y0096)" image.

(a) Proposed System (1/2/3) (b) Swain (1/29/152)



(c)



(d)

Fig. 5.4 Retrieved results for "Starfish (y0096)" image (continued).

(c) CCV (1/44/70) (d) Correlogram (1/3/5)



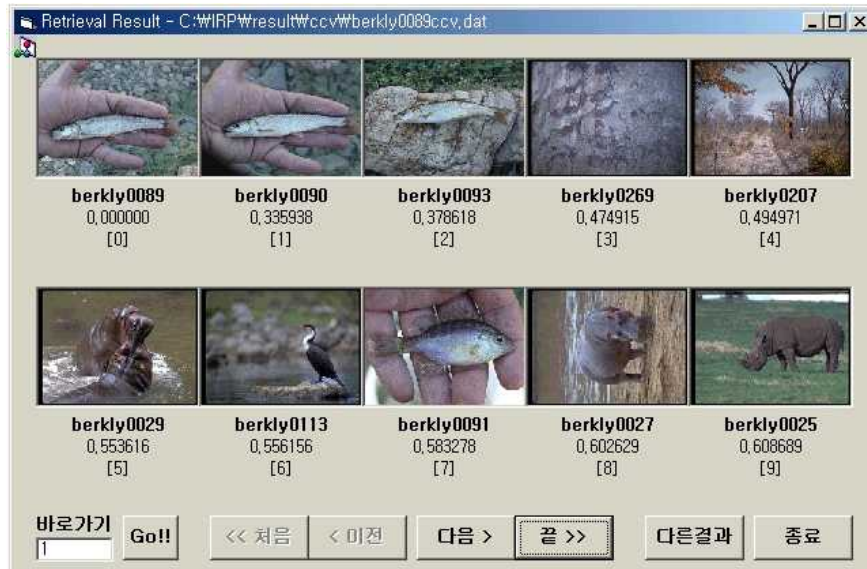
(a)



(b)

Fig. 5.5 Retrieved results for "Fish (y0089)" image.

(a) Proposed System (1/2/3/4/6/10) (b) Swain (1/2/7/8/3/66)



(c)



(d)

Fig. 5.5 Retrieved results for "Fish (y0089)" image (continued).

(c) CCV (1/2/15/8/3/80) (d) Correlation (1/2/3/16/4/22)



(a)



(b)

Fig. 5.6 Retrieved results for "Butterfly (y0154)" image.

(a) Proposed System (1/2/3/4/5/6/7/9/10) (b) Swain (1/2/3/6/8/10...)



(c)



(d)

Fig. 5.6 Retrieved results for "Butterfly (y0154)" image (continued).

(c) CCV (1/2/4/9..) (d) Correlogram (1/2/3/4/6/..)



(a)



(b)

Fig. 5.7 Retrieved results for "Flower (y0158)" image.

(a) Proposed System (1/2/3/4/9) (b) Swain (1/3/5/11/235)



(c)



(d)

Fig. 5.7 Retrieved results for "Flower (y0158)" image (continued).

(c) CCV (1/2/4/17/229) (d) Correlogram (1/2/3/11/142)



(a)



(b)

Fig. 5.8 Retrieved results for "Butterfly (y0177)" image.

(a) Proposed System (1/2/3/4) (b) Swain (1/2/59/91)



(c)



(d)

Fig. 5.8 Retrieved results for "Butterfly (y0177)" image (continued).

(c) CCV (1/2/3/147) (d) Correlogram (1/2/13/30)



(a)



(b)

Fig. 5.9 Retrieved results for "Landscape (y0197)" image.

(a) Proposed System (1/2/3) (b) Swain (1/38/67)



(c)



(d)

Fig. 5.9 Retrieved results for "Landscape (y0197)" image (continued).

(c) CCV (1/3/7) (d) Correlogram (1/2/4)



(a)



(b)

Fig. 5.10 Retrieved results for "Flower (y0373)" image.

(a) Proposed System (1/2/3/4/5/6/7/8/9/10) (b) Swain (1/2/4/5/7/..)



(c)



(d)

Fig. 5.10 Retrieved results for "Flower (y0373)" image (continued).

(c) CCV (1/2/3/4/5..) (d) Correlogram (1/2/3/4/5/6/7/8/..)

VI. 결 론

영상검색방법 중에서 현재까지 가장 보편적으로 사용되고 있는 문자기반 검색은 주석을 수작업으로 진행해야 하는 단점 외에도 여러 가지 문제점을 지니고 있다. 그리하여 영상의 이름이나 주석을 기억하지 않아도 필요한 영상을 검색할 수 있으며, 주제어로 표현할 수 없는 영상의 고유한 특성까지 검색에 이용할 수 있는 영상의 내용에 기반 한 접근방법이 필요하게 되었다. 그리하여 본 논문에서는 영상의 특징을 자동으로 추출할 수 있으며, 영상의 속성을 근거로 유사한 영상을 검색하는 칼라영상의 내용을 기반으로 한 영상검색방법에 대하여 연구하였다.

내용기반 영상검색기법의 연구 중에서 가장 핵심부분은 영상내의 특징을 추출하여 특성벡터를 구성하는 과정이다. 그러나 이 과정에서 가장 많이 사용되고 있는 히스토그램방법은 영상이 다르더라도 유사한 히스토그램을 가질 수 있으므로 검색되는 영상의 개수가 많아지면 질의영상과는 전혀 유사하지 않는 영상을 검색하였다. 이러한 히스토그램방법에서 검색성능을 저하시키고 있는 요인은 히스토그램 특성상 영상자체에서 지니고 있는 개체의 위치 또는 개체간의 공간적인 정보를 포함할 수 없다는 문제점이었다. 본 논문은 이와 같이 기존의 히스토그램방법에서 공간정보를 포함할 수 없다는 문제점을 해결하기 위하여 영상의 개체를 분할하며, 분할된 영역에서 색상과 공간정보를 함께 얻을 수 있는 방법을 제안하였다.

제안한 내용 중에서 영상의 개체를 분할하는 방법으로는 영상내의 유사한 색상을 가지는 화소들을 같은 그룹으로 군집하는 패턴인식분야에서 사용하고 있는 순차영역분할기법을 영상검색에 도입하였으며, 분할된 영역에서 특성벡터를 추출하는 방법으로는 영역의 중심점으로부터 화소들의 떨어진 정도를 구분하여 거리 히스토그램을 이용함으로써 공간정보를 확보하였다. 그리고 색상 정보는 각각의 영역에서 포함하고 있는 색상의 평균값을 특성벡터로 색인함으로써 시스템의 구현이 용이하고, 계산이 단순한 색상-공간정보를 함께 고려할 수 있는 영상검색방법을 제안하였다.

제안한 방법의 유효성을 증명하기 위하여 주관적인 방법과 객관적인 평가

도구를 병행하여 사용하였으며, 비교대상으로는 기존의 연구 중에서 색상정보만을 고려한 Swain의 히스토그램 인터섹션방법과 색상과 공간정보를 함께 포함한 Pass 등이 제안한 CCV 및 Huang의 Correlogram 방법을 이용하였다.

제안한 방법에서 색인하는 특성벡터는 색상평균값 9개와 거리 히스토그램 9개를 합하여 총 18개의 요소를 사용하고 있는 반면, 다른 방법인 Swain에서는 2048, Correlogram은 256, 그리고 CCV는 128개의 특성벡터를 사용하고 있음에도 불구하고, 제안한 방법이 다른 방법보다 작은 저장 공간을 가지면서도 *Eff*에 의한 성능평가에서 5위, 10위, 15위 및 20위의 순위에서 모두 성능이 우수하였다. 특히 순위 20위의 검색결과에서 Swain은 69.16 %, CCV는 75.49 %, Correlogram은 83.14 %로 나타나고 있으나, 제안한 방법은 91.65 %의 성능을 보였다. 또한 재현율 대 정확율의 평가에서도 각 질의영상에서 대부분의 재현율에서 정확율의 값이 상위를 점하고 있어 제안한 방법이 모두 우수함을 알 수 있었으며, 주관적인 평가에서도 제안한 방법의 우수성을 영상표시 윈도 우창을 통해 시각적으로 확인할 수 있었다.

참고 문헌

- [BACH96] J. R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, and C.-F. Shu, "The Virage Image Search Engine: An Open Framework for Image Management," *ISPIE Proceedings Vol. 2670*, pp. 76-87, 1996.
- [BELONIE98] S. Belongie, C. Carson, H. Greenspan, and J. Malik, "Color and Texture-Based Image Segmentation Using the Expectation-Maximization Algorithm and Its Application to Content-Based Image Retrieval," *International Conference on Computer Vision (ICCV'98)*, Bombay, India, January 1998.
- [BIMBO96] A. Del Bimbo, "Image and Video Database: Indexing, Browsing and Retrieval," *Journal of Visual Language and Computing*, 7, pp. 353-359, 1996.
- [BIMBO97] A. Del Bimbo and P. Pala, "Visual Image Retrieval by Elastic Matching of User Sketches," *IEEE Trans. Pattern analysis and Machine Intelligence*, 19(2), February 1997.
- [CARSON98] C. Carson, S. Belongie, H. Greenspan, and J. Malik. "Color-and Texture-based image segmentation using EM and its application to image querying and classification," Manuscript, submitted to *pattern Analysis and Machine Intelligence*, 1998.
- [CASTLMAN96] K. R. Castleman, *Digital Image Processing*, New Jersey: Prentice-Hall, 1996.
- [CHANG87] S. K. Chang, Q. Y. Shi, and C. W. Yan, "Iconic indexing by 2-D strings," *IEEE Trans. Patt. Anal. Mach. Intell. PAMI-9(3)*, pp. 413-428, May 1987.
- [CHANG96] S.-F. Chang, A. Eleftheriadis, and D. Anastassiou, "Development of columbia's video on demand testbed," *Image Communication*, 1996.
- [COHEN99] Scott Cohen, "FINDING COLOR AND SHAPE PATTERNS IN IMAGES," Ph.D Thesis, Stanford University, 1999.
- [COX96a] I. J. Cox, M. L. Miller, S. M. Omohundro, and P. N. Yianilos, "PicHunter: Bayesian Relevance Feedback for Image Retrieval," In: *Proceedings of International Conference on Pattern Recognition*. Vienna, Austria, pp. 361-369, August 1996.
- [COX96b] I. J. Cox, M. L. Miller, S. M. Omohundro, and P. N. Yianilos, "Target Testing and the PicHunter Bayesian Multimedia Retrieval System," In:

- Proceedings of Third Forum on Research & Technology Advances in Digital Libraries*, IEEE, pp. 66-75, 1996.
- [DELODDERE94] D. Deloddere, W. Verbiest, and H. Verhille. Interactive video on demand, May 1994.
- [DIMAI97] A. Dimai, "Spatial Encoding Using Differences of Global Features," *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, Volume 3022, February 1997.
- [DENG99] Y. Deng and B.S. Manjunath, "An efficient low-dimensional color indexing scheme for region-based image retrieval", *Proc. of IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 3017-20, 1999.
- [FALOUTSOS95] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equiz, "Efficient and Effective Querying by Image Content," *Journal of Intelligent Information System(JIIS)*, pp. 231-262, July 1994.
- [FRANKEL96] C. Frankel, M. Swain, and V. Athitsos. "Webseer: An image search engine for the world wide web.," Technical Report TR-96-14, University of Chicago, July 1996.
- [FORSYTH97] David Forsyth, Jitendra malik, and Robert Wilensky, "searching for Digital Pictures," *Scientific American*, pp. 12-20, June 1997.
- [FREEMAN91] W. T. Freeman and E. H. Adelson, "The Design and Use of Steerable Filters," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 13(9), pp. 891-906, 1991.
- [GROSKY92] W. I. Grosky, and Z. Jiang, "A hierarchical approach to feature indexing," *Proc. SPIE: Storage and Retrieval Systems 1662*. pp. 9-20, February 1992.
- [GOLDSTEIN99] E. B. Goldstein, *Sensation & Perception*, 5th ed, Pacific Grove: Brooks/Cole Publishing Company, 1999.
- [GONZALEZ92] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, 1992.
- [GONG96] Y. Gong, H. C. Chua, and X. Guo, "Image Indexing and Retrieval Based on Color Histograms," *Multimedia Tools and Application*. 2, pp. 133-156, 1996.
- [GUDIVADA95] V. Gudivada, V. Raghavan, "Content-Based Image Retrieval Systems," *IEEE Computers*, 28(9), pp. 18-22, 1995.
- [GUIBAS95] L. J. Guibas, B. Rogoff, and C. Tomasi, "Fixed-window image descriptors for image retrieval," *Proc. SPIE: Storage Retrieval Image video*

- Databases III* 2420, pp. 352-363, February 1995.
- [HALL67] A. V. Hall, "Methods for demonstrating resemblance in taxonomy and ecology," *Nature*, Vol. 214, pp. 830-831, 1967.
- [HAFNER95] Jane Hafner, H. S. Sawhney, W. Equitz, M. Flickner, and W. Niblack, "Efficient color Histogram Indexing for Quadratic form Distance functions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 17, Number 7, pp. 729-736, July 1995.
- [HIRATA92] K. Hirata and T. Kato, "Query by Visual Example-Content Based Image Retrieval," *Advances in Database Technology(EDBT '92)*, pp. 56-71, 1992.
- [HOFMANN93] G. R. Hofmann, "The modelling of images for communication in multimedia environments and the evolution from the image signal to the image document," *The Visual Computer*, 9, 1993.
- [HOU92] T. Y. Hou, A. Hsu, and M. Y. Chiu, "A content based indexing technique using relative geometry features," *Image Storage Retrieval Systems* 1662, pp. 59-68, February 1992.
- [HSU95] W. Hsu, T. S. Chua, and H. K. Pung, "An Integrated Color-Spatial Approach to Content-based Image Retrieval," *Proceedings of ACM Multimedia*, pp. 305-313, Sanfrancisco, November 1995.
- [HUANG97] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, "Image Indexing Using Color Correlograms," In: *IEEE Computer Vision and Pattern Recognition Conference*, San Juan, Puerto Rico, June 1997.
- [HUANG98] J. Huang, R. Zabih, "Combining Color and Spatial Information for Content-based Image Retrieval," In: *European Conference on Digital Libraries*, September 1998.
- [IDRIS97] F. Idris and S. Panchannathan, "Review of Image and Video Indexing Techniques," *Journal of visual communication and Image Representation*, VOL. 8, NO. 2, pp. 146-166, June 1997.
- [JACOBS95] C. E. Jacobs, A. Finkelstein, and D. H. Salesin, "Fast Multiresolution Image Query," *Proc. ACM SIGGRAPH*, New York, 1995.
- [JAIN89] A. K. Jain, *Fundamentals of Digital Image Processing*, London: Prentice-Hall, 1989,
- [JAIN95] Ramesh Jain, Rangachar Kasturi, and Brian G. Schunck, *Machine vision*, McGraw-Hill, Inc., 1995.
- [JAIN96] R. Jain, A. Gupta, "Computer Vision and Visual Information Retrieval," *Festschrift for Prof. Azriel Rosenfeld*, IEEE Computer Society, Available at

<http://vision.ucsd.edu/papers/rosenfeld/>, 1996

- [JULESZ73] B. Julesz, H. L. Frisch, E. N. Gilbert, and L. A. Shepp, "Inability of Humans to Discriminate Between Visual Texture That Agree in Second-Order Statistics Revisited," *Perception*. 2., pp. 391-405, 1973.
- [KATO92] T. Kato, T. Kurita, N. Otsu, and K. Hirata, "A Sketch Retrieval Method for Full Color Image Database," *Proceeding of 11th ICPR*, IEEE, pp. 530-533, 1992.
- [KUEHUI97] R. G. Kuehni, *Color: An Introduction to Practice and Principles*, New York: John Wiley & Sons, 1997.
- [LAI00] Ting-Sheng Lai, CHROMA: A Photographic Image Retrieval System, Ph.D. University of Sunderland, Sunderland, United Kingdom, January 2000.
- [LITTLE93] T. D. C. Little, G. Ahanger, R. J. Folz, J. F. Gibbon, F. W. Reeve, D. H. Schelleng, and J. F. Venkatesh, "A digital on-demand video service supporting content-based queries," In *Proc. ACM Intern. Conf. Multimedia*, June 1993.
- [LIU96] F. Liu and R. W. Picard, "Periodicity, Directionality, and Randomness: Wold Features for Image Modeling and Retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(7), pp. 722-733, 1996.
- [LIU99] F. Liu and R. W. Picard, "A Spectral 2-D Wold Decomposition Algorithm for Homogeneous Random Fields," In: *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, Phoenix, Arizona, 15-19 March 1999.
- [MA97a] W.-Y. Ma, *Netra: A Toolbox for Navigating Large Image Databases*. Ph.D. Dissertation, University of California at Santa Barbara, June 1997.
- [MA97b] W. Y. Ma, and B. S. Manjunath, "A Texture Thesaurus for Browsing Large Aerial Photographs," *Journal of the American Society for Information Science*, 1997.
- [MA97c] W. Y. Ma and B. S. Manjunath, "Edge flow: a framework of boundary detection and image segmentation," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 744-49, San Juan, Puerto Rico, June 1997.
- [MANDAL96] M. K. Mandal, T. Aboulnasr, and S. Panchanathan, "Image Indexing Using Moments and Wavelets," *IEEE Transactions on Consumer Electronics*, vol. 42, no. 3, pp. 557-565, August 1996.
- [MANJUNATH95] B. S. Manjunath, "Image browsing in the Alexandria Digital Library (ADL) Project," *D-Lib Magazine*, August 1995.

- [MEHROTRA95] R. Mehrotra, and J. E. Gary, "Similar-Shape Retrieval in Shape Data Management," *IEEE Computer*, 28(9), pp. 57-62, September 1995.
- [MEHTRE95] B.M. Mehtre, M.S. Kankanhalli, A. D. Narsimharu, and G. C. Man, "Color matching for image retrieval," *Pattern Recognition*, vol.16, pp. 325-331, 1995.
- [MEHTRE98] B.M. Mehtre, M.S. Kankanhalli, and W.-F. Lee, "Content-Based Image Retrieval Using A Composite Color-Shape Approach," *Information Processing & Management*, Volume34, Number 1, pp. 109-120, January 1998.
- [MENG96] J. Meng and S.-F. Chang, " Tools for compressed-domain video indexing and editing," In *IS&T/SPIE Symposium on Electronic Imaging: Science and Technology-Storage & Retrieval for Image and Video Databases IV*, San Jose, CA, February 1996.
- [MENG96b] J. Meng and S.-F. Chang, "CVEPS - a compressed video editing and parsing system," In *Proc. ACM Intern. Conf. Multimedia*, pp. 43-53, Boston, MA, November 1996.
- [MILANESE99] Ruggero Milanese and Michael Cherbuliez, "A Rotation, Translation, and Scale-Invariant Approach to Content-Based Image Retrieval," *Journal of visual communication and Image Representation 10*, pp. 186-196, 1999.
- [MINKA96] T. P. Minka, "An Image Database Browser that Learns from User Interaction," Master of Engineering Thesis, 1996.
- [MUNSELL69] A. H. Munsell, *A Grammar of Color: A Basic Treatise on the Color System*, (Albert H. Munsell, 1858-1918). Edited and with an introduction by Faber Birren, New York: Van Nostrand Reinhold Company, 1969.
- [NIBLACK93] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos and G. Taubin, "The QBIC Project: Querying Image by Content using Color, Texture, and Shape," *Proc. SPIE Storage and Retrieval for Image and Video Databases*, pp. 173-187, February 1993.
- [OGLE95] Virginia Ogle and Michael Stonebraker, "Chabot: Retrieval from a relational database of images," *IEEE computer*, 28(9): pp. 40-48, September 1995.
- [OOI97] B.C. Ooi, K.-L. Tan, T.S. Chua and W. Hsu, "Fast Image Retrieval Using Color-Spatial Information," *The VLDB Journal*, Number 7, pp. 115-128, 1997.

- [ORTEGA97] Michael Ortega, Yong Rui, Kaushik Chakrabarti, Sharad Mehrotra, and Thomas S. Huang, "Supporting Similarity Queries in MARS," *ACM Multimedia 97*, November 1997.
- [PASS96a] Greg Pass and Ramin Zabih, "Histogram Refinement for Content-based Image Retrieval," *Workshop on the Application of Computer Vision*, 1996.
- [PASS96b] Greg Pass, Ramin Zabih, and Justin Miller, "Comparing Images Using Color Coherence Vectors," *ACM Conference on Multimedia*. Boston, Massachusetts, November 1996.
- [PASS98] Greg Pass and Ramin Zabih, "Comparing Images Using Joint Histograms," *Computer Science Department*, Connell university, 1998.
- [PENTLAND94] A. Pentland, R. Picard, and S. Sclaroff, "Photobook: Content-Based Manipulation of Image Databases," *Storage and Retrieval for Image and Video Databases II*, SPIE Proceedings Vol. 2185, pp. 34-47, 1994.
- [PICARD94] R. W. Picard and M. Gorkani, "Finding Perceptually Dominant Orientation in Natural Textures," *Spatial Vision, Special Issue in Honor of Bela Julesz*, 8(2), pp. 221-253, 1994.
- [PICARD95] R. W. Picard and T. P. Minka, "Vision Texture for Annotation," *ACM/Springer Journal of Multimedia Systems*, 1995.
- [PITAS95] Ioannis Pitas, *Digital Image Processing Algorithms, Information Retrieval*. Prentice Hall Europe, 1995.
- [RIJSBERGEN79] C. J. van Rijsbergen, *Information Retrieval*. 2nd ed. London: Butterworths, 1979.
- [ROWE93] L. A. Rowe and R. R. Larson, "A video-on-demand system," Technical report, UC Berkeley technical report, 1993.
- [RUI97] Y. Rui, T. S. Huang, S. Mehrotra, and M. Ortega, "Automatic Matching Tool Selection via Relevance Feedback in MARS," In: *Proceedings of the 2nd International Conference on Visual Information System*, San Diego, California, 15-17 December 1997.
- [SCASELLATI94] B. Scasellati, S. Alexopoulos, and M. Flickener, "Retrieval images by 2-D shape: A comparison of computation methods with human perceptual judgements," *Proc. SPIE: Storage Retrieval Image video Databases III 2185*, pp. 2-14, February 1995.
- [SEABORN98] Matthew Seaborn, "On the use of colour in content based image retrieval," submitted for 1st year report of PhD, Brunel University, 15 October 1998.

- [SEKULER94] R. Sekuler and R. Blake, *Perception*, Third edition, New York: McGraw-Hill, 1994.
- [SMITH96a] J. R. Smith and S. F. Chang, "Automated Binary Texture Feature Sets for Image Retrieval," *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Proceeding (ICASSP'96)*, Atlanta, USA, May 1996.
- [SMITH96b] J. R. Smith and S.-F. Chang, "Tools and techniques for color image retrieval," In Ishwar K. Sethi and Ramesh C. Jain, editors, *Storage & Retrieval for Image and Video Databases IV*, volume 2670 of *IS&T/SPIE Proceedings*, pp. 426-437, San Jose, CA, USA, March 1996.
- [SMITH96c] J. R. Smith and S. F. Chang, "VisualSEEk: a Fully Automated Content-Based Image Query System," In: *Proceedings of ACM Multimedia'96 Conference*, Boston, USA, November 1996.
- [SMITH96d] J. R. Smith and S.-F. Chang, "Webseek: a content-based image and video search engine for the world-wide web," *IEEE Multimedia*, August 1996.
- [SMITH97] J. R. Smith and S. F. Chang, "Querying by Color Regions Using the VisualSEEk Content-Based Visual Query System," Maybury M.T. (ed.), *Intelligent Multimedia Information Retrieval*, California: AAAI Press, pp. 23-42, 1997.
- [SMITH97b] J. R. Smith, "Integrated Spatial and Feature Image Systems: Retrieval, Analysis and Compression," Ph. D. thesis, Columbia University, New York, NY, February 1997.
- [SMITH97c] J. R. Smith and S.-F. Chang, "An image and video search engine for the World-Wide Web," *Symposium on Electronic Imaging: Science and Technology - Storage & Retrieval for Image and Video Databases V*, San Jose, CA, February 1997.
- [SMITH98] J. R. Smith, "Image Retrieval Evaluation," In: *Proceedings of IEEE Workshop on Content-Based Access of Image and Video Libraries*, pp. 112-113, Santa Barbara, California, 21 June 1998.
- [STRICKER94] M. A. Stricker and M. J. Swain, "The Capacity of Color Histogram Indexing," In: *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 704-708, 1994.
- [STRICKER94b] M. A. Stricker, "Bounds for the discrimination power of color indexing techniques," *Storage and Retrieval for Image and Video Databases II*, SPIE Proceedings Series Vol. 2185, pp. 15-24, 1994.

- [STRICKER95] M. A. Stricker and M. Orengo, "Similarity of color images," *Storage and Retrieval for Image and Video Databases III*, volume 2420 of SPIE Proceedings Series, pp. 381-392, February 1995.
- [STRICKER96] M. A. Stricker and A. Dimai, "Color indexing with weak spatial constraints," *In Storage and Retrieval for Still Image and Video Databases*, volume 2670, pp. 29-40, 1996.
- [SWAIN91] M. J. Swain and D. H. Ballard, "Color Indexing," *International Journal of Computer Vision*, 7(1), pp. 11-31, 1991.
- [TAO99] Yi Tao and William I. Grosky, "Object-Based Image Retrieval Using Point Feature Maps," *Short Paper Proceedings of the 8th IFIP 2.6 Working Conference on Database Semantics (DS8)*, Rotorua, New Zealand, pp. 59-73, 5-8 January , 1999.
- [TAO99b] Yi Tao and William I. Grosky, "Spatial Color Indexing: A Novel Approach for Content-Based Image Retrieval," *accepted in IEEE 6th International Conference on Multimedia Computing and Systems (ICMCS'99)*, Florence, Italy, pp. 530-535, 7-11 June , 1999.
- [TAMURA78] H. Tamura, S. Mori, and T. Yamawaki, "Texture Feature Corresponding to Visual Perception," *IEEE Transaction on Systems, Man, and Cybernetics*. 8(6), pp. 460-472, 1978.
- [THEODORIDOS99] Sergios Theodoridos and Kostantinos Koutroumbas, *Pattern Recognition*, Athens: Academic Press, 1999.
- [THORNTON98] A L Thornton Beng AMIEE, "Colour object recognition using a complex colour representation and the frequency domain," Ph. D Thesis, The University of Reading, 1998.
- [VAIL96] A. Vailaya, Y. Zhong, and A. K. Jain, "A Hierarchical System for Efficient Image Retrieval," *Proceeding of 13th ICPR*, pp. 356-359, IEEE, 1996.
- [VERTAN00] C. Vertan and N. Boujemaa, "Spatially constrained color distributions for image indexing," *International Conference on Color in Graphics and Image Processing-CGIP'2000*, INRIA Rocquencourt, France, 2000.
- [VERTAN00a] C. Vertan and N. Boujemaa, "Color texture classification by normalized color space representation," *Processing of ICPR'2000*, Barcelona, Spain, 3-8 September 2000.
- [VERTAN00b] C. Vertan and N. Boujemaa, "Using fuzzy histograms and distances for color image retrieval," *Processing of CIR'2000*, Brighton, United Kingdom, 4-5 May 2000.

- [VERTAN00c] Constantium VERTAN and V. Buzuloiu, "Fuzzy nonlinear filtering of color images: A survey," In E. Kerre and M. Nachtegaele, editors, *fuzzy Techniques in Image Processing*, Heidelberg, Germany, 2000.
- [WATT98a] A. Watt and F. Policarpo, *Computer Image*. Harlow: Addison Wesley Longman, 1998.
- [WOELK86] D. Woelk, W. Kim, and W. Luther, "An object-oriented approach to multimedia databases," *ACM Databases*, 1986.
- [YOUNG1902] T. Young, "An account of some cases of the production of colours, not hitherto described." *Philosophical Transactions* 1902.
- [ZHANG94] H.-J. Zhang, S. W. Smoliar, J. H. Wu, C. Y. Low, and A. Kan-kanhalli, "A Video Database System for Digital Libraries," *ISS*, National University of Singapore, June 1994.
- [ZELANSKI89] P. Zelanski and M. P. Fisher, *Colour for Designers and Artists*, London: The Herbert Press, 1989.
- [WYSZECK82] Gunter Wyszecki, and W. S. Stiles, *Color Science: Concepts and Methods, quantitative Data and formulae*, Second Edition, John Wiley & Sons, 1982.
- [장정동2001] 장 정동, 이 태홍, "색상-공간정보를 고려한 내용기반 영상검색," 한국통신학회논문지, vol.26, No.3B, pp. 315-322, 3월, 2001년.
- [장정동2001a] 장 정동, 이 태홍, "색상분할영역에서 거리 히스토그램을 이용한 영상검색," 한국통신학회논문지, vol.26, No.7B, pp. 968-974, 7월, 2001년.

Ph. D. Thesis

Content-Based Image Retrieval of Color Image

Jung-Dong Jang

Department of Electronic Engineering

Graduate School

Yeungnam University

(Advised by Processor Taihong Yi)

Abstract

Consider that everyday we view various forms of visual media, such as photographs, graphics, animations and videos. These media are provided increasingly in digital form. For example, the WorldWideWeb (WWW) is one such source for viewing hundreds of gigabytes of digital visual information. While there is great accessibility to large stores of digital imagery, new systems need to be developed to manage and search better for the visual information. To cope with this problem, there have been two major approaches to study image retrieval from somewhat different angles.

The one is to annotate each image manually with text. This approach has however insuperable problems that creating annotations by hand can become hopeless because databases can be very large. Manual indexing is that the content of an image is a highly subjective concept, that is, different people may perceive it differently.

Another approach is to index images directly based on the visual image content.

Content-based image retrieval was proposed in the early 90's and it has been under intensive research from then on. In general, visual feature extraction is fundamental bases in the content-based image retrieval paradigm. Visual features are usually divided into four classes depending on what kinds of properties they describe. The classes are color, texture, shape, and spatial relationship.

Automatic feature extraction would be desired in large-scale image retrieval system. Besides, the manual annotation of text is likely to be error prone with incompatible or inconsistent text description. In addition, some of the image databases are not confined to specific domains. In these situations, content-based retrieval approaches are preferred.

However, the use of color information has been used widely in content-based retrieval systems. It is clear that color plays a very important role in defining an image. But color alone is insufficient to characterize an image. In order to overcome these problems, we discuss a technique of considering color-spatial information with spatial knowledge to obtain an overall impression of the image.

In this paper, content-based image retrieval considering color spatial information is proposed. Our goal is to improve the above mentioned shortcomings through following process. In the first, Sequential-Clustering schemes, which are quite straightforward and fast method, are introduced to cluster the homogeneous property of any images. Then, clustered color regions are used as the basis for computing distance histogram with relevant color-spatial information. Finally, Color mean of each regions and distance histogram computed from the center of each clustered color regions are used as feature vectors.

Our proposed method needs only 18 features and required storage space is smaller than with conventional methods. However, our experimental results showed not only that retrieval efficiency improved 8.5 percent over conventional methods but also much computation and data storage was saved. In addition, this method showed excellent Precision vs. Recall evaluations and was judged better by human perception.