creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

[UCI]I804:24011-200000236682

2008年 8月
博士學位論文

컬러와 에지 정제에 기반한 잡음 불변 영상 검색

朝 鮮 大 學 校 大 學 院

情報通信工學科

朴 泰 洙

2008年 2月 博士學位論文

컬러와 에지 정제에 기반한 잡음불변 영상검색

朴 泰 洙

컬러와 에지 정제에 기반한 잡음 불변 영상 검색

Noise-invariant Image Retrieval
based on Color and Edge Refinement

2 0 08 年 8月 25日

朝 鮮 大 學 校 大 學 院

情報通信工學科

朴 泰 洙

컬러와 에지 정제에 기반한 잡음 불변 영상 검색

指導敎授 朴 鍾 安

이 論文을 工學博士學位申請 論文으로 提出함.

2 0 08 年 4月 日

朝 鮮 大 學 校 大 學 院

情報通信工學科

朴 泰 洙

朴泰洙의 博士學位 論文을 認准함

委　員 光州科技院 敎授　崔　泰 善　印

委　員 朝鮮大學校 敎授　韓 承 朝　印

委　員 全南大學校 敎授　朴 勝 珍　印

委　員 順天大學校 敎授　文 庸 善　印

委　員 朝鮮大學校 敎授　朴 鍾 安　印

2 0 08 年　6 月　日

朝 鮮 大 學 校　大 學 院

목 차

도 목 차 ····· iii
표 목 차 ····· vi
ABSTRACT ····· ix

Ⅰ. 서 론 ····· 1
A. 국내외 연구 동향 ····· 2
B. 영상 검색의 문제점 분석 ····· 5
C. 연구의 내용 ····· 6
Ⅱ. 내용기반 영상 검색 시스템 ····· 8
A. 내용기반 영상 검색 ····· 8
B. 영상 검색을 위한 내용기반 상호 작용 ····· 10
C. 내용기반 영상 검색을 위한 특성 표현 ····· 11
D. MPEG-7 영상 특징 정보 ····· 14
Ⅲ. 내용기반 영상 검색 기법 ····· 19
A. 컬러 기반 영상 검색 ····· 20
1. 컬러 히스토그램 인터섹션 ····· 20
2. 색상 모멘트를 이용한 기법 ····· 22
3. 그레이 영상 처리 방법 ····· 23
4. 벡터 정합법 ····· 23
5. 다차원 구배도 방법 ····· 26
6. 벡터적 방법 ····· 28

B. 에지 기반 영상 검색 ····· 30
1. 에지의 방향 히스토그램을 이용한 기법 ····· 31
2. 저나이크 모멘트를 이용한 기법 ····· 32
3. 지역적인 미분 불변치를 이용한 기법 ····· 33
Ⅳ. 컬러와 에지 정제 기반 CBIR 설계 ····· 39
A. 기존 컬러/에지 알고리즘 방법 ····· 39
B. 컬러와 에지 정제 기반 영상 검색 알고리즘 ····· 41
1. 컬러 특성 벡터 ····· 42
2. Post-processing ····· 43
3. 에지 특성 벡터 ····· 44
4. 8 빈 히스토그램 특성 벡터 ····· 45
5. 컬러 코렐로그램 ····· 48
Ⅴ. 시뮬레이션 및 결과 ····· 50
A. 시뮬레이션 환경 및 평가 ····· 50
1. 환경 ····· 50
2. 영상 검색 성능 척도 평가 ····· 50
B. 시뮬레이션 결과 및 분석 ····· 52
1. 컬러 정제에 따른 시뮬레이션 결과 및 분석 ····· 52
2. 에지 정제에 따른 시뮬레이션 결과 및 분석 ····· 81
3. 컬러 및 에지정제방법에 따른 시뮬레이션 결과 및 성능 분석 ····· 95
Ⅵ. 결론 ····· 98
참고문헌 ····· 99

도 목 차

그림 2-1 내용기반 영상 검색 시스템 구성도 ····· 8
그림 2-2 컬러 히스토그램 ····· 12
그림 2-3 MPEG-7 표준화의 범위 ····· 15
그림 2-4 MPEG-7의 표준화 요소와 응용 시스템 기능 구성 ····· 17
그림 3-1 퓨전 방법의 처리과정 ····· 24
그림 3-2 채널별 구배도를 위한 마스크들 ····· 25
그림 3-3 다차원 구배도 방법의 처리과정 ····· 27
그림 4-1 기존 컬러 정제 기반 영상 검색 알고리즘 ····· 40
그림 4-2 기존 에지 정제 기반 영상 검색 알고리즘 ····· 40
그림 4-3 제안한 컬러 및 에지 정제 기반 영상 검색 알고리즘 ····· 42
그림 4-4 컬러 빈도수로 구성된 컬러 히스토그램 ····· 47
그림 4-5 8빈도수로 구성된 컬러 히스토그램 ····· 47
그림 4-6 오토 코렐로그램 획득 방법 ····· 49
그림 4-7 컬러 크로스 코렐로그램 획득 방법 ····· 49
그림 5-1 재현율 및 정확도의 구성 방법 ····· 50
그림 5-2 정제된 컬러영상 IMG0001 ····· 53
그림 5-3 정제된 컬러영상 IMG0002 ····· 54
그림 5-4 정제된 컬러영상 IMG0003 ····· 55
그림 5-5 정제된 컬러영상 IMG0004 ····· 56
그림 5-6 정제된 컬러영상 IMG0005 ····· 57
그림 5-7 정제된 컬러영상 IMG0006 ····· 58

그림 5-8 정제된 컬러영상 IMG0007 ····· 59
그림 5-9 정제된 컬러영상 IMG0010 ····· 60
그림 5-10 정제된 컬러영상 IMG0011 ····· 61
그림 5-11 정제된 컬러영상 IMG0012 ····· 62
그림 5-12 정제된 컬러영상 IMG0013 ····· 63
그림 5-13 정제된 컬러영상 IMG0014 ····· 64
그림 5-14 정제된 컬러영상 IMG0015 ····· 65
그림 5-15 정제된 컬러영상 IMG0025 ····· 66
그림 5-16 정제된 컬러영상 IMG0026 ····· 67
그림 5-17 정제된 컬러영상 IMG0027 ····· 68
그림 5-18 정제된 컬러영상 IMG0028 ····· 69
그림 5-19 정제된 컬러영상 IMG0032 ····· 70
그림 5-20 정제된 컬러영상 IMG0038 ····· 71
그림 5-21 정제된 컬러영상 IMG0039 ····· 72
그림 5-22 정제된 컬러영상 IMG0040 ····· 73
그림 5-23 정제된 컬러영상 IMG0041 ····· 74
그림 5-24 정제된 컬러영상 IMG0043 ····· 75
그림 5-25 정제된 컬러영상 IMG0045 ····· 76
그림 5-26 정제된 컬러영상 IMG0046 ····· 77
그림 5-27 정제된 컬러영상 IMG0047 ····· 78
그림 5-28 정제된 컬러영상 IMG0048 ····· 79
그림 5-29 정제된 컬러영상 IMG0049 ····· 80
그림 5-30 정제된 에지영상 IMG0001 ····· 81
그림 5-31 정제된 에지영상 IMG0002 ····· 82

그림 5-32 정제된 에지영상 IMG0003 ····· 83
그림 5-33 정제된 에지영상 IMG0005 ····· 84
그림 5-34 정제된 에지영상 IMG0011 ····· 85
그림 5-35 정제된 에지영상 IMG0025 ····· 86
그림 5-36 정제된 에지영상 IMG0026 ····· 87
그림 5-37 정제된 에지영상 IMG0027 ····· 88
그림 5-38 정제된 에지영상 IMG0028 ····· 89
그림 5-39 정제된 에지영상 IMG0029 ····· 90
그림 5-40 정제된 에지영상 IMG0038 ····· 91
그림 5-41. 정제된 에지영상 IMG0039 ····· 92
그림 5-42 정제된 에지영상 IMG0040 ····· 93
그림 5-43 재현율 & 정확도 성능 분석 이미지 1 ····· 96
그림 5-44 재현율 & 정확도 성능 분석 이미지 2 ····· 97

표 목 차

표 2-1 내용기반 영상 검색 시스템의 비교 ····· 9
표 2-2 오디오 비주얼 데이터의 대표적인 특징 및 기술자의 예 ····· 16
표 5-1 재현율 및 정확도의 구성의 기호 ····· 51
표 5-2 정제된 컬러 영상 IMG0001 기술자 값 ····· 52
표 5-3 정제된 컬러 영상 IMG0002 기술자 값 ····· 53
표 5-4 정제된 컬러 영상 IMG0003 기술자 값 ····· 54
표 5-5 정제된 컬러 영상 IMG0004 기술자 값 ····· 55
표 5-6 정제된 컬러 영상 IMG0005 기술자 값 ····· 56
표 5-7 정제된 컬러 영상 IMG0006 기술자 값 ····· 57
표 5-8 정제된 컬러 영상 IMG0007 기술자 값 ····· 58
표 5-9 정제된 컬러 영상 IMG0010 기술자 값 ····· 59
표 5-10 정제된 컬러 영상 IMG0011 기술자 값 ····· 60
표 5-11 정제된 컬러 영상 IMG0012 기술자 값 ····· 61
표 5-12 정제된 컬러 영상 IMG0013 기술자 값 ····· 62
표 5-13 정제된 컬러 영상 IMG0014 기술자 값 ····· 63
표 5-14 정제된 컬러 영상 IMG0015 기술자 값 ····· 64
표 5-15 정제된 컬러 영상 IMG0025 기술자 값 ····· 65
표 5-16 IMG0025가 180° 회전(IMG0026) 정제된 컬러 영상 기술자 값 66
표 5-17 IMG0025가 좌우반전(IMG0027) 정제된 컬러 영상 기술자 값 67
표 5-18 IMG0025가 좌우늘임(IMG0028) 정제된 컬러 영상 기술자 값 68
표 5-19 IMG0025가 상하반전(IMG0032) 정제된 컬러 영상 기술자 값 69
표 5-20 정제된 컬러 영상 IMG0038 기술자 값 ····· 70

표 5-21 IMG0038가 좌우반전(IMG0039) 정제된 컬러 영상 기술자 값 71
표 5-22 IMG0038가 상하반전(IMG0040) 정제된 컬러 영상 기술자 값 72
표 5-23 IMG0038가 90° 회전(IMG0041) 정제된 컬러 영상 기술자 값 73
표 5-24 IMG0038가 30° 비틀림(IMG0043) 정제된 컬러 영상 기술자 값74
표 5-25 IMG0038가 120° (IMG0045) 정제된 컬러 영상 기술자 값 ····· 75
표 5-26 IMG0038가 150° 비틀림(IMG0046) 정제된 컬러 영상 기술자 값76
표 5-27 IMG0038가 60° 비틀림(IMG0047) 정제된 컬러 영상 기술자 값77
표 5-28 IMG0038가 45° 비틀림(IMG0048) 정제된 컬러 영상 기술자 값78
표 5-29 IMG0038가 좌우늘임(IMG0049) 정제된 컬러 영상 기술자 값 79
표 5-30 정제된 에지 영상 IMG0001 기술자 값 ····· 81
표 5-31 정제된 에지 영상 IMG0002 기술자 값 ····· 82
표 5-32 정제된 에지 영상 IMG0003 기술자 값 ····· 83
표 5-33 정제된 에지 영상 IMG0005 기술자 값 ····· 84
표 5-34 정제된 에지 영상 IMG0011 기술자 값 ····· 85
표 5-35 정제된 에지 영상 IMG0025 기술자 값 ····· 86
표 5-36 정제된 에지 영상 IMG0026 기술자 값 ····· 87
표 5-37 정제된 에지 영상 IMG0027 기술자 값 ····· 88
표 5-38 정제된 에지 영상 IMG0028 기술자 값 ····· 89
표 5-39 정제된 에지 영상 IMG0029 기술자 값 ····· 90
표 5-40 정제된 에지 영상 IMG0038 기술자 값 ····· 91
표 5-41 IMG0038이 좌우반전(IMG0039) 정제된 에지 영상 기술자 값 92
표 5-42 IMG0038이 상하반전(IMG0040) 정제된 에지 영상 기술자 값 93
표 5-43 컬러 및 에지 정제 방법에 따른 시뮬레이션 및 성능 분석 ····· 95
표 5-44 기존 알고리즘과 제안한 알고리즘의 비교결과 ····· 96

ABSTRACT

Noise-invariant Image Retrieval based on Color and Edge Refinement

Park, Tae-Soo
Advisor : Prof. Park, JongAn, Ph.D.
Department of Information & Communications,
Graduate School of Chosun University

Standard histograms are very efficient image features and they are widely used for content based image retrieval. They are insensitive to small changes in the image appearances. Due to the fact that they are not very unique and robust, many images of different appearances can have similar histograms. Hence, the main disadvantage of histograms is that they provide coarse characterization of an image. Color histograms, too, are widely used and suffer from the same problem. Color histograms are employed in systems as QBIC [1], Chabot [2] etc.

In order to avoid these shortcomings of histograms, a modified scheme based on histogram refinement method [3] is presented. The histogram refinement method says that the pixels within a given bucket is to be split into classes based upon some local property, and these split histograms are then compared on bucket by bucket basis, just like normal histogram matching. The pixels within a bucket with same local property are compared. Therefore, the results are better than the normal histogram matching. Hence, not only the color features of the image are used but also spatial information is incorporated to refine the histogram. The results are obtained by testing the algorithm on a database of images provided in

MPEG-7 data.

The technique defined in this paper is based on histogram refinement, and is called as 'color and edge refinement'. The proposed method splits the pixels in a given bucket into several classes similar to histogram refinement method. These classes are quantized and compared based upon their color and edge coherence vectors. Color refinement is based on histogram refinement method suggested by Pass and Zabih [3]. However, edge refinement involves a different mechanism.

Pass and Zabih state that color histogram buckets are partitioned based on spatial coherence. A pixel is coherent if it is a part of some sizable similar colored region, otherwise it is incoherent. Therefore, in color refinement, the pixels are classified as coherent or incoherent within each color bucket. If a pixel is part of a large group of pixels of the same color which form at least one percent of the image then that pixel is a coherent pixel and that group is called the coherent group or coherent cluster. Otherwise it is incoherent pixel and the group is incoherent group or incoherent cluster. Two more properties are calculated next for each bin. Firstly, the numbers of clusters are found for each case, i.e., coherent and incoherent case in each of the bin. Secondly, the average size of clusters is computed. Hence, there are six values calculated for each bin: one each for percentage of coherent pixels and incoherent pixels, number of coherent clusters and incoherent clusters, average of coherent cluster and incoherent cluster.

Edge refinement is also based on histogram refinement method. The buckets are formed based on edge direction in the pre-processing stage, and then, total number of on pixels is computed for each bucket. Sobel operator is used to compute the horizontal and vertical edges. Hence, two buckets are formed, which are further classified as coherent or incoherent. Four more values are calculated in each bucket. The numbers of clusters

are found in each bucket by 8-neighborhood rule, and then the average of the cluster is computed in each bucket. Moreover, the number of straight edges and slanted edges is computed. Therefore, for each bin, there are five values: one each for total number of on pixels, number of clusters, average size of cluster, number of straight edges and number of slanted edges.

Hence, we can conclude that color refinement method takes care of the color as well as the spatial relation feature. Moreover, edge refinement method implementation to the results of color refinement method presents significant improvement and provides more accurate and precise results as proved through rigorous experimentation. Furthermore, experiments have proved that the proposed algorithm provides better results than the equivalent methods.

Ⅰ. 서 론

최근 디지털 정보화 사회에서 가장 크게 변하고 있는 것 중의 하나는 방대해진 양의 디지털 정보들이 끊임없이 증가하는 현상이다. 그 정보의 양은 커지고 컬러, 질감, 형태 또한 다양화되어 과거 문서 위주의 정보들로부터 영상, 비디오, 오디오, 음성 등으로 바뀌어 생성되고 있다. 이렇게 커진 데이터를 저장하는 방법과 사용자가 원하는 정보를 쉽고 정확하게 얻을 수 있는 방법이 연구되고 있다. 또한, 디지털 형태로 변환되고 저장이 되는 영상 자료들에 대한 표현이나 검색 방식은 과거의 방법과는 다르다. 텍스트에 기반 하여 영상을 표현하고 검색하던 과거의 영상 검색 방식은 텍스트 기반의 문제점으로 텍스트 주석이 각 영상에 추가되고 텍스트를 검색 질의어로 사용하기 때문에 사용자는 찾고자 하는 영상의 키워드 범위를 알고 있어야 한다. 그리고 나라마다 언어와 환경에서 나오는 문화적인 특성에 의해 키워드의 의미 및 용어가 다르게 됨으로써 텍스트 기반의 한계를 인식하고 있다. 그리하여, 현재 영상 데이터의 표현과 검색을 위한 주된 접근 방식중의 하나는 영상에 대한 컬러, 질감, 형태 등의 특징에 의해 검색하는 방식이다.

컬러 특징과 공간 정보를 이용한 영상 검색은 모든 영상을 컬러의 특징과 공간 정보를 이용해 데이터베이스로 구현하고 정확성을 측정하여 동일 영상과 유사 영상들을 검색하는 데에 있다. 내용기반 영상 검색과 함께 시각적인 정보인 컬러가 많이 연구되고 있으며, 여러 가지 광범위한 응용에서 중요한 역할로 활용되고 있다 [1][2].

많은 데이터베이스로부터 영상을 검색하기 위해서 영상의 전체적인 특징들을 인식하여 검색을 하면 시간적인 문제에 부딪혀서 어렵기 때문에 영상의 컬러와 공간의 특징들을 추출하여 사용하였다. 컬러는 물체의 표면을 반영할 뿐 아니라 그 표면의 방향, 카메라의 기하학 관점과 조명의 변화에 영향을 받는다. 하지만 컬러 분포, 공간의 거리와 각에 의한 상호 관계가 고려되면 그 환경의 요소의 전체적 변화로 픽셀은 거의 같은 비율로 서로 관련 있는 컬러를 변화시킨다. 만약 한 픽셀이 최대값을 가지고 있다면, 영상 안에 컬러 구성 요소는 환경의 변화 후에도 픽셀은 그 가치의 변화만으로 반드시 다시 최대값인 컬러 구성 요소를 가지게 된다. 이 같은 부분은 새로운 알고리즘에 적용되어 제안된다.

또한 에지에 기반한 영상 검색은 에지에 의한 특징벡터를 추출하여 에지맵을 구성

함으로써 감소된 특징벡터로 검색시간을 감소 시켜 효율성을 높이고 있다.
이와 같은 에지기반 영상 검색은 에지정제기법을 사용함으로써 정확도를 높이고 있다[3].

본 논문에서는 컬러와 에지 특징추출의 성능을 개선하여 이들 컬러정제와 에지정제기법을 융합하여 영상 검색 시뮬레이션을 통하여 그타당성을 입증하고자 한다. 이을 위하여 제1장에서는 국내외 연구 동향을 살펴보고 영상 검색의 문제점을 분석한다. 제2장과 제3장에서는 내용기반 영상 검색 시스템과 기법을 분석한다 그리고 제4장에서는 컬러와 에지 정제에 기반한 영상 검색 알고리즘을 설계하고 제5장에서는 시뮬레이션을 통하여 그 타당성을 확인하며 제6장에서는 결론을 맺는다.

A. 국내외 연구 동향

국내에서는 컬러와 공간정보를 이용한 연구가 많이 이루어지고 있다. 이는 영상 검색에서 정보 서비스의 중요한 부분으로 검색 시스템은 물론이거니와 의학, 지리정보, 광고, 예술, 패션, 멀티미디어에 관련된 모든 것에 응용이 되고 있다. 컬러의 특징과 공간 정보를 이용한 영상 검색은 추출된 최적의 특징값과 유사 영상 중에 유사성이 최대가 되는 특징값을 정합을 시킨다. 최근까지도 컬러 히스토그램, 컬러 코렐로그램, CCV 등이 내용기반 영상 검색으로 연구가 활발히 진행되고 있다.

국외에서도 영상 검색에 관한 연구 활동은 과거나 현재에도 이루어지고 있다. Swain과 Ballard가 제안한 컬러 히스토그램 방법은 각각의 객체에 대해 인덱스를 생성한 후 데이터베이스 내의 객체와 입력된 객체와의 히스토그램을 비교하여 공유되는 최소값을 구하는 방식이다. 이 방법은 모델 기반의 검색 방식이므로 영상에 제한을 가하지 않는 내용기반 시스템에서 적용하는데 있어서 문제점을 안고 있다[4].

Nagasaka와 Tanaka가 제안한 컬러 쌍의 경우 영상내의 경계선 사이에서 변화되는 컬러성분의 차로 히스토그램을 형성하는 방식을 사용하였다. 이들 연구에서는 기존의 히스토그램 방식이 전역적인 특징을 갖는다는 단점을 보완하는 데는 효율적이나 정합된 화상에 일괄적인 비중을 주는 오류를 범하고 있다[5].

Grosky 및 Mehrotra 등은 영상에 포함하고 있는 객체들의 윤곽선을 벡터화하여

이를 영상 데이터의 특징 값으로 사용하였다. 이 특징 값들은 거리, 방향, 인접요소와 이루고 있는 각도 등을 계산하여 영상간의 유사성을 비교하였다. 검색에 소요되는 시간을 단축하기 위해 특징 값들을 트리형태로 색인한다. 이 방법은 영상내의 객체의 크기와 위치에 관계없이 적용할 수 있다는 장점은 있으나, 객체의 형태가 제한된 환경에서만 사용할 수 있다는 단점이 있다[6].

Jacobs 등은 특성벡터의 추출과정에서 처음으로 웨이블릿 변환을 사용하였다. 그들의 연구에서는 색상영상에 대하여 직접적으로 웨이블릿 변환을 적용하여 얻은 계수들을 영상의 R, G, B 각 채널마다 따라 특성벡터를 구성하였다. 이는 특성벡터의 차원이 너무 크므로 처리시간이 많이 걸려 기존의 색인기술을 이용하는데 많은 어려움이 있다.

IBM사의 QBIC 프로젝트에서는 텍스트와 더불어 스케치, 사용자가 구성한 질의 영상, 컬러 레이아웃이나 구조적 묘자, 실험 영상 아이콘이나 그래픽 정보를 기반으로 한 대규모 데이터베이스 질의가 가능한 검색방법을 개발하였다. 이러한 연구들은 대부분 특정한 응용에만 적합한 상태이다.[1]

버클리 대학에서 개발한 Cabot VisualSEEK에서는 주석정보와 영상의 특성벡터를 함께 고려한 시스템으로 특성벡터로는 색상과 히스토그램을 이용하였다. 영상의 종류 키워드, 키워드와 더불어 스케치, 텍스쳐, 모양을 사용한 방대한 데이터베이스 상에서 검색할 수 있도록 연구되었다[8].

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 사각형꼴로 표현한 후, 영역의 대표색상과 공간정보를 분석하여, 통합된 색상-공간정보에 기반 한 검색방법을 제안하였다[3].

Stricker 등은 영상을 중앙의 타원형과 겹치는 영역을 포함하여 5영역으로 구분하고, 각 영역에서 히스토그램 및 1, 2, 3차 모멘트를 추출하여 비교하였다[9]. 이들은 HSV공간에서 히스토그램과 평균색상, 분산, 왜곡을 사용하여 계산량을 줄이는 효과를 가져온 반면, 영상에 나타나는 색상정보를 상세하게 반영하지 못하는 단점이 있다.

Zhang 등은 영상의 내용상 전체적으로 색상이 비슷한 영상을 검색하기 위해 고정된 크기의 국부적인 색상 히스토그램을 사용함으로써 계산이 간단하여 다양한 응용분야에서 효과적임을 보여주고 있다[10].

Mehtre는 클러스터링 방법에 의해 영역을 분할하고 이 영역에서 7가지의 색상 모멘트를 계산하여 복합적인 특성추출방법을 제안하였다[11].

영상을 고정 또는 가변블록으로 분할하지 않고 공간적인 색상분포를 표현하기 위한 다른 접근방법이 있다.

Huang 등도 공간정보를 고려하기 위한 방법으로, 주어진 색상의 화소가 임의의 거리에 존재할 확률을 계산한 테이블을 작성하여 색인하는 Color Correlogram을 제안하였다[12].

Mandal 등은 색상정보를 반영하기 위해서 색상 히스토그램에 대한 모멘트를 이용하였고, 공간정보를 고려하는 방법으로는 웨이브릿 부 영상을 가우시안 분포로 근사화한 뒤, 그 계수를 이용하였다[13]. 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 사용함으로서 공간적으로 구속된 색상분포를 나타내고 있으며, 확률적으로 계산하는 히스토그램 표현방식과는 달리 멤버십 함수로 그룹별 대표성을 표현하는 퍼지 히스토그램으로 영상의 내용을 표현하고 있다[14].

Smith & Chang's 방법은 또한 이미지들을 분할했다. 그러나 그들은 색상법이 [8]에 언급했던 한가지의 지배적인 컬러(color) 대신에 각 지역이 각 지역이 여러을 내포할 수 있도록 했다. 그들은 3100개의 이미지 데이터베이스(database)를 이용했다. Rickman과 Stonham[15]은 고정된 변을 가진 작은 정삼각형을 기초로 한 방법을 제공했다. Stricker와 Dimai[9]은 이미지 안에 일어나는 세 개의 순간(three moment) 색 분포(color distribution)을 발견했다.

박인규, 윤일동, 이상욱 등이 제안한 기법은 영상에서 컬러들의 분포를 MCAG (Modified Color Adjacency Graph)와 SVG(Spatial Variance Graph)의 복합 그래프로 표현하여 검색하는 기법이다[39]. MCAG는 인접한 컬러들의 분포를 그래프로 표현하고, SVG는 컬러들의 위치 정보들을 계산하여 Graph Intersection 방법을 이용하여 검색한다. 컬러를 그래프로 표현하는 과정에서 생성되는 노드의 수를 제한된 수로 표현하는 것이 가장 큰 문제점이다.

김홍염, 하영호 등이 제안한 기법은 박인규에서 양자화된 컬러공간에서 서로 다른 빈에 속하는 컬러들의 인접한 확률을 이용하지만, 균일 양자화에 의하여 비슷한 색들이 다른 색으로 잘못 구분될 수 있는 단점을 보완하기 위해서 Androutsos가 제안한 Vector Angular Distance를 이용하여 인접화소들의 실제 색상차를 구하여

문턱치값보다 큰 경우는 컬러의 인접성을 작을 경우는 컬러의 기울기를 추출하여 2차원 배열에 저장한 다음 graph intersection 방법을 이용하여 유사도를 측정하였다[40]. 하지만 이 방법 또한 컬러의 수를 제한된 수로 표현하는 것이 문제점이다.

김태희, 황태현 등이 제안한 엔트로피(entroph)를 이용한 검색 기법은 Chang이 제안한 PIM(Picture Information Measure) 값을 이용한다[41]. 한 영상에서 임의의 고정된 크기의 하위 블록들로부터 밝기값에 대하여 평균과 분산을 계산하여 엔트로피 특징으로 정한다. 만일, 엔트로피 특징값의 분산이 문턱치 분산값보다 큰 경우에는 엔트로피 특징값의 평균을 기준으로 두 영역으로 분리한다. 각 영역으로부터 RGB 컬러 히스토그램의 대역별 평균 컬러 값들을 계산한다. 검색 시에는 두 영상에서 계산되어진 엔트로피와 컬러 특징값들을 비교하며 유사한 영상을 검색한다.

박종안,

이외에도 국내·외에서의 영상 검색 하는 방법의 연구는 진행되어져왔고 최근까지도 계속하여 새로운 알고리즘을 개발 중에 있다.

B. 영상 검색의 문제점 분석

현대 사회가 멀티미디어 정보화 사회로 변화하는 가장 큰 현상 중의 하나는 방대해진 양의 정보들이 계속해서 증가하는 현상이다. 이에 따라 단순한 텍스트 정보만이 아닌 영상, 비디오, 오디오, 음성 데이터 등과 같은 멀티미디어 정보들이 넘쳐나고 있다. 영상에서 컬러는 과거의 컬러와는 비교가 되지 않을 정도로 많아졌다. 컬러의 종류가 많아졌을 뿐 아니라 환경이 변화하면서 그 요소들도 변화해 갔다. 이에 관련하여 변화하는 컬러 안에서 인터넷 및 데이터베이스 안에 존재하는 정보를 찾는 어려움을 극복하고자 많은 검색 엔진들이 끊임없이 개발되고 있으며 다양한 기법을 이용하여 정보를 원하는 사용자에게 많은 도움이 되고 있다.

컬러의 특징은 영상 검색을 위해 가장 널리 사용되는 시각요소이다. 컬러의 3차원 값은 영상의 1차원 값인 그레이 값보다 잠재적으로 더 높은 특징값을 만들어 준다. 적절한 컬러 기술을 선택하기 전에, 컬러 공간은 결정되어야 한다. 이들 컬러 특징자의 RGB 컬러 공간은 영상 출력을 위한 컬러 공간으로 널리 사용된다.

컬러 히스토그램의 컬러 패턴이 데이터를 추출하는데 유일한 비교 방법이라면

컬러 영상 검색시스템의 효과적인 방법이다. 컬러 히스토그램은 관점 축과 크기, 채색과 보는 각도의 완만한 변화와 회전에 강인하다. 영상의 어떤 픽셀들이라도 컬러의 공간에서 구성 요소들에 의해 기술하게 될 수 있는데, 양자화된 빈을 위한 픽셀수의 분포는 각 구성 요소들을 위해 정의된다. 다량의 빈들을 가지는 히스토그램은 그 계산적인 비용을 증가시킬 뿐만 아니라, 영상의 데이터베이스를 위한 효율적인 추출에 부적절하다. 게다가, 컬러 히스토그램은 픽셀의 공간 정보를 고려하지 않는다. 그래서 매우 다양한 영상들이 비슷한 컬러 분배를 가지고 있을 수 있다. 이 같은 히스토그램의 문제는 거대한 규모의 데이터베이스에서 더욱 심해진다.

컬러 코렐로그램은 픽셀의 컬러 분포 뿐 아니라 컬러들 쌍의 공간의 상관관계까지도 포함할 것을 제안했다. 공간의 상관관계에서 첫 번째와 두 번째는 픽셀의 쌍인 컬러이고 세 번째는 특별한 공간 거리이다. 만일, 우리가 컬러 한 쌍의 모든 결합을 고려하면, 컬러 코렐로그램의 크기는 매우 클 것이다. 그리고 컬러 코렐로그램에서 방향과 거리를 이용하는데 큰 거리는 결과적으로 많은 계산과 커다란 기억장치요구를 가져오기 때문에 작은 거리에서 최적의 성능을 보여야 한다. 또한 컬러 코렐로그램은 부분 영상의 크기에 민감한 반응을 보인다. 단지 주위에 같은 컬러가 나올 확률 분포이기 때문에 양자화 된 영상 크기는 같지만 우리가 찾고자 하는 부분 영상의 크기가 같지 않다면 확률 분포도로 나타내기 때문에 약간의 다른 영상으로 인식하는 문제점이 있다.

C. 연구의 내용

내용기반 영상 검색은 영상정보를 효율적으로 처리하기위해 많이 이용되고 있다. 이러한 영상을 추출하기위해 히스토그램을 생성하여 컬러 에지를 추출하여야 하며 컬러 영상을 그레이 영상으로 변환 후 에지를 추출하여 보면 컬러영상에서는 명확하게 에지로 보이는 것이 그레이 영상에서는 명도 값이 유사하기 때문에 에지로 검출되지 못하는 에지 정보의 손실을 볼 수 있다. 또 기존 컬러 히스토그램 검색 방법들은 양자화 오류 등의 이유로 정확성이 떨어지고, 공간정보가 부족한 문제점이 있다. 이를 해결하기 위해 색상 정보를 HSV 공간으로 변환하여 순수 색상 정보인 색상(hue) 성분만을 양자화하여 히스토그램을 구해 명암, 이동, 회전 등에 강

인한 검색 특징으로 사용한다. 한편 공간정보가 부족한 문제점을 해결하기위해 색상 특징과 영역간의 상관관계를 고려하여 객체영역을 선정한다. 본 논문에서는 이러한 기존의 단점들을 극복하고자 컬러 영상 정보와 에지 정제에 기반한 개선된 컬러 에지 정제 기법을 제안한다.

본 논문에서 제안하는 컬러와 에지 정제 기반 영상 검색 알고리즘의 구현은 다음과 같다.

첫 번째로 RGB 컬러를 HSV 컬러 공간으로 변환하고 R, G, B, 256 : 256 : 256에서 HSV 8 : 2 : 2를 얻기 위해 양자화를 하고, 이 중 색상 정보만을 고려하여 히스토그램을 얻는다.

두 번째 단계로 컬러 영상을 그레이 스케일의 밀도 영상으로 변환하여 8개 양자화 값으로 양자화 한 후 히스토그램을 얻는다.

세 번째로 컬러 영상의 에지 정제는 수평과 수직을 찾기 위해 소벨 연산자(Sobel operator)를 적용한다. 소벨 연산자는 특징벡터 크기를 줄이기 위한 전처리(Pre-processing)로 간주된다.

네 번째 단계로 본 논문에서 제안한 알고리즘을 테스트하기 위해 테스트용 멀티미디어 데이터베이스를 구축하였다.

마지막으로 성능 평가를 위하여 구축된 데이터베이스를 이용하여 제안된 검색 알고리즘을 분석하였다.

이를 위해서 II장에서는 내용기반 영상 검색 시스템에 대한 내용에 대해서 살펴보고, III장에서는 내용기반 영상 검색 기법에 대한 연구에 대해 고찰한다. 그리고 IV장에서는 제안한 시스템을 설계하고, V장에서 이를 시험 영상을 통해 시뮬레이션을 하여 이를 분석하고, VI장에서 결론을 맺는다.

Ⅱ. 내용기반 영상 검색 시스템

A. 내용기반 영상 검색

내용기반 영상 검색(CBIR: Content-Based Image Retrieval) 시스템은 영상 자체의 시각 정보를 이용하여 영상 데이터베이스를 색인하고, 예제 질의를 통하여 사용자가 찾고자하는 영상과 유사한 것들을 검색하는 기능을 제공한다. 저마다의 영상 검색 시스템은 그에 사용되어지는 특징 종류와 표현 방법이 다르고, 특징들 간의 유사성을 측정하는 방법 또한 다르다. 내용기반 영상 검색 시스템의 성능은 영상을 대표할 특징의 선택 그리고 추출 방법, 유사성 측정 방법 등으로 결정된다. 일반적인 내용기반 영상 검색 시스템은 그림 2-1과 같이 구성된다.

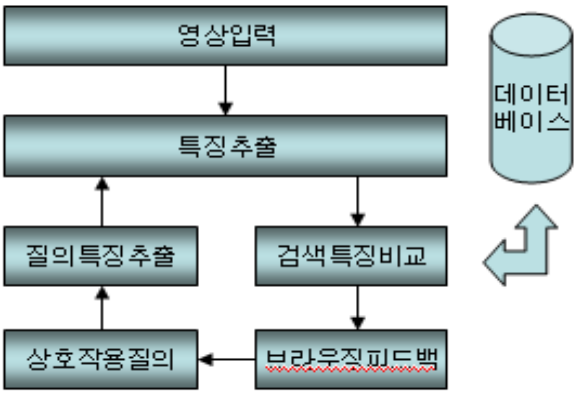


그림 2-1 내용기반 영상 검색 시스템 구성도

내용기반 영상 검색 시스템은 사용자에게 질의를 입력받는 질의 입력부와 영상을 분석하여 특징을 추출하는 특징 추출부, 특징간의 유사도를 계산하는 유사성 측정부 그리고 검색된 결과를 사용자에게 보여주는 결과 출력부로 구성된다.

질의 입력부는 사용자와 효율적인 상호작용을 할 수 있도록 시스템에서 제공하는 사용자 인터페이스 부분으로 사용자가 쉽게 질의할 수 있도록 영상에 의한 질

의 등과 같은 시각적 예제를 통한 질의 형태를 사용한다.

특징 추출부는 검색 시스템에서 사용하는 특징을 추출하는 기능을 수행한다. 시스템에 사용되는 영상 특징은 컬러, 질감, 모양 등으로 다양하며 특징의 종류와 표현 형태에 따라 영상에서 특징을 추출하는 알고리즘은 달라진다. 유사성 측정부는 사용자가 입력한 질의에서 추출된 특징과 데이터베이스 내의 저장된 특징간의 유사도를 계산하여 유사 여부를 결정한다. 일반적으로 특징들은 벡터 공간에서 표현되고, 특징 벡터들 간의 유사성은 벡터 거리 계산 척도에 의해 결정된다.

마지막으로 결과 출력부는 검색된 결과를 사용자가 잘 확인할 수 있도록 시각화하여 보여준다.

내용기반 영상 검색 시스템의 핵심 기술인 특징 추출부의 보다 효율적인 검색 시스템에 있어서 가장 중요한 요소는 영상을 어떠한 특징을 가지고 표현할 것인가와 추출된 특징들을 어떠한 방법으로 비교할 것인가 하는 문제이다. 검색에서 사용되는 특징은 원시적이거나 또는 구문적(semantic)이며, 특징 추출은 전적으로 자동적으로 수행되어져야 한다. 키워드가 영상의 내용을 나타낸다고 하더라도 수동적으로 할당된 키워드에 의한 검색은 내용 기반 영상 검색이라 하지 않는다.

기존의 내용 기반 영상 검색 시스템으로는 QBIC, VisualSEEK, Chabot, Photobook, SIMPLIcity 등이 있다.

표 2-1은 위의 시스템 및 기타 시스템을 비교한 것이다.

표 2-1 내용기반 영상 검색 시스템의 비교

시스템	검색방법	특징정보 추출방식	단점
QBIC	사용자 스케치 검색, 샘플 이미지에 의한 유사성 검색, 색상/모양/질감 기반 검색	점(point)으로 특징 벡터 표현, R*-트리를 이용한 인덱싱	통합되지 못한 검색 인터페이스, 의미 기반 정보 검색 불가능, 지식기반 검색 불가능
VisualSEEK	평균 색상 및 공간 검색, 시각적 특징에 의한 검색	영역적 평균 색상 히스토그램 추출과 자동화된 영역 추출	색상정보 검색의 큰 비중, 비효율적인 공간 관계 연산 시간
Chabot	텍스트 및 색상 기반 검색	평균 색상 히스토그램과 수동적인 키워드 입력	복잡한 질의 인터페이스, 내용정보간의 관계성 부족
Photobook	객체 기반 검색, 색상 및 질감, 형태에 기반한 검색	객체(윤곽선) 정보 추출, 색상/질감/형태 특징 추출	통합되지 않은 검색 인터페이스, 의미 기반 검색 불가능, 객체간의 관계성 부족
SIMPLIcity	영역 매칭 검색	영역 세그멘테이션 분할, 의미적 분류, 영역특징 추출	지식기반 검색 불가능, 영역적 연산 시간 소요

B. 영상 검색을 위한 내용기반 상호작용

내용기반 영상 검색 시스템은 효율적인 검색을 위하여 내용에 의한 질의를 통한 사용자와의 효과적인 상호작용을 제공한다. 내용에 의한 질의는 특정한 내용을 포함하는 영상들을 검색하기 위하여 사용되어진다. 질의는 질의 형태에 따라 일치형 질의와 범위형 질의, K-최근접 이웃 질의로 구분된다.

일치형 질의는 질의와 동일한 것 또는 복사본들을 찾기 위한 목적으로 사용되어진다. 여기서 복사본이란 원 영상에 일반적인 왜곡이 가해진 영상이거나 프린트 또는 스캐너에 의해 발생되는 잡음 등을 포함한 영상을 의미한다.

범위형 질의는 특징에 대한 범위를 제공함으로서 이루어진다. 예를 들어 “초록색(green)과 파란색(blue) 사이의 컬러를 가지는 모든 영상을 찾아라.”와 같이 특징 공간에서 질의의 범위에 포함되는 영상들은 질의와 유사하다고 간주되어진다.

K-최근접 이웃 질의는 검색 결과 항목을 K개만을 원할 때 사용한다. 예를 들어 “예제와 가장 유사한 10개의 영상을 찾아라.”와 같이 질의가 주어지면 시스템은 질의와의 유사성에 따라 항목들을 정렬하고 상위 10개의 항목만을 검색 결과로 제공한다. 이러한 질의 형태들에 대하여, 시각 예제에 의한 질의가 가장 적절한 형태이다.

시각 예제에 의한 질의는 시각 내용의 저수준 특징의 지각적 관점을 표현하는데 아주 적절한 패러다임 이다. 이 패러다임은 사용자가 프로토타입 영상을 참조 예제로 제공하도록 요구한다. 시각 예제에 의한 질의의 종류에는 아이콘닉 질의와 페인팅에 의한 질의, 스케치에 의한 질의, 영상에 의한 질의 등이 있다. 아이콘닉 질의는 고수준 개념기반 검색에 적절하다. 아이콘은 객체/개체의 속성을 표현한다. 아이콘닉 질의는 미리 정의된 아이콘 집합과 불리언 연결어를 통한 시각 언어에 따르는 시각 문장을 생성하여 사용한다. 이 방법은 영상에서 객체들의 공간적 관계를 표현하기에 매우 유용하다. 페인팅에 의한 질의는 컬러기반 검색에서 유용하게 사용되며 사용자가 영상에서 두드러진 컬러나 컬러 영역과 공간적 배열을 재현하여 페인팅 함으로써 이루어진다. 이러한 방법은 사용자에게 색상에 대한 충분한 지식을 요구하며, 또한 질의가 사용자의 페인팅 능력에 의존한다. 스케치에 의한 질의는 예제와 유사한 형태를 가지는 객체를 포함하는 영상을 검색할 때 일반적으로 사용되어진다. 이 방법 또한 페인팅에 의한 질의와 마찬가지로 사용자의 스케치 능

력에 의존하는 경향이 있다. 이러한 문제는 객체의 윤곽을 미리 제공하는 방법과 원본 객체를 아래에 두고 객체의 윤곽을 쫓아가면서 복잡한 객체의 실루엣을 스케치하는 방법에 의하여 어느 정도 해결될 수 있다. 영상에 의한 질의는 전역 컬러, 질감 또는 구조 속성을 기반 한 질의나 검색하고자 하는 영상을 재현하기가 쉽지 않을 때 사용한다.

C. 내용기반 영상 검색을 위한 특성 표현

색상 정보는 영상에서 객체를 식별하게 하는 중요한 시각 요소이며, 영상 검색에 가장 널리 사용되어지는 시각 특징 중의 하나이다.

색상 특징을 나타내는 대표적인 방법으로는 컬러 히스토그램 (Color Histogram)에 의한 방법과 컬러에 의한 지각 (Color-Induced Sensations)에 의한 방법이 있다. 컬러 히스토그램은 영상의 하위 수준 색상 속성을 기술하는 가장 전통적인 방법이다. 컬러 히스토그램은 영상의 색상 내용에 대한 매우 효과적인 표현이며, 컬러 히스토그램의 계산 방법은 각 색상 값에 얼마나 많은 화소가 속하는지 계소하는 통계적 방법을 사용한다.

영상간의 유사성은 각 화소가 갖는 색상 값의 차이에 의해 결정된다. 색상을 이용한 대부분의 연구들은 색상 값을 얻기 위하여 컬러 히스토그램 방법을 이용한다.

그림 2-2와 같이 R, G, B의 세 가지 색상 채널별 히스토그램을 정렬한 컬러 히스토그램은 주어진 컬러 영상은 초기에 특별한 변환을 가하지 않으면 R, G, B 값을 보유하고 있기 때문에, 영상을 구성하는 각 픽셀의 RGB 색상 공간에서의 분포 특성은 검색과정에서 유용하게 사용될 수 있다.

컬러 히스토그램으로서 표현되는 색상 분포는 영상을 나타내는 객체의 각 구성요소에 대한 내용보다는 전체적인 영상의 내용을 나타내는데 적합하다. 또한 컬러 히스토그램 방법은 객체의 회전이나 이동 등과 같은 기하학적 변형에도 비교적 영향을 받지 않으며 구현 알고리즘도 간단하다.

그러나, 빛의 밝기와 영상의 배경과 객체의 크기에 민감하고, 전혀 다른 종류의 영상도 같은 색상 분포를 가질 수 있다는 단점이 있다. 또한 색상은 화소들 사이의 위치 영역에 대한 정보가 부족하기 때문에 객체의 대략적인 구분 정보만을 가진다.

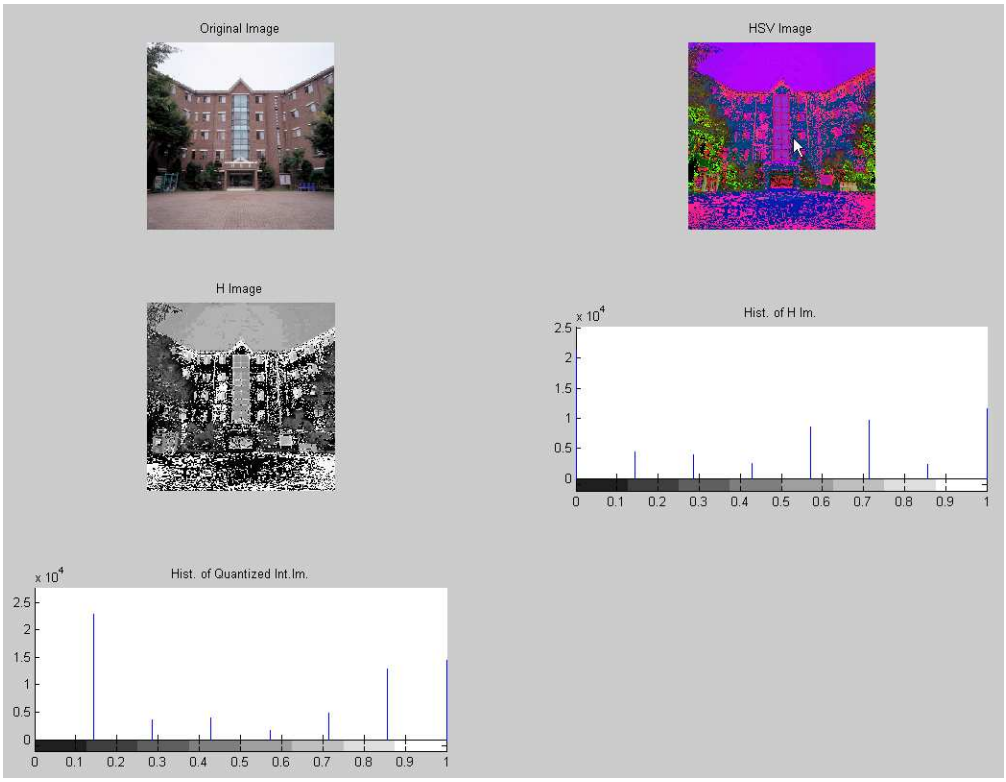


그림 2-2 컬러 히스토그램

색상 정보 기반 인덱싱 기법은 영상 데이터베이스의 검색에 있어서 아주 중요한 방법 이다. 영상의 다른 정보들과 비교할 때 색상 특징은 영상에 대한 인덱싱을 구축하는데 가장 중요한 저수준의 정보를 가지고 있다. 색상 정보는 잡음이나 다른 배경의 복잡도에 덜 민감하며 영상 내의 물체의 이동, 회전 등의 변화에 영향을 적게 받는 통계정보를 계산하는데 사용될 수 있다. Swain과 Ballard[31]는 히스토그램 인터섹션(histogram intersection)이라는 색상 검색 방법을 제안하였다. 이 방법에서는 각 영상에 대해 색상 히스토그램을 구성한 후 각 히스토그램간의 공통부분을 취하는 방식으로 거리를 구하여 비교하였다. 이 방법은 계산과 정보 표현에 있어서 간결하기 때문에 많은 상황에서 유용하게 사용되고 있지만, 조명이 변화하는 경우에서는 성능이 저하되는 단점을 가지고 있다. 이러한 문제를 해결하기 위해서 Funt와 Finlayson은 일정한 색상 인덱싱(constant color indexing)이라는 알고리즘을 개발하여 확장된 Swain 방법을 발표하였다. 이 방법은 색상 비율의 히스토그램을 비교함으로써 검색을 수행한다. Funt 등의 논문에서는 제안한 알고리즘이 일정

한 조명에서는 Swain의 방법보다 조금 성능이 떨어지지만 조명이 변화하는 경우에서는 성능이 크게 향상됨을 보였다. 또한 Stricker 등[9]도 히스토그램 인터섹션 방법의 단점을 보강하여 누적 히스토그램을 이용하는 방법 및 각 색상 채널의 모멘트를 이용하여 색상 정보를 표현하고 유사도를 비교하는 방법을 제시하였다. 반면에 Mehtre 등은 색상 비교를 위해 거리 방법(distance method)과 기준 색상표 방법(reference color table method)을 개발하였다. 거리 방법은 특징으로서 세 가지의 색상 성분의 각각의 1차원 히스토그램의 평균값을 사용하였다. 두 특징 벡터간의 거리가 두 특징 사이의 유사도를 나타내기 위해 사용된다. 기준 색상표 방법에서는 한 세트의 기준 색상들이 먼저 색상표로 정의된다. 영상내의 각 색상 픽셀들은 색상표 상의 가장 가까운 색상으로 정의된다. 새롭게 할당된 색상을 가진 픽셀들의 히스토그램이 영상의 색상 특징으로 사용된다.

그러나 색상 히스토그램을 사용하는 방법들은 다음과 같은 단점들을 가지고 있다. 첫째, 히스토그램은 차원을 줄이기 위해 양자화를 필요로 한다. 전형적인 24 비트 컬러 영상은 224 개의 빈을 발생시키며, 이것은 최소한 2MB의 저장 공간을 필요로 한다. 따라서 빈의 수를 줄이기 위해 양자화가 필요하며, 이것에 의해 색상 정보의 손실이 일어나며 또한 어느 정도의 양자화 과정이 최적인지, 어느 정도까지 양자화를 수행해야 하는지에 대한 규칙이 없다. 둘째, 히스토그램이 만들어지는 색상공간이 검색결과에 상당한 영향을 줄 수 있으며 또한 양자화의 방법 및 양자화 정도를 결정한다. 예를 들면, 균일한 양자화는 RGB 공간에서는 유용하게 사용될 수 있는데, 그것은 RGB 공간에서 각 색상 성분의 분포가 균일하기 때문이다. 그러나 Mensell이나 LUV 공간에서는 균일한 양자화로는 불충분하다. 셋째로 색상 히스토그램을 사용하면 색상의 제외를 고려하기가 어렵다. 양자화의 정도에 따라 빈들이 나타내는 색상들이 변화하기 때문에 제외하고자 하는 색상을 나타내는 빈을 찾아내기가 어려우며 또한 제외하고자 하는 색상과 다른 색상들이 제외되는 경우가 발생할 수 있다. 마지막으로 히스토그램은 영상의 전체적인 색상 정보만을 추출한다. 즉, 어떠한 공간적인 정보도 이용되지 않는다는 단점을 가진다.

지금까지 내용기반 영상 검색에 관련된 기존의 연구를 분석하였다. 앞서 지적한 바와 같이 색상 히스토그램은 전혀 다른 영상이지만 유사한 히스토그램을 가질 수 있으므로 이를 영상 검색에 적용할 경우 사용자의 의도와는 전혀 다른 영상을 검

색하는 결과를 가져온다. 이러한 히스토그램의 단점인 공간정보 부재문제를 해결하기 위해 영역을 고정 또는 가변 분할하는 방법을 시도하고 있으나, 유사한 영역의 분할이 용이하지 않아 검색결과가 좋지 않았다.

D. MPEG-7 영상 특징 정보

"한 장의 그림은 천 마디의 말로도 표현이 어렵다(A picture is worth a thousand words)."든지, "보기(듣기) 전에는 믿기 어렵다[Seeing(Hearing) is believing]." 등의 말은 멀티미디어 정보의 속성을 단적으로 드러내고 있다. 일반적으로 비디오 Archive로부터 "Terminator II 영화에서의 오토바이가 질주하는 장면"을 찾으려고 하거나, 또는 "오늘 내가 즐겨 시청한 것과 유사한 종류의 프로그램"을 방송하는 TV 채널을 찾으려고 하는 경우, 이를 키워드 기반의 표현 및 검색 기술로 구현하기란 대단히 어렵다. 이러한 문제를 해결하기 위해 내용 기반 멀티미디어 정보 검색을 효율적으로 지원하기 위한 기술을 개발하고 이를 국제 표준화하고자 하는 것이 MPEG-7(Moving Picture Expert Group)이다. 즉, 기존에 표준화되었거나 지금 표준화가 진행되고 있는 MPEG-1/2/4 등은 오디오 비주얼 데이터의 데이터 압축을 그 목표로 하였으나, MPEG-7은 데이터 그 자체가 아닌 데이터의 내용에 대한 표현 방법을 다루는 것이다. 이를 다른 말로 "메타데이터(Metadata)", 또는 "Bits about bits"라고 표현하기도 한다.

MPEG-7에서는 주로 오디오 비주얼 정보(정지화상, 픽처, 그래픽, 3D 모델, 오디오, 스피치, 비디오)의 표현을 그 대상으로 하고 있으나, HTML, SGML, 또는 RDF 등이 목표로 하고 있는 텍스트 문서의 표현에 대한 것은 표준화 범위에 포함하지 않는다. 필요한 경우 이러한 문서 포맷에 관한 표준화가 적용될 수 있도록 하는 기능은 갖게 될 것이다. 그림 2-3은 MPEG-7과 관련된 정보 처리 과정과, 이 중 MPEG-7이 표준화 하고자 하는 범위를 개략적으로 보여준다.

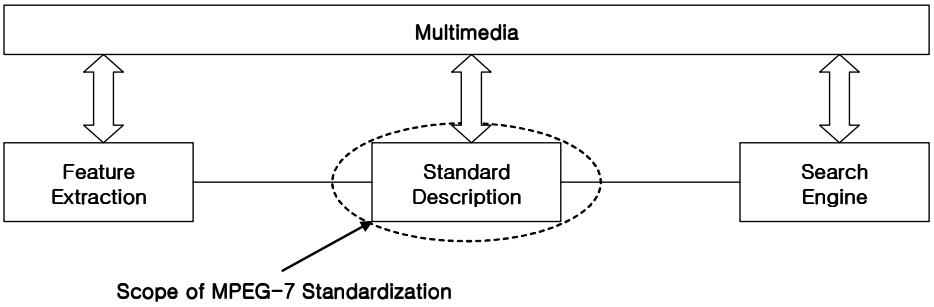


그림 2-3 MPEG-7 표준화의 범위

이 중 특징 추출(Feature Extraction) 방식 및 검색 엔진(Search Engine)은 표준화의 대상으로 하지 않는다. 왜냐 하면, 이들은 여러 멀티미디어 정보 처리 시스템 간의 호환성을 획득하는 것과는 상관이 없으며, 업체 간의 경쟁이 가능한 부분으로 남겨 놓는 것이 향후 계속적인 기술 발전을 유도하기 위하여 바람직하기 때문이다. MPEG-7은 멀티미디어 정보의 특징 및 내용 표현 방식에 초점을 맞추어 다음의 사항을 표준화하고자 한다.

- 기술자(Descriptors : D)와 기술 구조(Description Scheme : DS)
- 기술구조를 표현하기 위한 기술정의언어(Description Definition Language : DDL)
- 색인/저장/전송을 효율적으로 하기 위하여 사용될 코딩된 기술(Coded Description)

MPEG-7에서 정의하는 용어는 다음과 같다. 데이터(Data)는 표현 형식에 상관없이 오디오 비주얼 정보를 의미하며, 특징(Feature)은 오디오 비주얼 정보의 성격이나 속성을 나타낸다. 기술자(Descriptor)는 특징을 어떤 값에 연결시켜주는 도구이며, 기술 구조는 데이터를 여러 개의 기술자로 나타내기 위한 그릇이다. 표 2-2에 이 용어들의 개념을 명확히 하기 위하여 대표적인 특징 형태, 특징 및 기술자의 예를 보였다.

표 2-2 오디오 비주얼 데이터의 대표적인 특징 및 기술자의 예

특징 형태 (*Feature Type*)	특징(*Features*)	기술자(*Descriptors*)
N-dimensional spatial-temporal structure	duration of a music segment	time code, etc.
	trajectory of objects	chain code, etc.
Statistical Information	color	color histogram, etc.
	audio frequency content	average of frequency components, etc.
Objective features	color of an object	color histogram, etc.
	shape of an object	a set of polygon vertices, a set of moments, etc.
Subject features	emotion(happiness, angry, sadness, etc.)	a set of eigenface parameters, text, etc.
	style	text, etc.
Concepts	event	text, etc.
	activity	text, etc.

한 종류의 기술자나 기술구조가 위에 열거한 모든 응용 분야를 다 만족시키는 것은 어려운 일이며, 효율적이지도 않다. 따라서 MPEG-7에서는 기술자나 기술구조의 제안 사항을 주요 응용 분야 별로 나누어 표준화되었다. 응용 분야와 관련하여 논의가 진행되고 있는 것 중의 하나는 Pull Model과 Push Model의 구분이다. 전자는 국지적으로나 네트워크상에 있는 데이터베이스로부터 사용자의 질의에 가장 가까운 데이터를 찾아주는 응용 분야이고, 후자는 방송에서와 같이 정보 제공자로부터 일방적으로 데이터가 공급될 때, 이로부터 원하는 정보를 필터링하여 주는 응용 분야이다. 두 응용 분야의 주요 특징상 차이점은 후자의 경우 정보 검색 시 실시간성이 강하게 요구된다는 것이다. 그림 2-4에는 MPEG-7의 표준화 요소와 응용 시스템 기능 구성 간의 관계를 보였다.

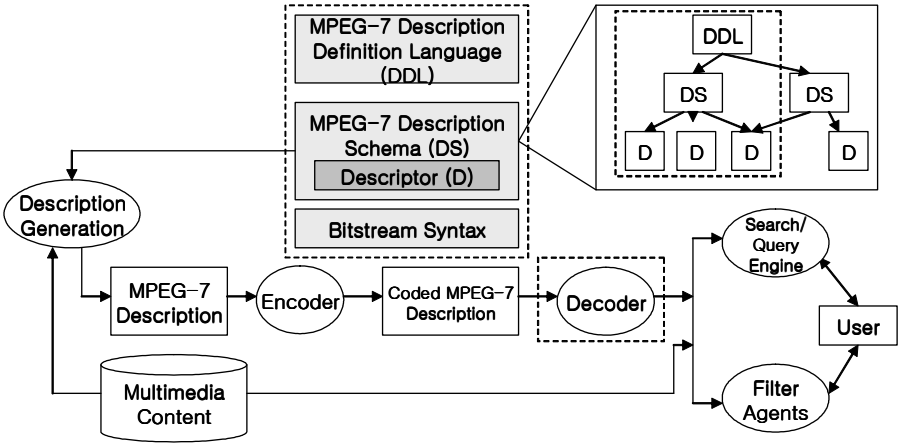


그림 2-4 MPEG-7의 표준화 요소와 응용 시스템 기능 구성

MPEG-7에 관한 요구 사항은 기술정의 언어에 대한 요구 사항, 오디오 비주얼 데이터에 공통적으로 적용되는 요구 사항, 그리고 오디오 또는 비주얼 데이터별 특성에 연관된 요구 사항으로 나누어진다.

우선 기술정의 언어에 대한 요구 사항으로는

① MPEG-7 응용 분야를 지원하는 어떤 기술구조도 쉽게 생성되어야 한다

② 문법이 모호하지 않고 인터프리터 등에 의해 쉽게 파싱(parsing)될 수 있어야 한다.

③ 오디오 비주얼 데이터를 포함한 복합적인 멀티미디어 정보에 대한 기술구조를 생성할 수 있어야 한다.

④ 기술구조 내에 그 코드를 포함시킬 수 있는 기능을 가져야 한다.

이 코드는 표준에서 정의되지 않은 새로운 기술자와 이 기술자에 대한 유사도 측정 기준 등을 사용자가 인식할 수 있도록 하는데 사용된다.

오디오 비주얼 데이터에 공통적으로 적용되는 기술자 및 기술구조에 대한 일반적인 주요 요구 사항으로는

① 데이터에 부가된 주석, 데이터의 시공간적인 구조, 주/객관적 특징, 또는 사건이나 활동 등의 개념까지도 포함하는 모든 특징의 표현을 지원하여야 한

다.

② 사용자의 요구에 적절히 대응할 수 있는 계층적인 내용 압축 및 표현 방식을 지원하여야 한다.

③ 비주얼 질의를 사용해서 오디오 정보를 찾고 오디오 질의를 사용하여 비디오 정보를 찾을 수 있도록 통합성(Cross-modality)을 지원하여야 한다.

④ MPEG-7의 여러 가지 기술구조 상호간의 호환성을 가져야 한다.

⑤ 특징별 우선순위를 지정하여야 한다.

⑥ 기술자의 확장성(Scalability)을 지원하여 한다.

⑦ 텍스트 기반 기술자의 경우 사용 언어의 지정, 모든 언어를 지원할 수 있는 글자 세트, 그리고 언어 사이의 변환을 위한 도구 등이 제공되어야 한다.

기능적인 관점에서의 요구 사항으로는

① 내용 기반 검색 및 유사도 기반 검색 기능이 제공하여야 한다.

② 데이터와 동기화 되어 제공되거나(Streamed) 또는 데이터와 별도로 제공되는(Stored) 기술자의 지원하여야 한다.

③ 대화형 질의의 지원하여야 한다.

④ 소스 데이터 또는 관련 정보의 위치를 알려줄 수 있는 기능을 가지고 있어야 한다.

⑤ 정보의 개략적인 파악(preview)이 가능하도록 브라우징(Browsing) 기능 등을 지원하여야 한다.

가 있다.

마지막으로 비주얼 또는 오디오 정보 각각의 특성에 따른 요구 사항으로는 지원되는 특징의 종류, 지원되는 데이터 포맷 및 클래스에 대한 요구 사항이 포함된다.

Ⅲ. 내용기반 영상 검색 기법

영상 검색에 사용되는 컬러는 이미지를 인지하는 기초적인 시각 특징이다. 인간의 눈을 통하여 들어온 빛은 시신경을 통하여 두뇌에 전달되고 두뇌에서는 이를 어떤 메커니즘을 거쳐 색으로 감지하게 된다. 이러한 메커니즘은 정확히 밝혀지지 않았고 이를 해석하는 방법은 여러 가지가 있을 수 있다.

빛의 삼원색은 빨강(Red), 녹색(Green), 파랑(Blue)을 말한다. 이 삼원색을 가산혼합하여 임의의 색을 만들 수 있다는 사실에 기초한다. TV나 컬러 모니터, 영사기, 무대조명 등에 사용되며 RGB 컬러공간의 기초가 된다.

색은 색상(Hue), 명도(Brightness), 채도(Saturation)의 3요소를 가지고 있는 것으로 생각한다. 색상이란 빨강, 주황, 노랑, 초록, 파랑, 남색, 보라 등과 같이 빛의 스펙트럼 상에 나타나는 색들을 말한다. 명도는 색의 밝고 어두운 정도를 의미하며 이것은 물리적으로는 시감반사율의 크고 작음을 의미한다. 여기서 시감 반사율은 물체표면에 입사되는 빛과 반사되는 빛의 비율을 말한다. 명도는 색의 3속성 중에서 인간의 눈에서 가장 예민한 부분이며, 그 다음으로 색상, 채도 순이다. 채도는 색의 지각적인 면에서 보았을 때 색의 강약이라 할 수 있다. 진한 색, 연한 색, 흐린 색, 맑은 색 등이 채도의 고저를 가리키는 말이다. 순색에 무채색의 혼합양이 많아지면 채도가 낮아지고 적어지면 채도가 높아지게 된다. 이를 이용한 컬러공간에는 HSI, HSV 공간이 있다.

색의 삼원색 청록색(Cyan), 심홍색(Magenta), 노랑(Yellow)을 말한다. 빛의 삼원색으로부터 만들어지는 2차색상인 청록, 심홍, 노랑으로부터 감산혼합으로 무한한 색상을 만들어 낼 수 있다는 사실에 기초한다. 예를 들면 노랑물감과 심홍색 물감이 종이위에 겹치게 되면 빨강색이 만들어 지는데 그 원리는 두 물감이 합쳐진 부분에서 녹색과 파랑의 빛의 파장을 흡수하고 (빼고) 빨강만 반사하기 때문이다. 이런 이유로 감산혼합이라 부른다. 물감으로 그림을 그리거나 프린터에서 프린트할 때 감산혼합방식이 적용되며 CMY 공간의 기초가 된다.

컬러 유사도를 이용하는 이미지 검색을 위해서는 컬러를 표현하는 공간 모델이 필요하다. 가장 일반적으로 사용되는 하드웨어 중심 공간은 컬러 모니터와 컬러 비디오, 카메라를 위한 RGB 공간, 조명 기기를 위한 XYZ 공간, 컬러 프린터를 위한 CMY(Cyan, Magenta, Yellow) 공간과 CMYK(Cyan, Magenta, Yellow, Black) 공

간, 컬러 텔레비전 방송을 위한 YIQ(Luminace, Inphase, Qu adrature) 공간 등이 있으며 컬러를 재현하는 장치의 특성에 따라 정의된다. 컬러를 표현하는 공간 모델 중에서 사용자 중심 공간은 HSV(Hue, Saturation, Value) 공간, HSI(Hue, Saturation, Intensity) 공간, HSB(Hue, Saturation, Brightness) 공간, L*u*v* 공간, L*a*b* 공간 등으로 분류할 수 있다. 사용자 중심 공간은 인간의 컬러 인식을 기초로 정의되며 인간이 느끼는 컬러간의 거리를 3차원 컬러 공간으로 표현한 다. 하나의 컬러 공간에서 다른 컬러 공간으로 컬러 데이터를 변환할 때 컬러 공간의 기하학적인 모양이 처리 과정에 영향을 미친다. RGB 공간은 이미지 처리장치의 하드웨어에 적합하여 많이 사용되고 있다. 여러 가지 컬러 공간 중에서 RGB 공간,HSV 공간, CMY/CMYK 공간, NTSC 공간, YCbCr 공간, XYZ 공간, YUV 공간에 대하여 기술한다.

A. 컬러 기반 영상 검색

1. 컬러 히스토그램 인터섹션

내용기반 영상 검색 방법 가운데 컬러를 이용한 대표적인 검색방법은 컬러 히스토그램 인터섹션(Color Histogram Intersection)이다. 컬러 영상의 경우 밝기영상에 비해 정보량이 많음을 이용하여 영상의 통계적 특성을 이용한 3차원 히스토그램을 생성한 후 히스토그램 공간상에서 데이터베이스 영상과 입력영상의 정합을 시도하는 색인화 알고리즘을 제안하였다. 이 알고리즘은 간단하고 적용하기 쉬우면서도 물체의 회전, 이동 및 어느 정도의 크기 변화, 그리고 영상 획득 시 각도의 변화 등에 강인하게 동작한다는 장점을 가진다. 이 방법에서는 컬러 공간으로 식 (3-1)과 같은 R, G, B로부터 얻어지는 rg, by, wb의 보색 축(opponent color axis)을 이용하였다.

먼저 영상의 컬러를 식 (3-1)과 같이 공간변환을 행한다.

$$rg = R - G \tag{3-1}$$

$$by = 2B - R - G$$

$$wb = R + G + B$$

영상의 컬러 히스토그램을 구하기 위해서 위 세 축의 구간을 나누어야 하는데, 여기서는 세기(intensity)축인 wb가 광원으로부터의 거리 및 밝기의 변화 등에 민감하기 때문에 영향을 줄이기 위해 구간을 더 크게 하였다. rg , by 는 16개의 구간으로 나누었으며 wb는 8개의 구간으로 나누어 히스토그램은 총 2048 (16x16x8)개의 그룹을 가진다. 히스토그램이 구해지면 두 영상의 색상 히스토그램은 히스토그램 인터섹션 방법에 의해 비교된다.

일반적으로, 영상을 N 개의 컬러로 양자화 한다면, 컬러 히스토그램 H 는 벡터 $\langle h_1, h_2, h_3 \ldots, h_N \rangle$로 구성된다. 이때 h_j는 영상 내의 컬러를 가진 화소수의 수를 나타낸다. 입력 영상(query image)이 I(image)이고, 데이터베이스 영상이 M(model)일 때, 히스토그램 인터섹션은 식 (3-2)와 같이 두 히스토그램의 교집합으로 정의된다.

$$H(I, M) = \sum_{j=1}^{n} \min(I_j, M_j) \tag{3-2}$$

여기서 j는 히스토그램에서의 각 구간을 나타낸다. 이 결과는 영상의 크기에 따라 값이 변하므로 식 (3-3)과 같이 정규화 시켜서 사용한다.

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

식(3-3)은 입력 영상과 데이터베이스 영상의 유사도를 나타내며, 0에서 1사이의 값으로 정규화 된 값이다. 따라서 서로 유사도가 높은 두 영상에 대해서 1의 값에 가까운 값을 가지게 된다. 만일 두 영상이 동일한 경우는 1의 값을 가지고, 두 영상이 전혀 다른 경우는 0의 값을 가진다. 보는 바와 같이 히스토그램 인터섹션은 알고리즘이 비교적 간단하고, 식 (3-3)에서 보다시피 화소의 위치에는 상관없이 유

사도가 측정된다는 것을 알 수 있다. 즉, 어떠한 영상이라도 히스토그램만 유사하다면 검색되기 때문에 영상 내에서 어떤 목적물의 이동이 있는 영상에 대해서도 잘 검색되고, 영상의 회전이 있을 때도 좋은 검색 결과를 나타내는 것을 알 수 있다. 그러나 영상 내의 컬러 빈도수만을 이용하기 때문에 히스토그램은 같으나 영상의 내용은 전혀 다를 경우에도 검색되어 질 수 있는 단점을 가지고 있다. 또, 밝기 변화나 컬러 변화가 없는 영상을 가정한 알고리즘이므로 영상내의 밝기나 컬러 변화에는 민감하다는 단점을 가지고 있다. 특히, 데이터베이스의 크기가 커질수록 그 문제가 심각해진다.

2. 색상 모멘트를 이용한 기법

색상 히스토그램이나 누적 히스토그램은 검색에 적용 시 색인 과정에서 양자화가 필요하며 최적의 양자화 과정을 찾아내는 것은 아주 어려운 일이다. 따라서 인덱스 생성 과정에서 파라미터에 의존하지 않는 방법이 요구되며 이 중의 하나가 색상 분포의 특징을 이용하는 것이다. Stricker 등은 색상 검색을 위한 색상 분포의 특징으로서 각 색상 채널의 모멘트를 이용하는 방법을 제안하였다. 색상 모멘트를 이용한 방법을 보면 이 방법은 확률 분포가 그것의 중심 모멘트(Central Moment)들에 의해 유일하게 표현될 수 있다는 사실을 바탕으로 하고 있다. 이미지의 색상 분포 즉 색상 히스토그램은 확률 분포로 해석될 수 있으며, 따라서 색상 분포는 모멘트에 의해 표현될 수 있다. 여기서는 색상 정보를 위한 인덱스로서 각 색상 채널의 1, 2, 3차 중심 모멘트를 이용하였다. j 번째 이미지 픽셀에서 i번째 색상 채널의 값이 pij 라면, 이 채널과 관계된 인덱스는 식 (3-4)과 같이 계산된다.

$$E_i = \frac{1}{N}\sum_{j=1}^{n} p_{ij} \qquad (3\text{-}4)$$

$$\sigma_i = \left(\frac{1}{N}\sum_{j=1}^{n}(p_{ij}-E_i)^2\right)^{\frac{1}{2}}$$

$$s_i = \left(\frac{1}{N}\sum_{j=1}^{n}(p_{ij}-E_i)^3\right)^{\frac{1}{3}}$$

여기서, N 은 이미지의 픽셀 총 수이다.

r 개의 색상 채널을 가지는 두 이미지 I, H 의 각각의 색상 정보 인덱스를 Ei, Fi, σi, ζi 그리고 si, ti 라고 하면 두 이미지의 색상 유사도는 식 (3-5)와 같이 얻어진다.

$$d_{mom}(H, I) = \sum_{i=1}^{r}(w_{i1}|E_i - F_i| + w_{i2}|\sigma_i - \zeta_i| + w_{i3}|s_i - t_i|) \qquad (3\text{-}5)$$

여기서 $w_{kl} \geq 0\,(1 \leq l, k \leq 3)$는 사용자가 정의한 계수이다. 계수를 조정함으로써 적용 환경에 맞게 유사도 함수를 정할 수 있다.

이 방법은 이미지의 색상 정보를 적은 수의 인덱스를 이용하여 나타낼 수 있는 장점을 가지며 성능에 있어서도 히스토그램을 이용한 방법 이상의 결과를 보여주었다. 그러나 각 색상 채널사이의 상관관계에 대한 어떠한 정보도 이용하지 않으므로 그 성능에 있어서 한계가 있다.

색상 정보를 이용한 대표적인 연구로는 색상 모델의 선택이 어떻게 히스토그램에 매칭 되는가를 분석한 연구, 이미지의 지역적 색상 속성을 특성화하기 위해 컬러 히스토그램을 집합을 제안한 연구, VisualSEEK 시스템에서 컬러 이미지의 위치 영역에 사용한 색상 집합에 대한 연구 등이 있다.

3. 그레이 영상 처리 방법

일반적으로 물체를 구분하는 것은 주로 그 물체의 에지 성분에 의하므로 영상처리에서 에지 검출은 중요한 작업이 된다. 에지는 어떤 화상의 특징, 보통 밝기에서 크고 갑작스런 변화로서 정의된다. 에지에 해당하는 픽셀을 구하는 방법을 에지 검출이라 하며 에지 검출 기법도 에지의 종류에 따라 달라지게 된다. 에지 검출 방법으로는 로버트(Roberts) 필터, 소벨 필터, 프리위트(Prewitt) 필터, 라플라시안(Laplacian) 필터 등을 들 수 있다.

4. 벡터 정합법

앞 항에서 설명하였던 에지 추출은 컬러 영상을 밝기 값으로 변환시킨 후 변환

되어진 그레이 영상에 대하여 에지를 추출하는 방법이다. 컬러 영상을 그레이 값으로 변환 후 에지를 추출하는 경우 컬러 영상이 지니고 있는 에지 정보의 손실이 발생 할 수 있다. 컬러 영상을 구성하고 있는 각 채널별 에지맵을 생성한 결과를 보면 일부 채널에서는 에지 정보를 손실하지 않고 찾아주고 있는 것을 볼 수 있다. 일반적으로 채널별 에지 정보를 이용하여 에지를 추출하는 컬러 에지 추출방법은 각 채널별 컬러 에지맵을 혼합하여 하나의 에지맵으로 변환하는 퓨전 방법[16][17]과 컬러 영상을 채널로 분류한 후 채널별로 구배도를 구하고 세 개의 구배도를 하나의 값으로 변환한 후 이것으로부터 최종 에지맵을 생성하는 다차원 구배도 방법[18], 마지막으로 픽셀 값을 하나의 벡터 값으로 취급하여 컬러 에지를 추출하는 방법[19]으로 크게 분류된다.

a. 영상 퓨전 방법*(output fusion method)*

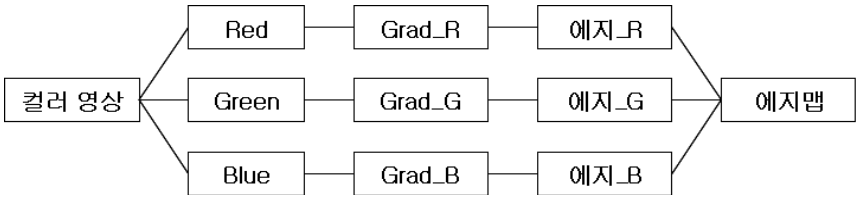


그림 3-1 퓨전 방법의 처리과정

퓨전 방법은 컬러 영상을 각각의 세 가지 채널로 분류한 후 각 채널별로 구배도를 구하고 에지를 생성한 후 여러 가지 방법을 이용하여 하나의 에지맵으로 출력하는 방법이다. 이때 채널 선택은 일반적인 R, G, B 채널이외에도 HSI, YUV등의 여러 가지 채널들이 사용된다.

jianping. Fba와 David.K. Y. Yau는 YUA 컬러 채널을 이용하여 컬러 에지를 검출하는 방법으로 다음과 같은 방법을 사용하였다.

컬러 영상의 휘도(luminance) 채널인 Y와 색차(chrominance) 요소인 U와 V요소의 각각 각 픽셀에 값을 갖고 있는 2차원 영상으로 표현한다.

이때, Y채널은 실제 픽셀 그대로 사용하고 U와 V채널 요소는 해상도를 반(half solution)으로 줄여서 사용한다. 픽셀 (x, y)에 대하여 HOE(x, y), VOE(x, y), NOE(x, y), SOE(x, y)로 정의된 패턴을 적용시켜 에지의 강도를 구한다.

(x-1, y-1)	(x-1, y)	(x-1, y+1)
(x, y-1)	(x, y)	(x, y+1)
(x+1, y-1)	(x+1, y)	(x+1, y+1)

(a) 픽셀(x, y)과 인접 픽셀들

1	2	1
0	0	0
-1	-2	-1

(b)HOE성분

1	0	-1
2	0	-2
1	0	-1

(c)VOE성분

2	1	0
1	0	-1
0	-1	-2

(d)NOE성분

0	1	2
-1	0	1
-2	-1	0

(e)SOE성분

그림 3-2 채널별 구배도를 위한 마스크들

이때, HOE(x, y)는 수평성분을, VOE(x, y)는 수직성분을, NOE(x, y)는 북동쪽 방향 성분을, SOE(x, y)는 북서쪽 방향 성분을 정의한다. 각 채널에 대하여 식 (3-6)부터 식 (3-9)의 패턴을 적용하여 채널별 구배도를 구한다.

$$\begin{aligned} HOE(x, y) = \mid & I(x-1, y-1) + 2I(x, y-1) \\ & + I(x+1, y-1) - I(x-1, y+1) \\ & - 2I(x, y+1) - I(x+1, y+1) \quad \mid \end{aligned} \tag{3-6}$$

$$\begin{aligned} VOE(x, y) = \mid & I(x-1, y-1) + 2I(x-1, y) \\ & + I(x-1, y+1) - I(x+1, y+1) \\ & - 2I(x+1, y) - I(x+1, y+1) \quad \mid \end{aligned} \tag{3-7}$$

$$\begin{aligned} NOE(x, y) = \mid & 2I(x-1, y-1) + 2I(x, y-1) \\ & + I(x-1, y) - I(x+1, y+1) \\ & - 2I(x+1, y) - I(x+1, y+1) \quad \mid \end{aligned} \tag{3-8}$$

$$SOE(x, y) = | 2I(x, y-1) + 2I(x+1, y-1) + I(x+1, y) - I(x-1, y+1) - 2I(x-1, y) - I(x, y+1) | \quad (3\text{-}9)$$

픽셀(x, y)의 지역 최대 에지 강도인 MOE(x, y)는 식 (3-6)부터 식 (3-9)의 네 개의 값중 가장 큰 값을 갖는 것으로 식 (3-10)와 같이 정한다.

$$MOE(x, y) = \max HOE(x, y),\ VOE(x, y),\ NOE(x, y),\ SOE(x, y) \quad (3\text{-}10)$$

다음 단계에서 각 YUV 채널별로 임계치를 이용하여 에지와 에지가 아닌 픽셀을 구분하여 에지맵을 생성한다. 각각의 채널별로 각 임계치에 대하여 다음의 함수를 이용하여 에지맵을 생성하다.

Y 채널의 에지판정 함수는 식 (3-11)과 같다. 이때 Ty는 임계값이다.
U채널과 V채널의 에지판정 함수도 식 (3-11)과 유사하다.

$$YE(x, y) = \begin{cases} 1, edgw\,\pi xel, \text{if}\, MOE(x,y) \le Ty \\ 0, nonedge\,\pi xel, \text{if}\, MOE(x,y) < Ty \end{cases} \quad (3\text{-}11)$$

앞에서 서술한 방법을 이용하여 각각의 Y,U,V 채널에 대하여 에지맵을 생성한 후 최종 퓨젼 시키는 방법은 식 (3-12)과 같다.

$$\begin{cases} 1.\text{if}(Y_E(x,y) = 1\, U_E(\frac{x}{2}, \frac{y}{2}) = 1)\, U(V_E((\frac{x}{2}, \frac{y}{2}) = 1 \\ 0, otherwise \end{cases} \quad (3\text{-}12)$$

5. 다차원 구배도 방법(*multiple dimensional method*)

다차원 구배도 방법은 세 가지 채널로부터 하나의 구배도를 생성해 내는 방법으로 그림 3-3과 같은 처리과정을 갖는다.

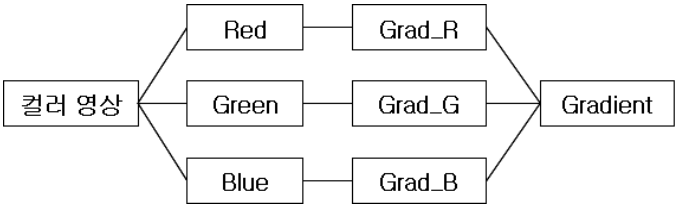


그림 3-3 다차원 구배도 방법의 처리과정

컬러 영상의 전체적인 구배도를 구하기 위하여 세 개의 채널에서 구한 채널별 구배도를 혼합하는 방법으로 Lee와 Cok은 컬러 영상의 3가지 채널의 구배도 G(i, j)의 크기를 계산하였다.

컬러 영상 y를 벡터 필드 y=[y1, y2, y3]라 한다. 이때, y1, y2, y3는 컬러 영상의 각 채널을 의미한다. 벡터필드 y의 공간 구배도(spatial gradient) 행렬 D 는 식 (3-13)과 같이 정의될 수 있다.

$$D(i,j) = \begin{bmatrix} \frac{\partial y1}{\partial u}, \frac{\partial y1}{\partial v} \\ \frac{\partial y2}{\partial u}, \frac{\partial y2}{\partial v} \\ \frac{\partial y3}{\partial u}, \frac{\partial y3}{\partial v} \end{bmatrix} \tag{3-13}$$

여기서 u와 v는 픽셀 (i, j)의 공간좌표이다.

$$D^T D(i,j) = \begin{pmatrix} p(i,j) l(i,j) \\ l(i,j) q(i,j) \end{pmatrix} \tag{3-14}$$

의 가장 큰 고유값(eigen value)의 제곱근과 그에 대응하는 고유벡터(eigen vector) 는 스칼라 필드에서의 구배도의 크기와 방향과 같은 의미를 갖게 된다. $D^T D$의 최대 고유값 λ는 식 (3-15) 와 같이 표현될 수 있다.

$$\lambda = \frac{1}{2}\, p(i,j) + q(i,j) + \sqrt{p(i,j) + q(i,j)^2 - 4(p(i,j)q(i,j) - t(i,j)^2))} \qquad (3\text{-}15)$$

여기에서 $p(i,j) = \left(\frac{\delta y1}{\delta u}\right)^2 + \left(\frac{\delta y2}{\delta u}\right)^2 + \left(\frac{\delta y3}{\delta u}\right)^2$

$$p(i,j) = \left(\frac{\delta y1}{\delta v}\right)^2 + \left(\frac{\delta y2}{\delta v}\right)^2 + \left(\frac{\delta y3}{\delta v}\right)^2$$

$$p(i,j) = \left(\frac{\delta y1}{\delta u}\right)\left(\frac{\delta y1}{\delta v}\right) + \left(\frac{\delta y2}{\delta u}\right)\left(\frac{\delta y2}{\delta v}\right) + \left(\frac{\delta y3}{\delta u}\right)\left(\frac{\delta y3}{\delta v}\right)$$

이다.

결과적으로 각 채널의 혼합된 구배도의 최대크기는 G(i, j) = $\sqrt{\lambda(i,j)}$ 으로 나타낼 수 있다. 이러한 다차원 구배도 방법을 이용하여 컬러 에지를 추출하기 위하여 여러 가지의 마스크가 사용될 수 있다. 보다 정확한 컬러 에지를 추출하기 위하여 다차원 구배도 방법을 이용할 경우 구배도를 구하기 위한 마스크로 컴패스 연산자를 이용한 방법과 Canny 연산자를 이용한 방법이 있다.

6. 벡터적 방법(vector method)

이 방법은 영상의 픽셀 값들을 벡터로 간주하여 처리하고자 하는 접근방법이다 여기에 포함되는 방법으로는 영상의 픽셀에 대한 고유값을 구하여 에지를 추출 하려는 방법과 통계학에서의 순서통계량(ordering statistics) 방법을 적용하는 벡터 순서통계량 방법(vector order statistics method)이 있다

벡터 순서통계량 방법은 통계학에서 제안하는 순서(ordering) 스키마를 사용하여 컬러 에지를 구하고자 하는 접근 방법이다.

크기가 N인 확률 표본 X_1, X_2, ···, X_N을 크기순으로 나열한 후 가장 작은 값을 $X_{(1)}$, 그 다음 작은 값을 $X_{(2)}$ ··· 이렇게 지정하며 $X_1 < X_2 <$ ···, X_N가 성립한다. 이때, X_1, X_2, ···, X_N을 순서통계량 이라 한다. R-순서는 순서통계량의 한 종류로 특정 벡터로부터의 거리 값 di를 이용하여 다른 벡터들의 값을 순서대로 나열 한 후 결과적으로 다중 순서량으로 표현된 것을 일차원 순서로 줄이는 방법이다. 이 방법을 그대로 컬러 영상에 사용하게 되면 컬러 영상이 갖는 세 가지 채널에 대한 것이 결과적으로 하나의 채널 값으로 결합되어 나타나게 되는 것이다.

R-순서 스키마를 사용하는 벡터 순수통계량 방법으로 접근할 때 컬러 영상을 픽셀 수만큼 가로, 세로 n×n 윈도우를 갖고 이때, 각 픽셀들은 컬러 벡터로 간주된다. 따라서 영상의 픽셀을 C_i, I = 1, 2, ..., N;N = n×n로 표현할 수 있다. R-순서 스키마에서, 각 벡터 C_i는 거리 표준(distance criterion)에 의하여 스칼라 값인 di로 축소된다. 거리 행렬에 의하여 각 벡터 집합인 C_1, C_2, ···, C_N의 총 거리 d_i는 식 (3-16)과 같이 나타낼 수 있다.

$$d_i = \sum | C_i - C_k | \tag{3-16}$$

오름차순으로 di의 값을 나열하면 식 (3-17)과 같다.

$$d_{(1)} < d_{(2)} < \ldots < d_{(N)} \tag{3-17}$$

위의 표현은 벡터 C_i에 대하여 다음과 같이 나열된 것과 같은 의미를 지닌다.

$$C_{(1)} < C_{(2)} < \ldots < C_{(N)} \tag{3-18}$$

이것을 이용하여 컬러 에지 검출기를 정의하며 식 (3-19)과 같다.

$$VR = | C_{(1)} - C_{(N)} | \tag{3-19}$$

위에 표현된 VR(vector range) 식은 벡터들의 중간값 밖으로 놓여 있는 벡터들의 편차를 나타낸다. 이들을 정규화 시키게 되면 모든 벡터들은 서로 보다 밀집되게 되고, VR의 결과 값은 점점 작아지게 된다.

벡터로 간주되는 픽셀이 에지의 한쪽에 있을 때를 $C_{(1)}$이라고 한다면 $C_{(N)}$은 에지의 반대쪽에 있게 된다. 이것이 VR이 되는 것이다. 벡터 범위결과에 임계치를 적용하여 최종적으로 실제 에지를 얻게 된다.

반면에 VR은 갑작스런 잡음에 매우 민감하므로 하나의 잘못된 잡음 픽셀로 인하여도 잘못된 에지가 검출될 수 있기 때문에 높은 쪽에 위치한 벡터들은 잡음으로 간주된다. 에지 검출 시에는 노이즈가 잘못 간주되는 것을 최소화하기 위하여

단일차(single difference)와 최소화연산자(minimization operation) 대신에 차집합 k를 이용한다.

$$\min(|C(n-j+1)-C(1)|) \quad \text{for} \quad j=1,2,...,k \quad k<N \tag{3-20}$$

최소 벡터 분산(minimum vector disperion)의 결과로 에지를 검출할 수 있으며 식 (3-21)과 같다.

$$MVD = \min\left(\left|C(N-j+1)-\frac{\sum_{i=1}^{L}C(i)}{L}\right|\right) \tag{3-21}$$

$$\text{or } j=1,2,\ldots,k \quad L<N$$

최소 벡터 분산 방법을 이용하여 에지를 구하게 되는 경우 최소화 연산자를 사용하는 경우 길게 이어지는 잡음(long tailed noise)에 약하고 평균연산자를 사용하는 경우보다 전체적으로 잡음에 강한 결과를 가져온다.

B. 에지 기반 영상 검색

물체의 형태 정보는 시각적으로 영상의 매우 다양한 정보를 가지고 있으며 형태 정보 기반 검색 기법은 구현하기 어렵지만 형태 정보가 물체를 가장 잘 표현하므로 내용 기반 검색 시스템에서 가장 중요한 부분이다. 형태 표현 기법은 크게 2가지 부류로 나뉘어질 수 있다[22]. 하나는 윤곽선 기반(boundary-based)기법이고 다른 하나는 영역 기반(region-based)기법이다. 윤곽선 기반 기법은 영상내의 물체의 외곽선에 기반하여 특징을 추출하는 방법으로서 체인 코드(chain code), 퓨리어 묘사(Fourier descriptor), UNL 퓨리어 특징, 물체의 윤곽선의 반경 특징점을 이용한 기법 등 다양한 기법들이 존재한다. 영역 기반 기법은 물체 영역의 기하학적인 특징을 추출하는 방법으로서 면적(area), 신장율(elongation), 오일러 수(Euler number), 불변 모멘트(invariant moments), 저나이크 다항식을 이용한 저나이크 모

멘트(Zernike moments)를 비롯하여 많은 방법들이 있다. 본 연구에서는 대표적인 형태 기반 검색 기법인 Jain 등의 에지의 방향 히스토그램을 이용한 방법[30], 저나이크 모멘트를 이용한 방법[21] 및 지역적인 미분 불변치(local differential invariants)를 이용한 방법들에 대해 알아본다[27][28].

1. 에지의 방향 히스토그램을 이용한 기법

Jain에 의해 제안된 방법으로 영상 내의 물체의 에지의 방향 히스토그램을 형태 정보로 이용하였다. 즉, 에지의 방향 히스토그램을 구한 후 히스토그램 인터섹션 방법을 이용하여 각 영상의 히스토그램을 비교하였다[30].

먼저 Canny 에지 연산자[29]를 이용하여 에지의 위치를 찾은 후 에지 위치에서 그레디언트의 방향을 구하면 에지의 방향 정보를 얻어낼 수 있다. 즉, 픽셀 (m,n)에서 가로 수평 방향으로의 변화량을 $g_h(m,n)$이라 하고, 수직 방향으로의 변화량을 $g_v(m,n)$라고 하면 방향은 식 (3-22)을 이용하여 계산할 수 있다.

$$\theta = \tan^{-1}\frac{g_v(m,n)}{g_h(m,n)} \tag{3-22}$$

이렇게 에지 위치에서 얻어진 방향들을 이용하여 히스토그램을 생성한 후 앞 절의 컬러 히스토그램 인터섹션을 이용하여 유사도를 비교한다. 하지만 에지의 방향을 히스토그램화 하는 것은 영상의 이동에는 불변하지만 축척이나 회전에는 변화한다는 단점을 가지고 있다. 따라서 Jain은 이러한 다음과 같은 방법을 이용하였다. 축척에 대한 문제는 영상내의 총 에지 수를 이용하여 히스토그램을 정규화 함으로써 해결할 수 있다. 하지만 회전에 대한 변화는 축척에 대한 것보다 다소 어렵다. 물체가 회전하면 에지의 방향 히스토그램에서는 이동(shift)이 일어나게 된다. 하지만 다음과 같은 문제가 발생한다. 즉, 에지 방향의 히스토그램의 그룹이 45。의 각도로 나뉘어져 있을 경우, 30。, 40。의 에지는 같은 그룹에 포함되게 된다. 그러나 10。의 회전이 이루어질 경우 각기 서로 다른 그룹에 속하게 됨으로써 원래의 히스토그램과는 다른 모양의 히스토그램이 만들어지게 된다. 이러한 문제를 해결하기 위해서 히스토그램에 대해 평활화(smoothing) 연산을 수행하였다. 하지 만, 히스토

그램을 매끄럽게 만들어도 급격한 회전에는 적용되지 못하는 단점을 가지고 있으므로, 이동 축척의 변화에는 어느 정도 좋은 결과가 나왔지만, 회전의 변화에는 좋지 못한 결과를 보였다.

2. 저나이크 모멘트를 이용한 기법

저나이크 모멘트는 단위원의 내부에서 완전한 직교집합을 형성하는 복소 저나이크 다항식으로부터 유도되며 저나이크 모멘트를 이용한 영상의 인식에 대한 연구가 Khontanzad에 의해 제안되었다[21]. 저나이크 다항식의 집합은 식 (3-23)의 $V_{nm}(x,y)$에 의해서 나타내어진다.

$$V_{nm}(x,y) = V_{nm}(\rho,\theta) = R_{nm}(\rho)\exp(jm\theta) \tag{3-23}$$

여기서,

n: 0과 양의 정수

m: $|m| \le n, n-|m| = even$ 조건을 만족하는 정수

ρ, θ: 중심까지의 거리와 시계 반대 방향의 각도

$R_{nm}(\rho)$: 방사형 다항식

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|} (-1)^s \frac{(n-s)!}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!} \rho^{n-2s} \tag{3-24}$$

디지털 영상 $f(x,y)$ 에 대한 n 차, 반복계수 m 의 저나이크 모멘트는 저나이크 다항식을 이용하여 식 (3-25)와 같이 정규 모멘트와 유사하게 정의된다.

$$A_{nm} = \frac{n+1}{\Pi} \sum_x \sum_y f(x,y) V^*_{nm}(x,y), \text{ 여기서 } x^2+y^2 \le 1 \tag{3-25}$$

각 영상에 대해서 이렇게 추출된 모멘트들을 비교함으로서 영상간의 형태의 유

사성을 비교할 수 있다.

저나이크 모멘트의 경우 회전에 의한 모멘트의 크기가 불변이고, 그것들로부터 영상의 재구성이 쉽다는 특징이 있다. 저나이크 모멘트를 적용하기 위해서는 반드시 크기와 이동에 대한 정규화 과정이 있어야 한다. 즉, 영상 내의 물체가 영역 $x^2+y^2 \le 1$ 내부에 위치하도록 조정되어야 한다. 저나이크 모멘트는 완전한 직교집합으로, 저나이크 모멘트를 이용하여 원 영상을 복원할 수 있으며 영상의 기하학적 변형(이동, 회전, 축척)과 조명변화에 매우 강인하고 영상 전체의 형태 정보를 효과적으로 표현할 수 있다는 장점이 있다. 그러나 영상 열화에 매우 민감하고 영상 내부의 세부 픽셀 값은 무시되는 약점이 있다.

3. 지역적인 미분 불변치를 이용한 기법

영상에서 많은 정보를 포함하고 있는 부분들 즉, 변화가 큰 부분들을 따로 추출하여 그 부분들에 대한 지역적인 특징을 검색에 이용한다. 이를 위해 영상에 있어서 특징점으로서 인터레스트 포인트(interest point)를 정의하고 그 특징점에 대해 미분 불변치를 구함으로서 지역적인 특징을 구하고 각 영상간의 유사한 지역적인 특징을 가지는 특징점들을 비교함으로써 유사성을 측정할 수 있다.

(1) 인터레스트 포인트 추출

인터레스트 포인트는 영상에서 신호가 2차원적으로 변하는 위치이며, 코너 포인트, T 교차점, 텍스쳐가 급격히 변하는 위치 등이 이에 해당된다. 인터레스트 포인트는 회전, 이동, 축척, 영상 열화, 조명 변화, 관점(viewpoint) 변화 등에도 기하학적으로 안정적이며, 높은 정보량을 가지므로 영상 검색에 효과적으로 이용된다 [23]-[25].

인터레스트 포인트의 성능은 검출기의 성능에 따라 좌우되며 검출기의 성능은 다양한 영상변화에도 반복적인 검출결과를 나타내는 반복성과, 랜덤 변수의 엔트로피 즉, 분별성을 나타내는 정보량이 비교기준이 된다. 현존하는 검출기는 Moravec, Harris, Heitge, Forstner, Horaud, Cottie 등에 의해 고안되었는데, 이중 반복성과 정보량 측면에서 Harris 검출기가 가장 우수함이 판명되었다[24].

Harris 검출기는 자기 상관함수와 관련된 행렬을 구하고, 자기 상관함수의 중요 곡률(curvature)이 되는 이 행렬의 고유값(eigenvalue)을 척도로 인터레스트 포인트를 검출한다. Harris 검출기는 Moravec 검출기의 문제점을 개선시킨 검출기이다. Moravec 검출기는 지역 윈도우 영상을 다양한 방향으로 이동하였을 때 영상 세기의 변화량 ($E_{x,y}$)을 기준으로 영상의 영역을 분류한다.

$$E_{x,y} = \sum_u \sum_v W_{u,v}(I_{u+x,v+y} - I_{u,v})^2$$

여기서, W: 단위 직각 윈도우 (3-26)

x, y: 이동($shift$), u, v: 윈도우 영역

윈도우 내의 영상 세기가 대략적으로 비슷한 경우 영상 세기의 변화량은 작다. 윈도우가 에지 성분에 걸쳐 있다면 에지 성분을 따라 이동할 경우 영상 세기의 변화는 적지만, 수직으로 이동시에는 변화량이 크게 된다. 윈도우 영역이 코너 포인트 등의 인터레스트 포인트에 해당하는 영역은 영상세기의 변화량은 항상 크게 된다. Moravec의 검출기는 변위가 $\{(1,0),(1,1),(0,1),(-1,1)\}$일 때의 변화량의 최소값이 문턱값 이상이고, 지역의 최대값일 때 인터레스트 포인트로 판별한다.

Harris 검출기는 Moravec 검출기의 세 가지 문제점을 개선하였다. Moravec 검출기는 단위 직각 윈도우를 사용해서 잡음에 민감하였으나 Harris는 원형인 가우시안 윈도우를 사용하여 잡음문제를 해결하였고, 윈도우의 이동이 45。마다 이루어지므로 변화량 결과가 이방성(anisotropic)이라는 문제를 해결하기 위해 모든 방향으로의 이동을 고려해 해석적인 팽창(analytic expansion)을 적용하였고, 검출 성능 향상을 위해 자기 상관함수 행렬 M 의 고유값을 이용하였다.

$$E_{x,y} = \sum_u \sum_v W_{u,v}(I_{u+x,v+y} - I_{u,v})^2 = \sum_u \sum_v W_{u,v}(xX + yY + O(x^2+y^2))^2 \quad (3\text{-}27)$$

윈도우 영역의 작은 이동에서 영상 세기의 변화량은 식 (3-28)과 같다.

$$E(x,y) = Ax^2 + 2Cxy + By^2 = (x,y)M(x,y)^T \quad (3\text{-}28)$$

여기서, $M = \begin{bmatrix} A & C \\ C & B \end{bmatrix}$

$$W_{u,v} = \exp\{-(u^2+v^2)/w\sigma^2\}$$
$$X = I \otimes (-1,0,1) = \partial I / \partial x$$
$$Y = I \otimes (-1,0,1)^T = \partial I / \partial y$$
$$A = X^2 \otimes W$$
$$B = Y^2 \otimes W$$
$$C = (XY) \otimes W$$

M 은 회전에 불변하며, x, y의 변화 시 원점에서 분포상태의 척도가 된다. M 의 고유값 α, β는 자기 상관함수의 중요 곡률에 비례하므로 이를 기준으로 영상에서의 영역을 분류한다. 두개의 고유값이 모두 작으면 윈도우 영역은 거의 동일한 픽셀값을 갖고, 고유값의 하나는 크고 하나는 작으면 윈도우 영역은 에지 성분에 해당되고, 고유값이 모두 크면 윈도우 영역은 인터레스트 포인트에 해당된다. 자기상관함수 매트릭스의 고유값은 다음 식들과 같은 성질을 가진다.

$$Tr(M) = \alpha + \beta = A + B \tag{3-29}$$
$$Det(M) = \alpha\beta = AB - C^2 \tag{3-30}$$
$$R = Det - kTr^2 \tag{3-31}$$

응답함수 R 은 A, B, C의 함수이며 k 값은 임의의 작은 양수이다. R 값을 척도로 R 값이 양수면 인터레스트 포인트, 음수면 에지 영역, 작으면 유사한 픽셀값을 갖는 영역으로 판별한다[24].

(2) 특징점에 대한 미분 불변치 특징 벡터 추출

영상으로부터 추출된 특징점에서 회전, 이동, 크기 변형에도 강인하게 영상 검색에 적용할 수 있는 미분 불변치 벡터는 다음과 같이 계산할 수 있다. 미분 불변치 계산은 원영상에 작은 크기의 잡음이 첨가된 경우 원영상과 잡음이 첨가된 영상의 미분치를 비교하면 매우 다르다. 심지어 작은 잡음이 첨가된 경우라도 고주파 잡음

을 포함하는 경우 미분치는 상당히 달라지는데, 고차항의 미분일수록 값은 더욱 상이하게 된다. 원신호를 $f(x)$라 하고 잡음이 첨가된 신호를 $h(x)$라 할 때,

$$h(x) = f(x) + \epsilon \sin(wx) \tag{3-32}$$

$$h(x)' = f(x)' + w \cdot \epsilon \cos(wx) \tag{3-33}$$

$$h(x)'' = f(x)'' \cdot w^2 \epsilon \sin(wx) \tag{3-34}$$

고주파 잡음인 경우 w 값이 매우 크므로 $h(x)'$와 $f(x)'$는 매우 상이하며, 고차항 미분일수록 더욱 상이하게 된다. 그러므로 원신호를 미분하기 전에 신호를 평활화(smoothing)를 할 필요가 있다. 평활화 함수로 가우시안 함수 G 를 이용하여 원신호로부터 잡음을 제거한 다음 미분을 수행한다. 또한 미분 연산과 콘볼루션은 교환법칙이 성립하므로 식 (3-35)와 같이 나타낼 수 있다.

$$\partial_i (G * F) = G * \partial_i F = \partial_i G * F \tag{3-35}$$

미분 계산을 안정시키는 가장 간단한 방법은 평활화 함수인 가우시안 함수를 미분한 것과 영상을 콘볼루션시키는 것이다[27].

가우시안 함수 $G(x, \sigma)$, 가우시안 함수의 미분치 $G_{i_1, i_2, \ldots, i_D}(x, \sigma)$ 일 때,

$$G(x, \sigma) = \frac{1}{(\sqrt{2\Pi\sigma^2})^D} exp(-\frac{x^2}{2\sigma^2}) \tag{3-36}$$

$$G_{i_1, i_2, \ldots, i_D}(x, \sigma) = \frac{\partial^D}{\partial_{i_1}, \partial_{i_2}, \ldots, \partial_{i_D}} G(x, \sigma) \tag{3-37}$$

여기서, $x = (x_1, x_2, \ldots, x_D)$ (영상에서는 $D = 2$)이며, 가우시안 함수내의 σ는 평활화의 정도를 결정한다. N 차까지 미분치를 계산하여 벡터로 표시할 수 있는데 이는 Koenderink에 의해 처음으로 로컬젯(Local Jet)이라는 벡터로 표현되었다[27].

영상의 특징점을 중심으로 하는 지역 윈도우 영상에 대한 특징을 미분치들의 집합으로 나타낼 수 있다. 미분치를 구할 때 가우시안 미분치와 지역 윈도우 영상의 콘볼루션에 의해 안정적으로 구할 수 있고 이러한 미분치 집합을 로컬젯이라 한다

[27]. 지역 윈도우 영상을 I 라고 하고, 주어진 크기가 일 때, 임의의 한 점 $x=(x_1, x_2)$ 에서 N 차 로컬젯은 식 (3-38)과 같다.

$$J^N[I](x,\sigma)=\left\{L_{i_1,i_2,\ldots,i_n}(x,\sigma)\,(x,\sigma)\in I*R^+;n=0,1,2,\ldots,N\right\} \tag{3-38}$$

σ는 가우시안 함수의 크기에 해당하며 영상의 크기가 변하는 경우 영상 검색에 적용할 때 중요하게 사용된다. 미분 불변치는 로컬젯으로부터 계산되어지며, 이러한 특성 벡터로 영상 신호를 표현할 수 있다. 특성 벡터는 여러 가지 변형에서도 강인하게 특징값을 유지한다. 2차 미분항까지 고려한 미분 불변치 벡터 DI 를 Einstein 합방식과 직교좌표계에서 표현하면 식 (3-39)와 같다.

$$DI[0,\ldots,4]=\begin{bmatrix} L \\ L_iL_i \\ L_i \\ L_iL_{ij}L_i \\ L_{ij}L_{ji} \end{bmatrix}=\begin{bmatrix} L \\ L_xL_x+L_yL_y \\ L_{xx}L_xL_x+2L_{xy}L_xL_y+L_{yy}L_yL_y \\ L_{xx}+L_{yy} \\ L_{xx}+L_{yy}+2L_{xy}L_{yx}+L_{yy}L_{yy} \end{bmatrix} \tag{3-39}$$

L_i 는 위에서 표시한 로컬젯의 성분이다. 즉, L 은 지역 윈도우 영상과 가우시안 함수를 콘볼루션한 성분이다. 다른 크기의 영상에 대해서도 적용하기 위해서 가우시안 함수의 크기를 변형하여 미분 불변치 벡터를 구할 수 있다. 영상의 경우 벡터 $x=(x,y)$ 로 2차원이며, L_{xy} 에서 첨자 x,y 는 x,y 방향으로 각각의 편미분에 해당된다. Einstein 합방식에서 첨자 i 는 변수들에 대한 미분의 합을 나타낸다. L_i 와 L_{ij} 는 식 (3-40), (3-41)과 같이 표현될 수 있다.

$$L_i=\sum_i L_i=L_x+L_y \tag{3-40}$$

$$L_{ij}=\sum_i\sum_j L_{ij}=L_{xx}+L_{xy}+L_{yx}+L_{yy} \tag{3-41}$$

3 차항의 미분을 고려한 경우의 미분 불변치 벡터 DI 를 Einstein 합방식으로 표현하면 식 (3-42)와 같다.

$$DI[5...8] = \begin{bmatrix} \epsilon_{ij}(L_{jkl}L_iL_kL_l - L_{jkk}L_iL_lL_l) \\ (L_{iij}L_jL_kL_k - L_{ijk}L_iL_jL_k) \\ -\epsilon_{ij}L_{jkl}L_iL_kL_l \\ L_{ijk}L_iL_jL_k \end{bmatrix} \tag{3-42}$$

$$(\text{단}, \epsilon_{12} = -\epsilon_{12}, \epsilon_{11} = -\epsilon_{22} = 0)$$

미분 불변치 벡터 DI 를 직교 좌표계에서 표현한 식을 간략히 하면 식 (3-43)과 같다.

$$\begin{aligned} DI[5] &= L_{xxx}L_yL_yL_y - L_{yyy}L_xL_xL_x + 3L_{xyy}L_xL_xL_y - 3L_{xxy}L_xL_yL_y \\ DI[6] &= L_{xxx}L_xL_yL_y + L_{yyy}L_xL_xL_y + L_{xxy}L_y(L_yL_y - 2L_xL_x) + L_{xyy}L_x(L_xL_x - 2L_yL_y) \\ DI[7] &= L_{xxx}L_xL_xL_y - L_{yyy}L_xL_yL_x + L_{xxy}L_x(2L_yL_y - L_xL_x) + L_{xyy}L_y(L_yL_y - 2L_xL_x) \\ DI[8] &= L_{xxx}L_xL_xL_x + L_{yyy}L_yL_yL_y + 3L_xL_y(L_{xxy}L_x + L_{xyy}L_y) \end{aligned} \tag{3-43}$$

위와 같이 영상들에 대해 Harris의 코너점 추출기를 이용하여 특징점을 추출한 후 각각의 특징점에 대해 미분 불변치를 계산했을 때, 각 영상간의 유사도 비교는 각각의 특징점간의 거리를 비교함으로써 수행된다. 비교되는 특징점간의 미분 불변치 특징 벡터의 거리가 미리 정한 문턱값을 넘지 않을 경우 두 특징점을 동일한 포인트로 간주하며, 두 영상간의 일치하는 특징점의 수를 이용하여 두 영상의 유사도를 계산한다. 유사도 계산 방법을 정리하면 다음과 같이 나타낼 수 있다. N 개의 특징점을 가지고 있는 두 영상 I_1, I_2 에 대해 I_1 내의 특징점의 집합을 $\{p_i\}, i = 1, 2, ..., N$, I_2 내의 특징점의 집합을 $\{p_j\}, j = 1, 2, ..., N$ 라 할 때 식 (3-44)와 같이 계산할 수 있다[26].

$$\begin{aligned} &\text{for } \ i = 1, 2, ..., N \\ &\text{if}(d_{\min}(I_1(p_i), I_2(p_j), \text{for } \ j = 1, 2, ..., N) < Threshold) \\ &\qquad count = count + 1 \\ &\text{유사도} = count/N \end{aligned} \tag{3-44}$$

유사도가 클수록 두 영상간의 형태 정보가 유사함을 나타낸다.

Ⅳ. 컬러와 에지 정제 기반 *CBIR* 설계

A. 기존 컬러/에지 알고리즘 방법

제안한 방법의 결과를 분석해 보기 위하여 두 가지 방법으로 에지 히스토그램을 생성하여 비교하고자 한다. 먼저 앞에서 언급했던 MPEG-7 에지 히스토그램 서술자를 이용하여 실험해 보고 제안한 방법 과 같이 컬러 영상으로부터 에지를 추출 후 결합하는 컬러 에지 추출 방법에 의하여 생성된 영상을 이용해 보고자 한다.

MPEG-7 에지 히스토그램 방법은 이미 앞에서 자세히 설명하였으므로 여기서는 생략하기로 하고 컬러 에지를 추출하여 히스토그램을 생성하는 방법에 대하여 먼저 간단히 설명하기로 한다.

컬러 에지를 추출하여 에지 히스토그램을 생성한 방법으로 기존의 컬러/에지 방법을 이용한 PicSOM시스템이 있다. 이 시스템에서는 에지 히스토그램을 구하기 위하여 컬러 영상을 HSI 컬러 공간으로 변환한다. 색상채널을 제외한 나머지 두 채널에 대하여 소벨 연산자를 적용하여 에지 맵을 생성한다. 명도 채널에 대하여는 15%의 임계치를, 채도 채널에 대하여는 35%의 임계치를 적용시켜 에지맵을 이진 영상으로 변환 한다.

두 개의 이진영상에 대하여 논리합 연산을 적용하여 구배도가 큰 쪽의 값이 갖는 방향값을 선택하고 최종적으로 8개의 빈으로 구성된 히스토그램을 생성 한다. 이때 논리합 연산을 적용하여 구배도가 큰 쪽의 값을 선택하면 최종적으로 두 개의 채널에 대한 에지 성분이 결합된 에지가 생성된 것을 볼 수 있다. 이 결합된 에지맵에 대하여 방향 성분을 구한 후 25° 간격으로 총 8개의 빈을 갖는 히스토그램을 생성 한다.

이 방법 이외에도 Anil.K.Jain과 A.Vailaya에 의하여 제안된 또 다른 에지 히스토그램으로 다음과 같은 방법이 있다. Canny오퍼레이터(σ = 1 그리고 가우시안 마스크 사이즈 = 9)를 사용하여 영상에 대한 에지를 추출한 후 에지 방향을 5° 간격으로 양자화 시켜서 전체 72개 빈으로 구성되는 에지 히스토그램을 생성하는 방법이다.

본 실험에서 제안한 방법과의 결과를 비교 분석하기 위한 기존의 컬러 정제 또는 에지 정제 기반 영상 검색 알고리즘은 그림 4-1 및 그림 4-2와 같다.

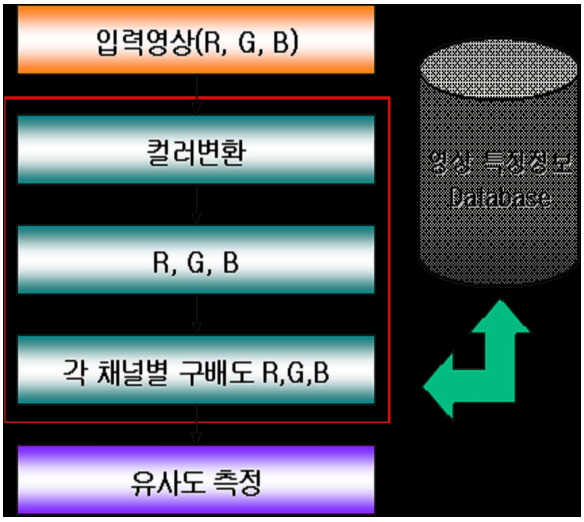


그림 4-1 기존 컬러 정제 기반 영상 검색 알고리즘

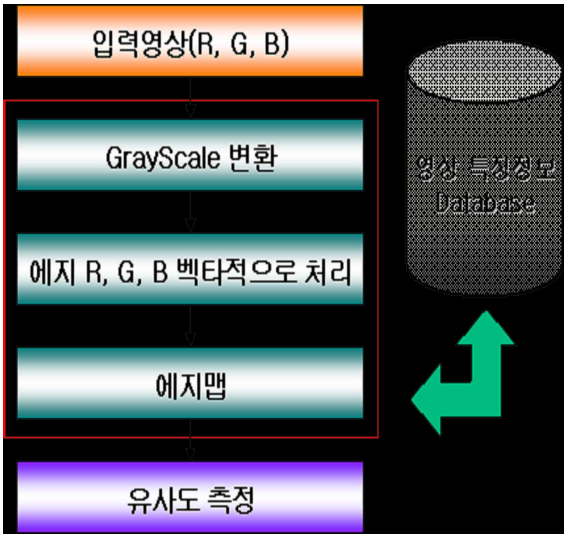


그림 4-2 기존 에지 정제 기반 영상 검색 알고리즘

B. 컬러와 에지 정제 기반 영상 검색 알고리즘

제안한 방법은 위에서 살펴보았던 영상 퓨젼 기법을 개선한 형태이다. 기존의 컬러/에지 방법에서는 채널별 구배도를 이용하여 채널별 에지맵을 생성한 후 이 세가지 에지맵에서 다섯 가지 방법(픽셀수, 클러스터수, 클러스터 평균수, 직선에지, 기울어진 에지수 정제)을 이용하여 하나의 에지맵으로 합쳤으나 제안한 방법은 각 채널별로 에지맵을 구할 때 단순한 미분 연산자들을 이용하여 구배도로만 구한 것과는 달리 기존의 에지를 계산한 후 에지 발생 확률 값을 픽셀에 대하여 실제 픽셀이 갖고 있는 컬러 값을 그대로 출력하도록 하였다. 따라서 출력 영상은 이진화 영상이 아니다.

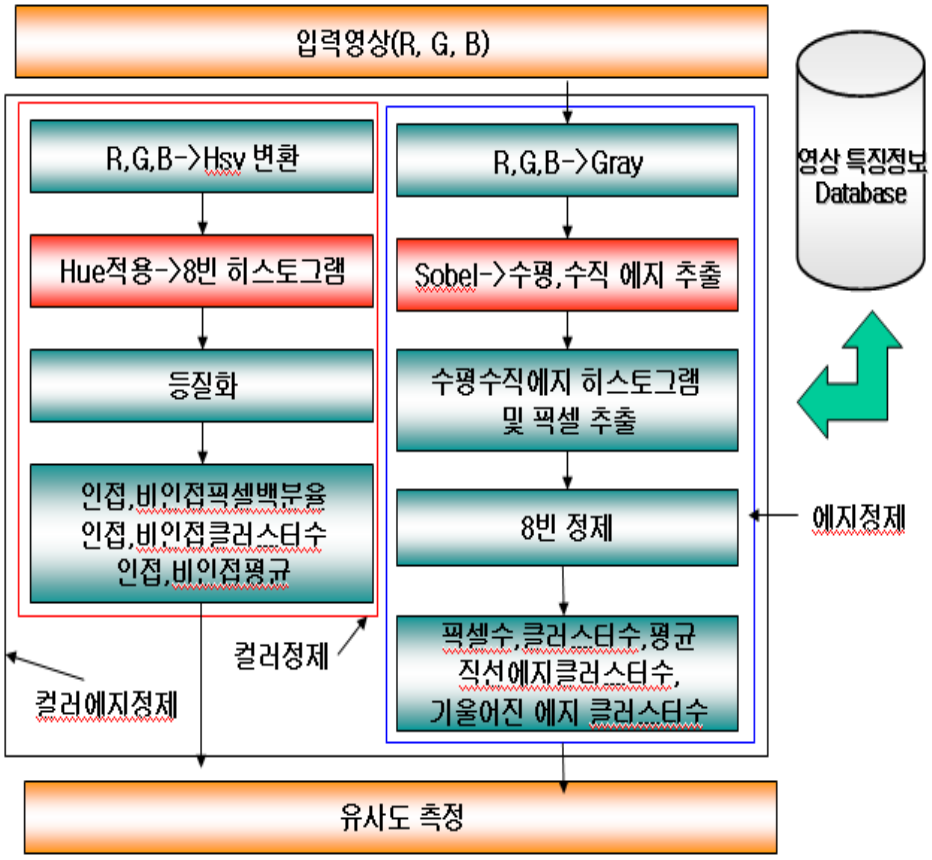
본 논문에서는 컬러 영상의 에지 히스토그램을 생성하여서 기존의 에지 정제 방법에서 사용하는 방법을 개선한 컬러 에지방법을 사용한다.

새로운 컬러 에지를 정제하기 위해서는 먼저 컬러 영상을 기존의 연구에서 처럼 R, G, B 세 개로 분류하고 기존의 연구와 다른 방법을 적용한다. 이렇게 얻어진 영상을 히스토그램 개선방법으로 주어진 버켓 안에 픽셀을 국부적인 특징에 기초하여 여러 분류로 나누고 이것을 전처리를 위해 세 가지 다른 방법이 사용된다.

첫 번째는 HSV공간으로 변환하고 R, G, B, 256 : 256 : 256에서 HSV 8 : 2 : 2를 얻기 위해 양자화를 한다. 색상 값만 고려하여 히스토그램을 얻는다.

두 번째는 그레이 스케일의 밀도 이미지로 변환 한다. 이같이 8개 양자화 값으로 양자화 한다. 그 다음 히스토그램을 얻는다. 그리고 컬러 영상의 에지정제는 수평과 수직을 찾기 위해 소벨연산자를 적용하고, 8개의 방향에서 에지를 찾기 위해서는 컴패스 연산자를 적용한다. 소벨연산자는 특징벡터 크기를 줄이기 위한 전처리로 간주된다.

그림 4-3은 제안한 컬러 및 에지 정제에 기반한 영상 검색 알고리즘의 순서도이다.



4-3 제안한 컬러 및 에지 정제 기반 영상 검색 알고리즘

1. 컬러 특성 벡터

컬러 히스토그램 버킷은 Pass와 Zabih[3]에 의해 계산된 것처럼 공간적 결합성에 의해 나누어진다. 픽셀은 어떤 규모 있는 컬러 영역의 일부라면 인접성이 있는 것으로, 그렇지 않으면 인접성이 없는 것이다. 그래서 픽셀은 각각의 컬러 버킷 안에서 인접성 또는 비인접성으로 나누어진다.

만약 픽셀이 이미지의 적어도 1%를 형성하는 같은 색의 큰 그룹의 픽셀의 일부라면, 그 픽셀은 인접 픽셀이 되고, 그 그룹은 인접그룹 또는 클러스터라 불린다. 그렇지 않으면 비인접 픽셀이 되고, 그 그룹은 비인접 그룹 또는 클러스터가 되는 것이다. 그리고 나서, 각각의 빈에서 두 영역이 더 계산되어 진다.

첫 번째는 각각의 경우 (즉 각 빈 내에서의 인접과 비인접의 경우) 클러스터의 수를 알게 된다.

두 번째는 각 클러스터의 평균이 계산된다. 그래서 각 빈에서 6개의 값을 갖게 된다. 인접 픽셀과 비인접 픽셀의 백분율, 인접 클러스터와 비인접 클러스터의 수, 인접 클러스터와 비인접 클러스터의 평균이 값들은 연결된 구성 요소들을 추정함으로서 계산되어진다. 연결된 구성요소 C는 어떤 두 픽셀이 P, P′∈ C, C내에서 P와 P′ 사이에 통로가 있을 때 픽셀의 최대 세트인 것이다. 8개의 연결된 이웃한 방법이 연결된 구성 요소들을 추정하는 데 사용된다. 하나의 픽셀은 τ(τ는 이미지 크기의 1%)와 같거나 큰 크기의 연결된 구성 요소의 일부라면 인접하다고 분류된다. 그렇지 않으면 비인접성으로 분류된다.

그리고 연결된 구성 요소는 τ와 같거나 초과될 때 인접한 연결 구성 요소로 분류된다. 마지막으로 인접과 비인접성 연결 구성 요소가 계산되어 진다. 각각의 Discretized color j에 대해, 인접 픽셀 수를 αj, 인접한 연결 구성 요소의 수를 Cαj, 그리고 인접한 연결 구성 요소의 평균을 μαj라 하자. 비슷하게, 비인접 픽셀 수를 βj, 비인접 연결 구성 요소의 수를 Cβj, 그리고 비인접 연결 구성 요소의 평균을 μβj라 하자. 각각의 Discretized color j에 대해, 픽셀 전체의 수는 αj + βj이고, 컬러 히스토그램은 <α1 + β1, · · ·, αn + βn>으로 이미지를 요약한다.

2. Post-processing

이미지 I와 I′를 비교하기 위해 거리 L1을 사용한다. 거리 L1을 사용하면, I와 I′ 간의 거리에 j번째 버킷의 기여도는

$$\triangle 1 = |(\alpha j \quad \alpha' j)| + |(\beta j \quad \beta' j)| \tag{4-1}$$

$$\triangle 2 = |(C\alpha j \quad C\alpha' j)| + |(C\beta j \quad C\beta' j)| \tag{4-2}$$

$$\triangle 3 = |(\mu\alpha j \quad \mu\alpha' j)| + |(\mu\beta j \quad \mu\beta' j)| \tag{4-3}$$

그래서, 우리는 이 방법을 통하여 개선된 차이점을 얻는 것이다.

최초의 안 [3]에서는 식 (4-1)만을 사용하였고, 비교를 위해서 다음 방정식이 사용되었다.

$$\triangle 1 = |(\alpha j + \alpha' j)| - |(\beta j + \beta' j)| \quad (4\text{-}4)$$

아울러 식 (4-1)에서 식 (4-3)은 측정 가능성을 구체화해 준다. 그리고 Huang et al[12]에 의해 [3]에서 정의된 CCV를 사용하는 것만으로는 제거할 수 없는 문제들을 제거할 수 있다. 에지 개선방법 으로서는 버켓은 전처리 단계에서 에지 방향에 기초하여 형성되고, 그 다음 각 버켓 마다 픽셀의 총수가 계산된다. 에지를 계산하기 위해 소벨연산자를 사용했고, 본 논문에서는 수평과 수직 에지만을 가지고 계산했다. 그러므로 앞에서 두개의 버켓이 만들어 진다. 여기서 버켓 안의 각 픽셀을 인접 또는 비인접으로 분류함으로서 또 다른 개선을 만들어냈다. 그런 다음 각 버켓 안에서 4개의 값을 더 계산해 내었다.

3. 에지 특성 벡터

에지 정제는 히스토그램의 정제법에 기반으로 하며. 먼저 선행 과정으로 두 가지의 다른 방법으로 사용할 수 있다

① 수평 에지, 수직 에지를 찾기 위해 소벨 연산자를 적용
② 이 기법은 특징벡터 사이즈를 감소시키기 위해 고려되어졌다.

첫째, 버켓들은 선행과정 단계에서 에지 방향에 기초를 두고 형성 된다. 그리고 각각의 버켓을 위해 총 픽셀의 수가 계산된다. 에지를 계산하기위해 소벨을 사용했다. 수직, 수평 에지 값만을 계산했다. 그러므로 이 경우 두 개의 버켓이 구성 되었다. 인접과 비인접 버켓으로 각각의 픽셀을 분류함으로써 다른 개선작업이 이루어졌다. 그리고 각각의 버켓에 추가로 4개의 값들이 계산되었다. 첫째로 8개의 인접 법칙에 의해 각각의 버켓에 클러스터의 개수가 도출 되었다.

두 번째로 각 버켓의 평균 클러스터의 값이 계산되었다. 그리고 픽셀의 수에 기초하여 클러스터, 직선 에지의 수와 기울어진 에지의 수가 계산되었다. 그리하여 각 빈은 5개의 값을 가진다. 각 빈으로 부터 계산된 픽셀의 총 수, 클러스터의 수, 평균 클러스터 값, 직선과 기울어진 에지의 수를 특징 값으로 한다.

마지막으로 이렇게 추출된 특징 값 들은 연관 구성요소를 계산함으로써 구해진다.

4. 8 빈 히스토그램 특성 벡터

영상 내용 특징화의 일반적인 방법은 컬러 히스토그램 방법이다. 영상에 대한 컬러 히스토그램은 각 컬러 픽셀 수의 계산으로 이루어진다. 대부분의 많은 응용들에서는 화상의 전체 외상적인 모양에 기초하여 두 개의 화상을 비교하는 단순한 방법을 요구하고 있다. 예를 들면, 사용자는 대규모 화상 데이터베이스로부터 주어진 화상과 유사한 모든 화상을 검색하기를 원한다. 그렇기 때문에 컬러 히스토그램은 이러한 문제에 대한 좋은 해결책으로 사용한다.

컬러 히스토그램은 계산이 효율적이고, 일반적으로 카메라의 위치 변화에 덜 민감하다. 컬러 히스토그램이 많이 사용되는 몇 가지 요인은 다음과 같다.

① 컬러 히스토그램은 계산 시간량이 적다.
② 컬러 히스토그램은 카메라 위치의 미세한 변화에 크게 영향을 받지 않는다.
③ 서로 다른 객체는 보통 구별되는 컬러 히스토그램을 갖는다.

다음과 같은 요인으로 컬러 히스토그램을 많이 사용하였지만, 컬러 히스토그램에 기반 한 검색은 여러 가지의 문제점이 있다. 첫 번째는 컬러 히스토그램은 높은 차수로 이루어져 있다. 컬러 공간의 양자화와 같이, 특징공간인 영상의 히스토그램은 영상의 공간에서 100차수이상을 차지할 수 있다. 이 높은 차수는 특징을 감소시키는 방법이고, pre-filtering과 계층적 인덱싱이 이행될 수 있게 한다. 또한 많은 차수는 복잡도와 거리함수를 증가시켜 계산량을 증가시킨다.

컬러 영상 히스토그램은 컬러 영상 명암의 가능성 밀집 함수로 간주 된다. 컬러 영상에 대해서 세 개의 컬러 채널 명암의 접속 가능성을 포착하게 확장된다. 공식적으로, 컬러 히스토그램은 식 (4-5)과 같이 정의 된다.

$$h_{r,g,b}(R,G,B) = N \cdot \mathrm{Prob}\{R=r,\ G=g,\ B=b\} \qquad (4\text{-}5)$$

R, G, B는 세 개의 채널을 표현하고, N은 영상내의 픽셀 수이다. 계산적으로, 컬러 히스토그램은 영상내의 분리된 컬러와 각 컬러 픽셀 수의 계산에 의해 만들어진다. 컬러 히스토그램은 영상내의 컬러를 분별하거나 영상 배열 내에 나타나는 컬러를 카운트함으로써 얻어 질 수 있다. 즉, 히스토그램이란 모든 가능한 컬러의 집합 내에서 상등함수로 정의 할 수 있다. 기존의 히스토그램 방식은 크게 두 가지로 나누어지는데, 영상 전체에 대한 전역적인 특징을 갖는 글로벌 히스토그램과 객체와 배경 사이의 경계선 사이에 존재하는 국부적 히스토그램으로 나뉜다. 이들을 구분 짓는 것은 전자의 경구, 사용되는 단위가 히스토그램이고 후자의 경우는 각 히스토그램간의 거리 값이다.

전역적으로 사용되는 히스토그램 방식은 누적히스토그램과 같은 형태로 나타나는데 3차원 컬러공간(x, y, z)에 누적되는 히스토그램을 표현하면 식 (4-6)과 같다.

$$C(x,y,z) = \sum_{i=1}^{x}\sum_{j=1}^{y}\sum_{z=1}^{z} H(x,y,z) \tag{4-6}$$

그동안, 컬러 히스토그램을 다룬 다수의 논문이 발표되었다. Soffer등은 컬러 히스토그램의 가중치 거리(weighted - distance) 인덱싱에 대한 효율적인 방법을 제시하였다. 모든 영상은 동일한 수의 픽셀 M을 가지고 있다고 볼 수 있다. 하나의 영상은 n개의 구별되는 컬러를 갖는 컬러공간으로 분리된다. 컬러 히스토그램 H는 각 빈의 hi 영상 내 컬러 j의 픽셀수를 포함하는 벡터(h1, h2,… hn)이다. 즉, 임의의 영상 I에 대하여 히스토그램 Hi는 화상의 모든 컬러 픽셀의 합이다. 따라서 컬러 히스토그램은 다음 식 (4-7)과 같이 나타낼 수 있다. 여기서, n은 컬러의 최대 분포레벨이다.

$$H_i = \sum_{i=0}^{n} hi \tag{4-7}$$

전형적으로 컬러 히스토그램은 일치되는 차이점의 합(L2-distance) 또는 차이점들의 절대치 값의 합(L1-distance)을 사용해서 계산한다. 그래서 I와 가장 유사한 영상은 영상 I′를 최소화한다.

$$|Hi - Hr| = \sum_{j}^{n}(Hi[j] - Hr[j])^2 \tag{4-8}$$

$$|Hi - Hr| = \sum_{j}^{n}|(Hi[j] - Hr[j])| \tag{4-9}$$

L2 - distance 는 대부분 $\sqrt{2}M$이고, L1 - distance 는 대부분 2M이다. 다음 그림 4-4는 임의의 화상에 대한 컬러 히스토그램을 보이고 있다.

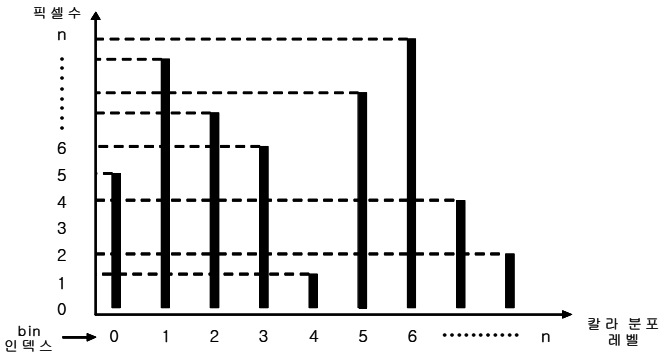


그림 4-4 컬러 빈도수로 구성된 컬러 히스토그램

그림 4-5는 본 논문에서 제안한 RGB영상을 HSV로 변한한 후에 각각 색상(H) 채도(S) 명도(I)를 분리하여 8개의 빈으로 구성된 컬러 히스토그램이다.

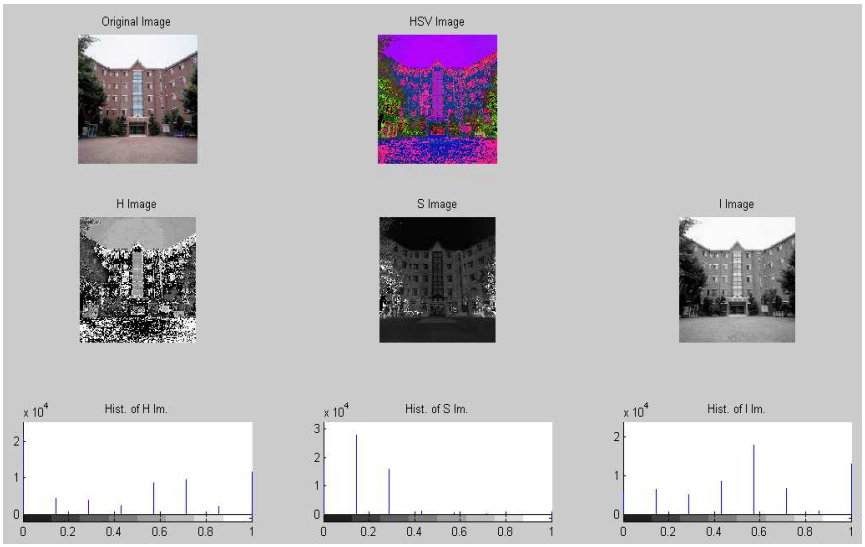


그림 4-5 8빈도수로 구성된 컬러 히스토그램

5. 컬러 코렐로그램

컬러 코렐로그램은 영상의 모든 화소의 컬러에 대해 일정 거리가 떨어진 화소에 나타나는 컬러를 2차원 확률로 나타낸다. m 개의 컬러로 양자화 된 영상 I 에서 화소 간의 거리가 k 인, 컬러 C_i 와 C_j 에 대한 컬러 코렐로그램 $r^k_{C_i,C_j}(I)$ 는 식 (4-10)과 같다.

$$r^{(k)}_{c_i,c_j} = \Pr_{p_1 \in I_{c_i},\, p_2 \in I_{c_j}} \left[p_2 \in I_{c_i = c_j} \middle| \left| p_1 - p_2 \right| = k \right] \tag{4-10}$$

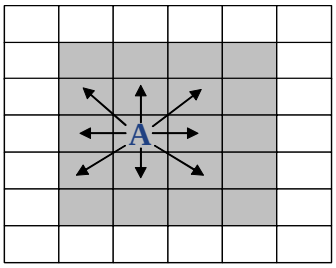
여기서, $\Pr[\cdot]$ 은 내부조건을 만족하는 확률, $i,j = 0,1,\cdots,m-1$ 이고 m 은 양자화 된 컬러의 개수이다 I_{C_i} 와 I_{C_j} 는 영상에서 i 와 j 번째 컬러를 가지는 모든 화소를 나타낸다. 화소 p_a 와 p_b 간의 거리 $|p_a - p_b|$ 는 식 (4-11)과 같다.

$$|p_a - p_b| = \max(|x_a - x_b|, |y_a - y_b|) \tag{4-11}$$

여기서, 각 화소의 좌표는 $p_a = (x_a, y_a) \in I$, $p_b = (x_b, y_b) \in I$ 이다. 이때, $C_i = C_j$ 인 영역을 오토-코렐로그램 이라 한다. 여기에는 영상의 한 화소가 가지는 컬러와 같은 컬러를 가지는 화소들이 그 주위에 얼마나 분포되어 있는지에 대한 정보를 가지는데, 하나의 컬러를 가지는 영역의 크기정보를 포함한다. 또, 컬러 코렐로그램의 대부분의 에너지는 이 영역에 나타난다.

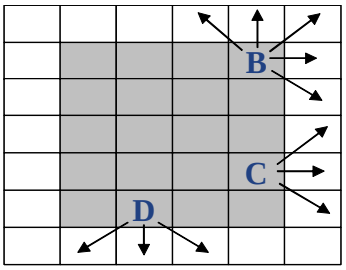
$$r^{(k)}_{c_i,c_j} = \Pr_{p_1 \in I,\, p_2 \in I} \left[p_2 \in I_{c_i \neq c_j} \middle| \left| p_1 - p_2 \right| = k \right] \tag{4-12}$$

식 (5-13)는 컬러 크로스 코렐로그램으로써 같은 컬러를 가지는 확률 분포를 제외한다. 즉 픽셀의 주위의 같은 컬러는 제외하고 다른 컬러가 나오는 부분의 확률을 나타낸다.



칼라	0	1	2	3
0	◎			
1		◎		
2			◎	
3				◎

그림 4-6 컬러 오토 코렐로그램 획득 방법



칼라	0	1	2	3
0		◎	◎	◎
1	◎		◎	◎
2	◎	◎		◎
3	◎	◎	◎	

그림 4-7 컬러 크로스 코렐로그램 획득 방법

다음 그림 4-6은 컬러 코렐로그램을 획득하는 방법이며, 컬러 코렐로그램은 거리와 각도에 따라 계산되는 알고리즘이다. 그림 4-6에서 보는 바와 같이 A픽셀을 기준으로 하고 거리를 1로 하여 기준 픽셀들과 주위의 같은 컬러의 픽셀을 찾아내는 확률을 나타낸 것이다. 그림 4-7에서의 픽셀 B, C, D가 크로스 코렐로그램 부분이다. B, C, D픽셀을 기준으로 하고 거리를 1로 하여 기준 픽셀들과 주위의 다른 컬러의 픽셀을 찾아내는 확률을 나타낸 것이다.

V. 시뮬레이션 및 결과

A. 시뮬레이션 환경 및 평가

1. 환경

본 연구에서는 펜티엄 CPU 3.4GHZ 프로세서를 장착한 개인용 컴퓨터시스템을 사용하여 제안한 방법들을 실험 하였고, 컬러와 에지 정제기법을 이용한 콘텐츠기반 영상 검색 설계 구축을 위한 소프트웨어 및 OS는 Windows XP 와 Matlab 7.4을 사용 하였다.

2. 영상 검색 성능 평가척도

본 논문에서는 영상 검색의 성능을 평가하기 위하여, 멀티미디어 컨텐츠 기술에 관한 표준인 MPEG-7에서 멀티미디어 검색 방법의 성능 평가 척도로 권고하는 ARR(Average Retrieval Recall)와 ARP(Average Retrieval Precision)을 사용하였으며 ARR와 ARP는 검색율을 평가하기 위하여 설계된 척도이다.

ARR은 영상 데이터베이스 내에서 질의 영상과 관련된 영상의 개수와 검색결과 중에서 질의 영상과 관련된 영상의 개수의 비율로 나타내며, ARP는 영상 데이터베이스 내에서 검색된 영상의 수 중 질의 영상과 관련된 영상의 개수의 비율로 나타내며, 둘 다 질의 영상과 관련된 영상이 얼마나 많이 검색되는지를 평가하는 척도이다.

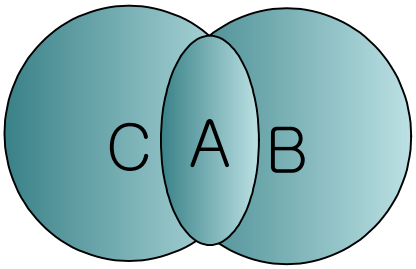


그림 5-1 재현율 및 정확도의 구성 방법

표 5-1 재현율 및 정확도의 구성의 기호

집합의 구성	기호	의 미
A+B+C+D	데이터 베이스	전체 영상의 수
A+C	T	전체 영상 중에서 질의와 관련된 영상의 총수
A+B	Ri	전체 영상 중에서 검색된 영상의 수
B+D	Nq	전체 영상 중에서 질의영상과 관련 없는 영상의 수
A	Rr	검색된 영상 중에서 질의영상과 관련된 영상의 수

재현율(Recall)은 영상 데이터베이스에서 질의와 관련된 영상의 총수 중 검색된 영상 중에서 질의 영상과 관련된 영상의 수의 비율로써 다음의 구성된 기호들을 이용해서 식 (5-1)로 나타낼 수 있다.

$$Recall = \frac{Rr}{T} \tag{5-1}$$

정확도(Precision)는 영상 데이터베이스에서 검색된 영상의 수 중 질의영상과 관련된 영상의 수의 비율로써 다음의 구성된 기호들을 이용해서 식 (5-2)를 나타낼 수 있다.

$$\mathrm{Pr}ecision = \frac{Rr}{Ri} \tag{5-2}$$

재현율과 정확도는 [0, 1]사이의 실수값을 가지며, 1에 가까운 값을 가질수록 검색율이 높다는 것을 의미한다. ARR과 ARP는 모든 질의 영상 q에 대한 재현율(q), 정확도(q)의 합의 평균을 의미하고 식 (5-3)과 식 (5-4)로 나타낼 수 있다.

$$ARR = \frac{1}{n(Q)} \sum_{q \in Q} Recall(q) \tag{5-3}$$

$$ARP = \frac{1}{n(Q)} \sum_{q \in Q} \mathrm{Precision}(q) \qquad (5\text{-}4)$$

ARR과 ARP는 [0, 1] 사이의 실수값을 가지며 1에 가까운 값을 가질수록 검색율이 높다는 것을 의미한다.

B. 시뮬레이션 결과 및 분석

1. 컬러 정제에 따른 시뮬레이션 결과 및 분석

아래 표 5-2는 IMG001의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은 2개, 인접 평균은 11189, 비인접 평균은 11로 나타나고, 빈6에서는 인접에서는 92%, 비인접에서는 8%, 인접 그룹은 4개, 비인접 그룹은 9개, 인접 평균은 828, 비인접 평균은 32개로 나타낸다. 그림 5-2는 시뮬레이션 히스토그램 및 양자화 시킨 결과 영상이다.

표 5-2 정제된 컬러 영상 IMG0001 기술자 값

IMG001	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	100	0	1	2	11189	15
빈 2	0	100	0	2	0	11
빈 3	0	0	0	1	0	0
빈 4	0	0	0	1	1	0
빈 5	0	0	0	1	0	0
빈 6	92	8	4	9	828	32
빈 7	99	1	3	2	3229	26

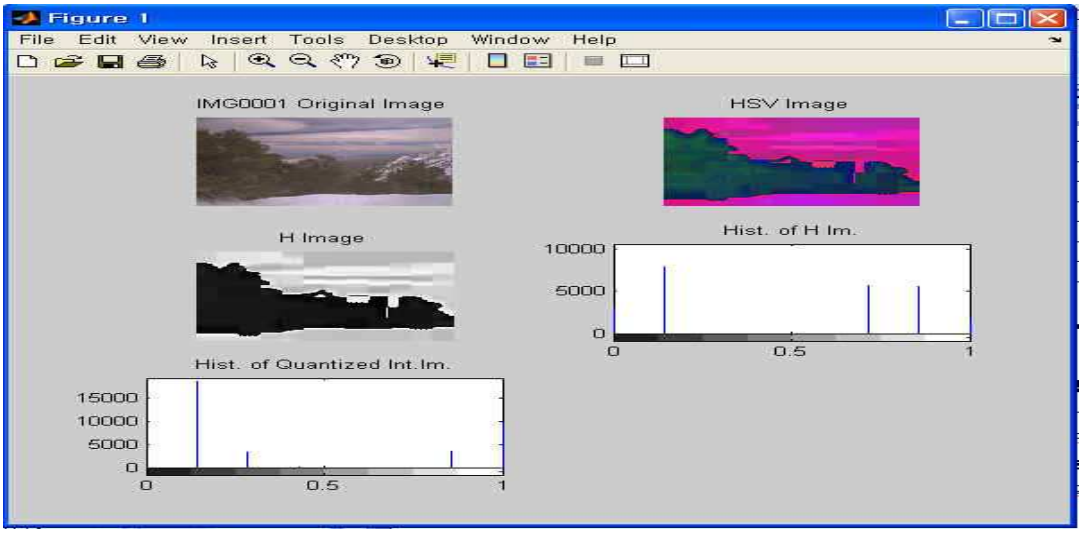


그림 5-2 정제된 컬러영상 IMG0001

아래 표 5-3 IMG002의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 99%, 비인접 백분율은 1%, 인접 그룹은 1개, 비인접 그룹은 7개, 인접 평균은 9938, 비인접 평균은 9로 나타나고, 빈6에서는 인접에서는 92%, 비인접에서는 8%, 인접 그룹은 4개, 비인접 그룹은 12개, 인접 평균은 851, 비인접 평균은 25개로 나타낸다. 그림 5-3은 시뮬레이션 히스토그램 및 양자화 시킨 결과 영상이다.

표 5-3 정제된 컬러 영상 IMG0002 기술자 값

IMG002	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	99	1	1	7	9938	9
빈 2	0	100	0	5	0	3
빈 3	0	0	0	1	0	0
빈 4	0	0	0	1	0	0
빈 5	0	0	0	1	0	0
빈 6	92	8	4	12	851	25
빈 7	99	1	3	16	3580	7

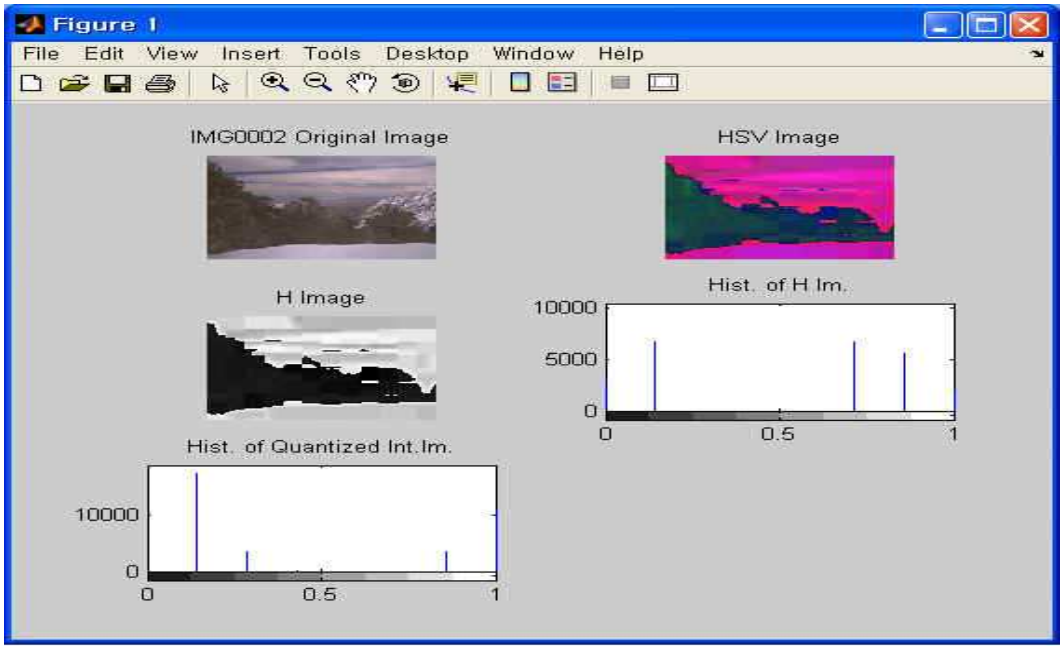


그림 5-3 정제된 컬러영상 IMG0002

아래 표 5-4 IMG003의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 98%, 비인접 백분율은 2%, 인접 그룹은 3개, 비인접 그룹은 4개, 인접 평균은 3889, 비인접 평균은 70으로 나타나고, 빈6에서는 인접에서는 81%, 비인접에서는 19%, 인접 그룹은 1개, 비인접 그룹은 21개, 인접 평균은 655, 비인접 평균은 7개로 나타낸다. 그림 5-4는 시뮬레이션 히스토그램 및 양자화시킨 결과 영상이다.

표 5-4 정제된 컬러 영상 IMG0003 기술자 값

IMG003	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	98	2	3	4	3889	70
빈 2	0	100	0	7	0	5
빈 3	0	0	0	1	0	0
빈 4	0	0	0	1	0	0
빈 5	0	0	0	1	0	0
빈 6	81	19	1	21	655	7
빈 7	100	0	1	1	11780	2

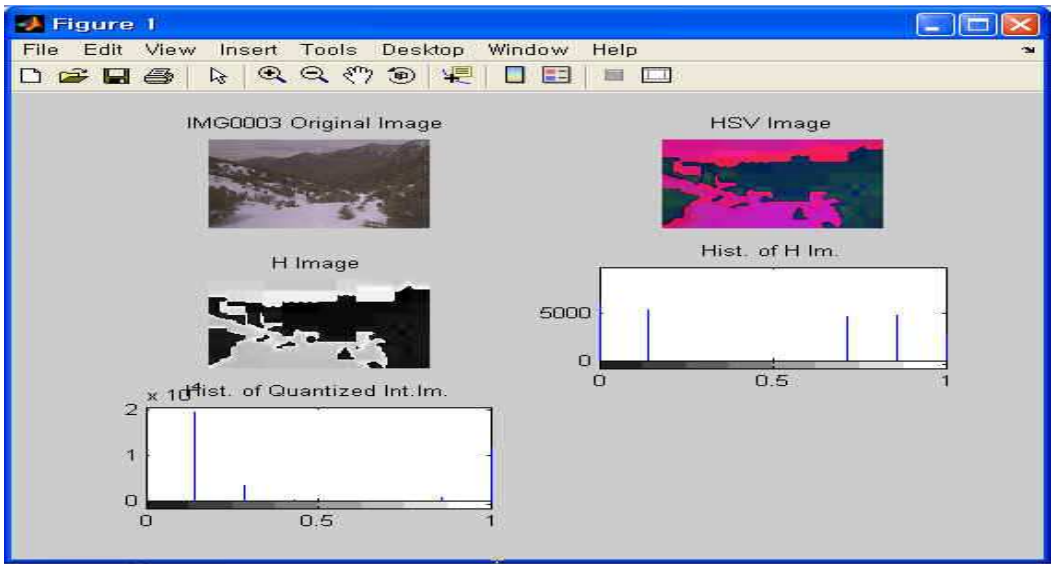


그림 5-4 정제된 컬러영상 IMG0003

아래 표 5-5 IMG004의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 85%, 비인접 백분율은 15%, 인접 그룹은 5개, 비인접 그룹은 6개, 인접 평균은 517, 비인접 평균은 74로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 1개, 비인접 그룹은 5개, 인접 평균은 0, 비인접 평균은 59개로 나타낸다.

표 5-5 정제된 컬러 영상 IMG0004 기술자 값

IMG004	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	85	15	5	6	517	74
빈 2	0	100	0	5	0	5
빈 3	0	0	0	1	0	0
빈 4	0	100	0	2	0	1
빈 5	0	100	0	2	0	21
빈 6	0	100	1	5	0	59
빈 7	100	0	0	0	21182	1

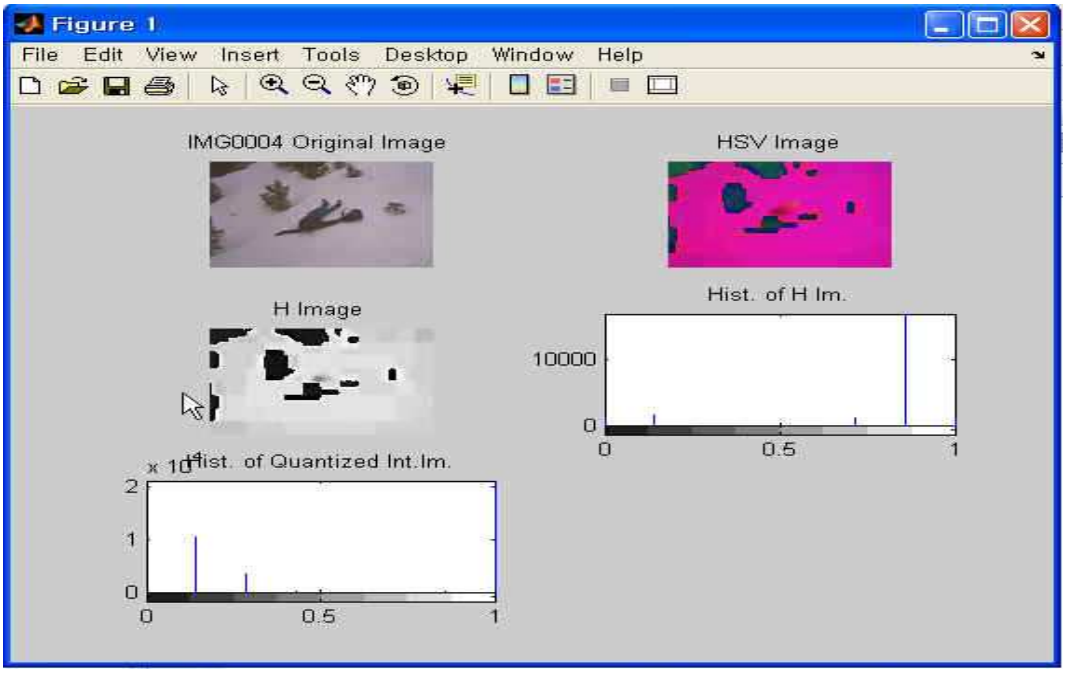


그림 5-5 정제된 컬러영상 IMG004

아래 표 5-6 IMG005의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은 11개, 인접 평균은16488, 비인접 평균은 3으로 나타나고, 빈6에서는 인접에서는 97%, 비인접에서는 3%, 인접 그룹은 1개, 비인접 그룹은 7개, 인접 평균은 5528, 비인접 평균은 25개로 나타낸다.

표 5-6 정제된 컬러 영상 IMG0005 기술자 값

IMG005	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	100	0	1	11	16488	3
빈 2	0	100	0	19	0	12
빈 3	0	100	0	2	0	2
빈 4	0	100	0	2	0	1
빈 5	0	100	0	30	0	16
빈 6	97	3	1	7	5528	25
빈 7	45	55	2	39	361	22

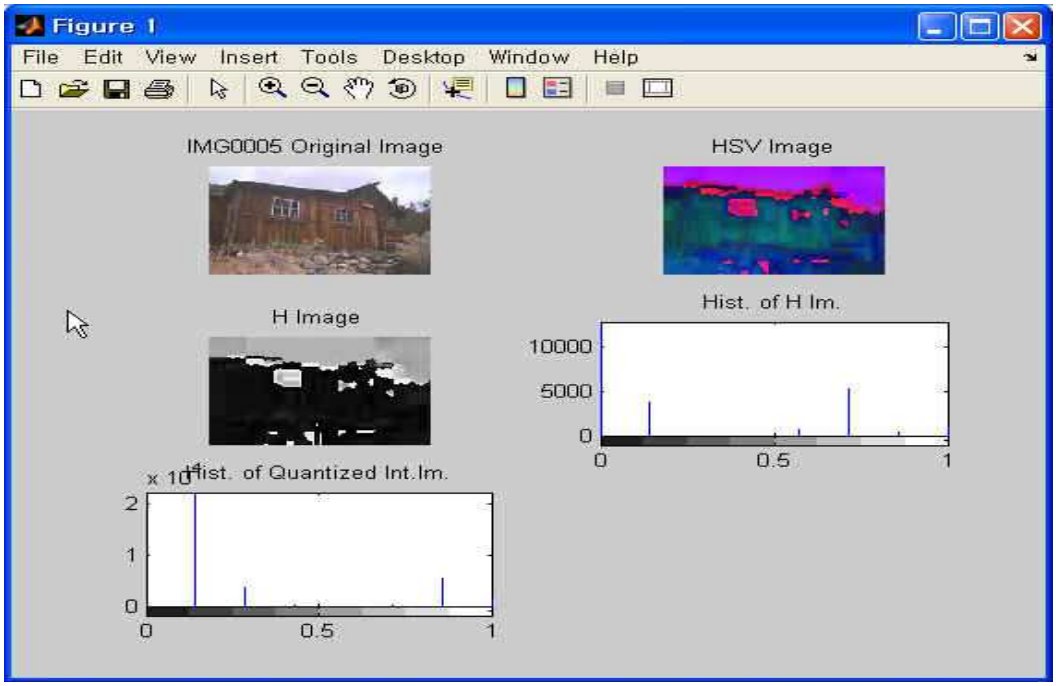


그림 5-6 정제된 컬러영상 IMG0005

아래 표 5-7 IMG006의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은 1개, 인접 평균은18393, 비인접 평균은 31로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 2개, 인접 평균은 0, 비인접 평균은 13개로 나타낸다.

표 5-7 정제된 컬러 영상 IMG0006 기술자 값

IMG006	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	100	0	1	1	18393	31
빈 2	91	9	2	8	489	12
빈 3	0	100	0	2	0	4
빈 4	0	100	0	2	0	28
빈 5	0	100	0	30	0	0
빈 6	0	100	0	2	0	13
빈 7	75	25	4	17	938	73

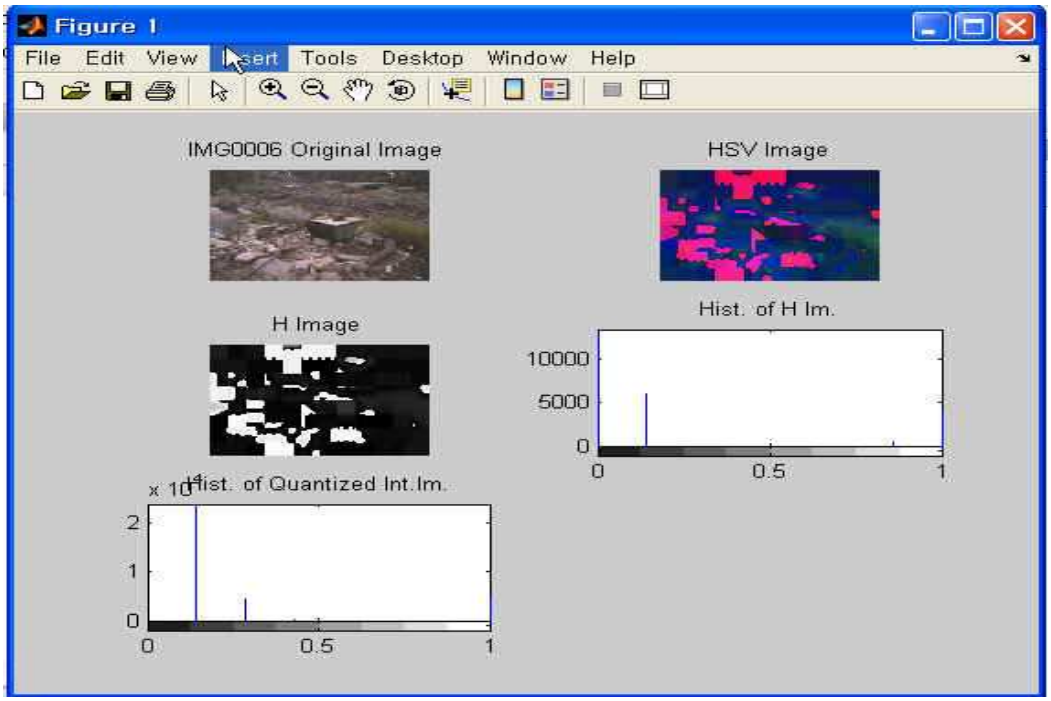


그림 5-7 정제된 컬러영상 IMG0006

아래 표 5-8 IMG007의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은 1개, 인접 평균은18393, 비인접 평균은 31로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 2개, 인접 평균은 0, 비인접 평균은 13개로 나타낸다.

표 5-8 정제된 컬러 영상 IMG0007 기술자 값

IMG007	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	100	0	1	1	18393	31
빈 2	91	9	2	8	489	12
빈 3	0	100	0	2	0	4
빈 4	0	100	0	2	0	28
빈 5	0	100	0	30	0	0
빈 6	0	100	0	2	0	13
빈 7	75	25	4	17	938	73

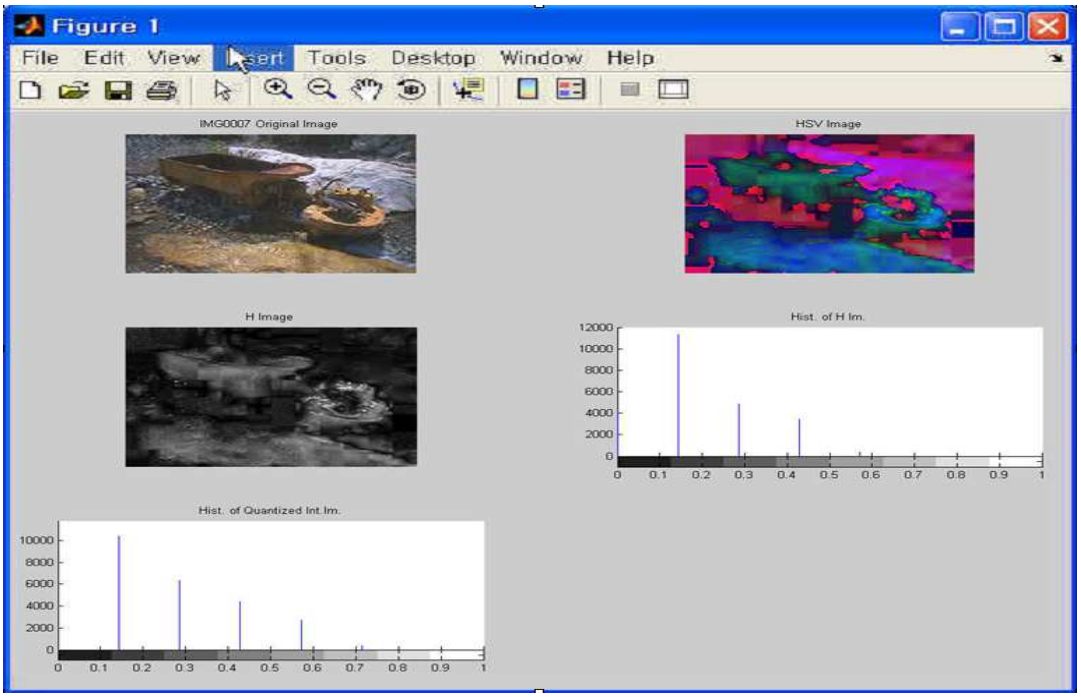


그림 5-8 정제된 컬러영상 IMG0007

아래 표 5-9 IMG0010의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 97%, 비인접 백분율은 3%, 인접 그룹은 1개, 비인접 그룹은28개, 인접 평균은11962, 비인접 평균은 11로 나타나고, 빈6에서는 인접에서는 82%, 비인접에서는 18%, 인접 그룹은 4개, 비인접 그룹은 47개, 인접 평균은 936, 비인접 평균은 17개로 나타낸다.

표 5-9 정제된 컬러 영상 IMG0010 기술자 값

IMG0010	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	97	3	1	28	11962	11
빈 2	16	84	1	78	318	21
빈 3	0	100	0	46	0	10
빈 4	0	100	0	69	0	9
빈 5	85	15	4	31	672	15
빈 6	82	18	4	47	936	17
빈 7	0	100	0	72	0	22

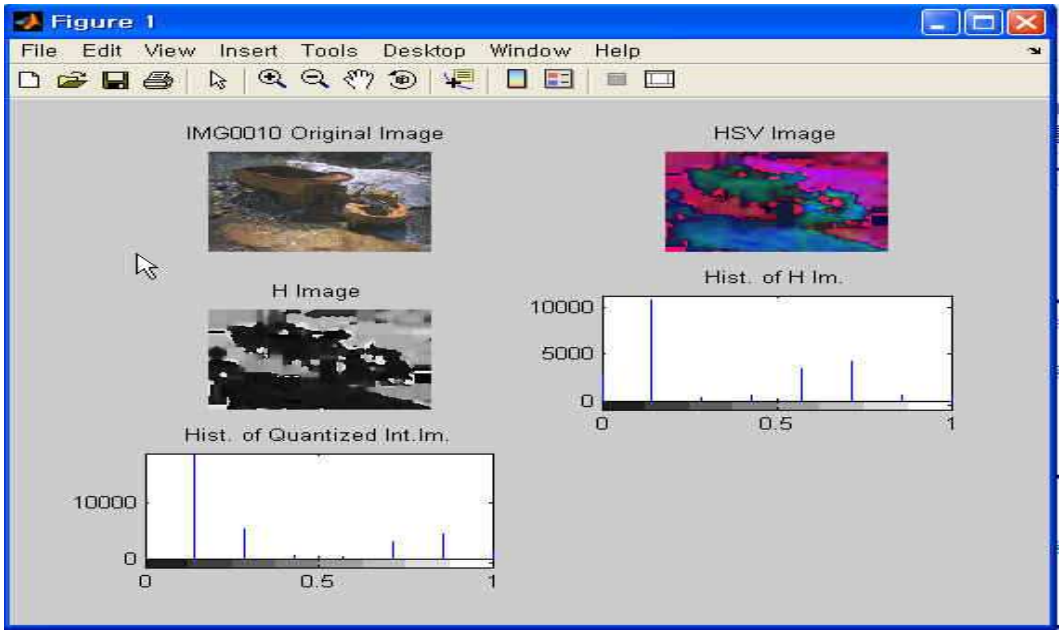


그림 5-9 정제된 컬러영상 IMG0010

아래 표 5-10 IMG0011의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 97%, 비인접 백분율은 3%, 인접 그룹은 1개, 비인접 그룹은28개, 인접 평균은11962, 비인접 평균은 11로 나타나고, 빈6에서는 인접에서는 82%, 비인접에서는 18%, 인접 그룹은 4개, 비인접 그룹은 47개, 인접 평균은 936, 비인접 평균은 17개로 나타낸다.

표 5-10 정제된 컬러 영상 IMG0011 기술자 값

IMG0011	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	70	30	2	44	611	12
빈 2	74	26	3	34	1289	40
빈 3	0	100	0	31	0	17
빈 4	0	100	0	19	0	5
빈 5	0	100	0	22	0	11
빈 6	57	43	2	18	477	40
빈 7	82	18	4	32	736	21

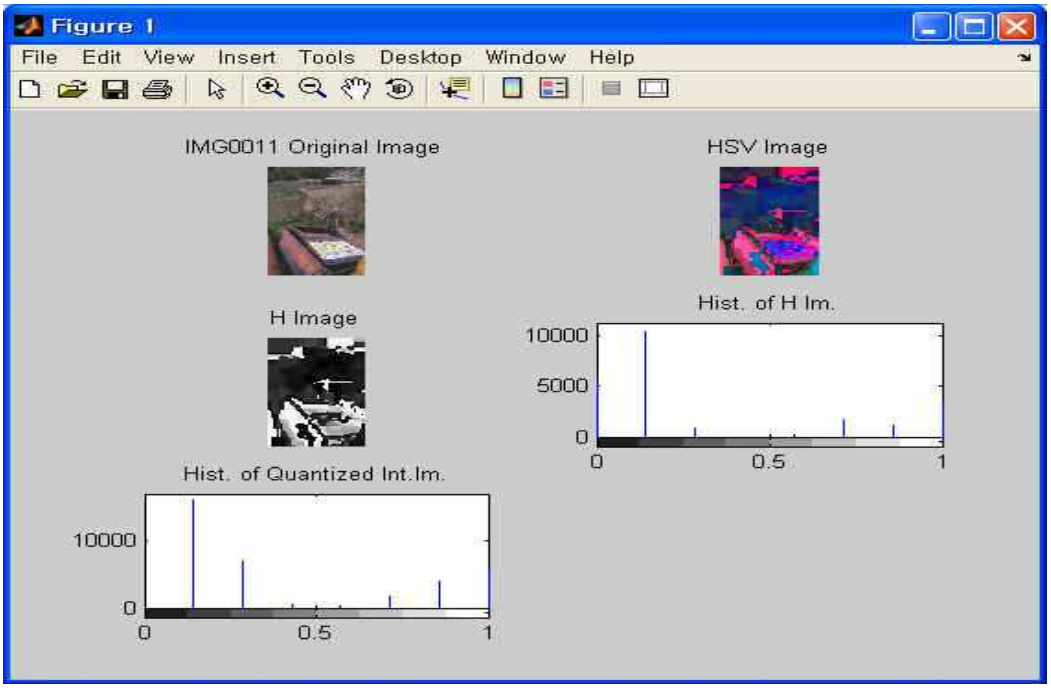


그림 5-10 정제된 컬러영상 IMG0011

아래 표 5-11 IMG0012의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 0%, 비인접 백분율은 100%, 인접 그룹은 0개, 비인접 그룹은32개, 인접 평균은0, 비인접 평균은 21로 나타나고, 빈6에서는 인접에서는 85%, 비인접에서는 15%, 인접 그룹은 2개, 비인접 그룹은 28개, 인접 평균은 1218 비인접 평균은 15개로 나타낸다.

표 5-11 정제된 컬러 영상 IMG0012 기술자 값

IMG0012	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	0	100	0	32	0	21
빈 2	69	31	1	27	2346	38
빈 3	0	100	0	20	0	20
빈 4	0	100	0	28	0	11
빈 5	0	100	0	21	0	24
빈 6	85	15	2	28	1218	15
빈 7	81	19	4	28	1036	35

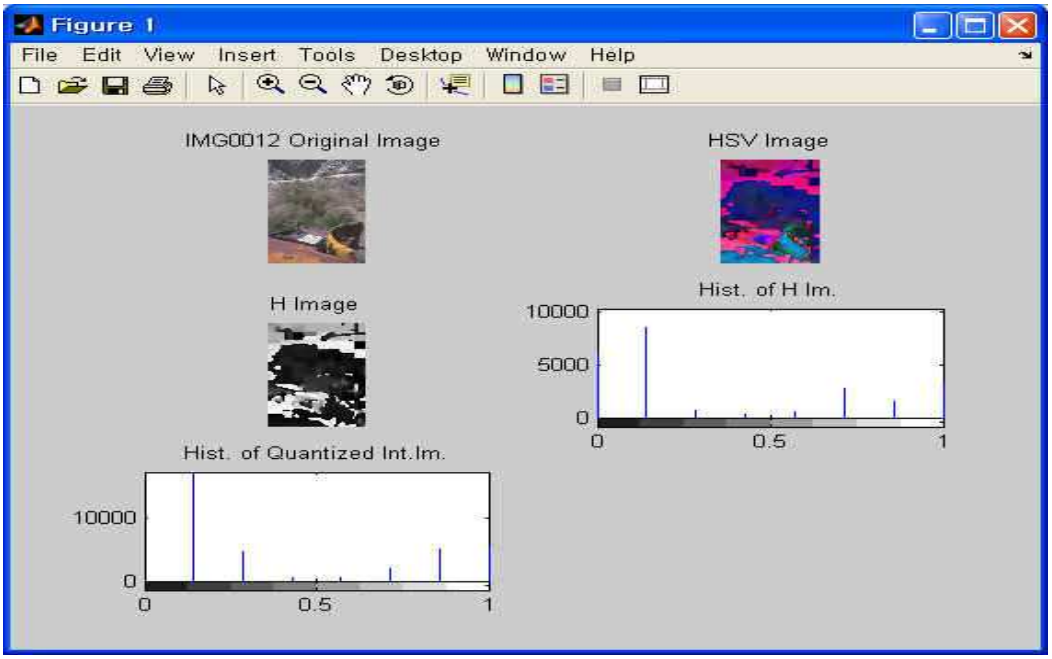


그림 5-11 정제된 컬러영상 IMG0012

아래 표 5-12 IMG0013의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 98%, 비인접 백분율은 2%, 인접 그룹은 1개, 비인접 그룹은1개, 인접 평균은15191, 비인접 평균은 233으로 나타나고, 빈6에서는 인접에서는 100%, 비인접에서는 0%, 인접 그룹은 2개, 비인접 그룹은 1개, 인접 평균은 1562 비인접 평균은 0개로 나타낸다.

표 5-12 정제된 컬러 영상 IMG0013 기술자 값

IMG0013	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	98	2	1	1	15191	233
빈 2	0	100	0	3	0	5
빈 3	0	0	0	1	0	0
빈 4	0	0	0	1	0	0
빈 5	0	0	0	1	0	0
빈 6	100	0	2	1	1562	0
빈 7	96	4	2	12	2900	18

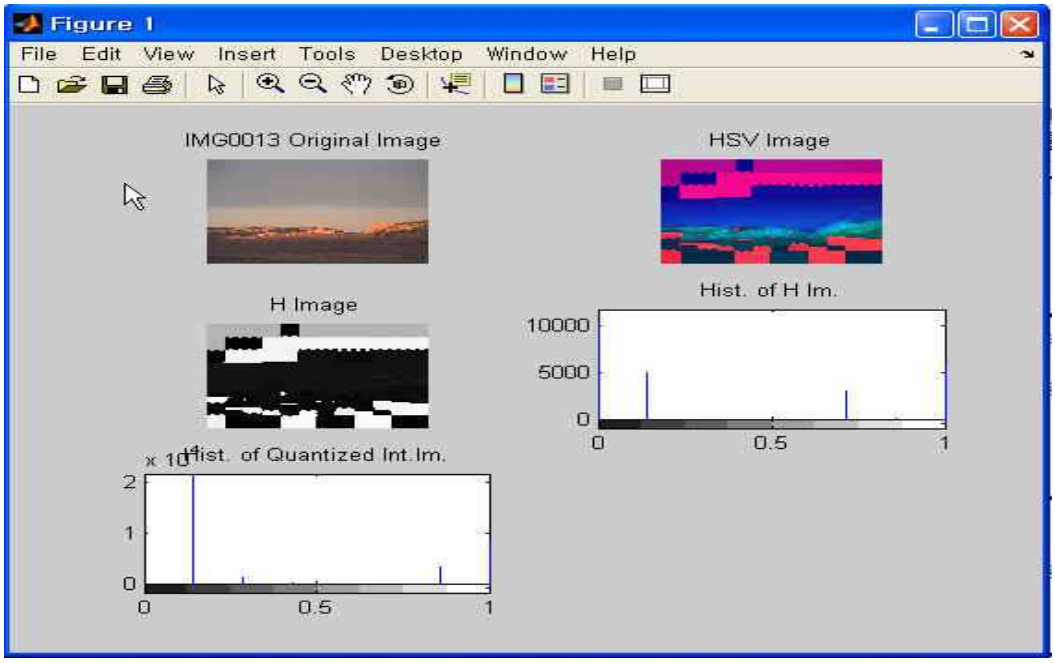


그림 5-12 정제된 컬러영상 IMG0013

아래 표 5-13 IMG0014의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은23개, 인접 평균은11772, 비인접 평균은 1로 나타나고, 빈6에서는 인접에서는 100%, 비인접에서는 0%, 인접 그룹은 1개, 비인접 그룹은 4개, 인접 평균은 5437 비인접 평균은 2개로 나타낸다.

표 5-13 정제된 컬러 영상 IMG0014 기술자 값

IMG0014	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	100	0	1	23	11772	1
빈 2	0	100	0	5	0	5
빈 3	0	0	0	1	0	0
빈 4	0	100	0	2	0	26
빈 5	0	100	0	11	0	22
빈 6	100	0	1	4	5437	2
빈 7	96	4	4	12	1686	23

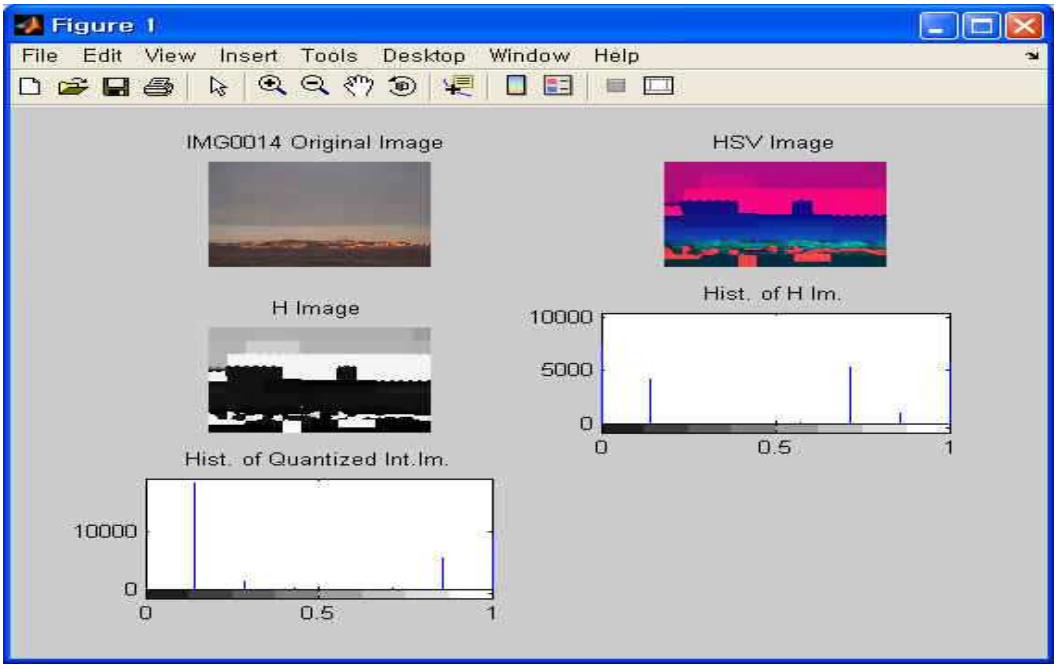


그림 5-13 정제된 컬러영상 IMG0014

아래 표 5-14 IMG0015의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 98%, 비인접 백분율은 2%, 인접 그룹은 1개, 비인접 그룹은20개, 인접 평균은16271, 비인접 평균은 17로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 8개, 인접 평균은 0 비인접 평균은 11개로 나타낸다.

표 5-14 정제된 컬러 영상 IMG0015 기술자 값

IMG0015	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	98	2	1	20	16271	17
빈 2	0	100	0	8	0	8
빈 3	0	0	0	1	0	0
빈 4	0	0	0	1	0	8
빈 5	0	100	0	5	0	3
빈 6	0	100	0	8	0	11
빈 7	96	2	2	11	3826	14

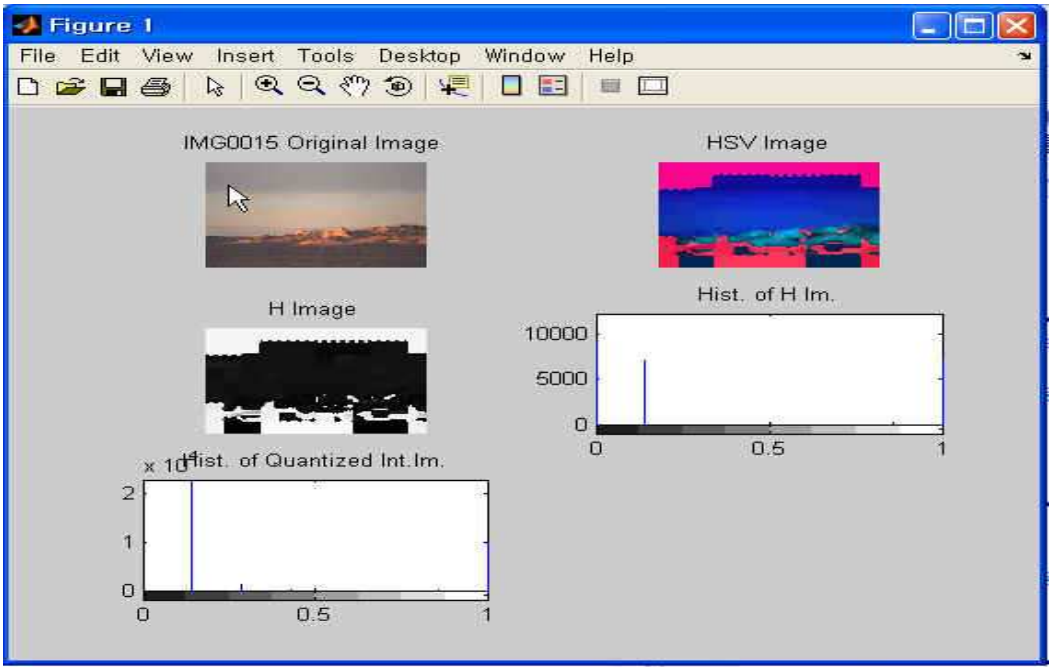


그림 5-14 정제된 컬러영상 IMG0015

아래 표 5-15 IMG0025의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은0개, 인접 평균은64768, 비인접 평균은 0으로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 0%, 인접 그룹은 0개, 비인접 그룹은 0개, 인접 평균은 0 비인접 평균은 0개로 나타낸다.

표 5-15 정제된 컬러 영상 IMG0025 기술자 값

IMG0025	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	0	0	0
빈 1	100	0	1	0	64768	0
빈 2	0	0	0	0	0	0
빈 3	0	0	0	0	0	0
빈 4	0	0	0	0	0	0
빈 5	0	0	0	0	0	0
빈 6	0	0	0	0	0	0
빈 7	0	0	0	0	0	0

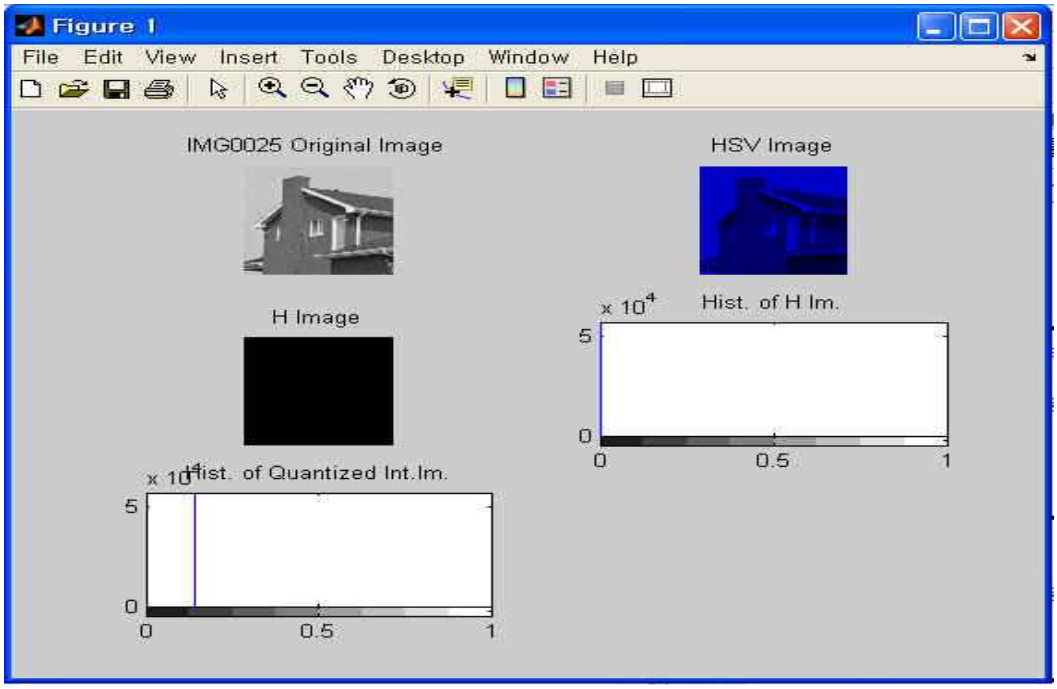


그림 5-15 정제된 컬러영상 IMG0025

아래 표 5-16 IMG0026의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은0개, 인접 평균은64768, 비인접 평균은 0으로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 0%, 인접 그룹은 0개, 비인접 그룹은 0개, 인접 평균은 0 비인접 평균은 0개로 나타낸다.

표 5-16 IMG0025가 180° 회전(IMG0026) 정제된 컬러 영상 기술자 값

IMG0026	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	0	0	0
빈 1	100	0	1	0	64768	0
빈 2	0	0	0	0	0	0
빈 3	0	0	0	0	0	0
빈 4	0	0	0	0	0	0
빈 5	0	0	0	0	0	0
빈 6	0	0	0	0	0	0
빈 7	0	0	0	0	0	0

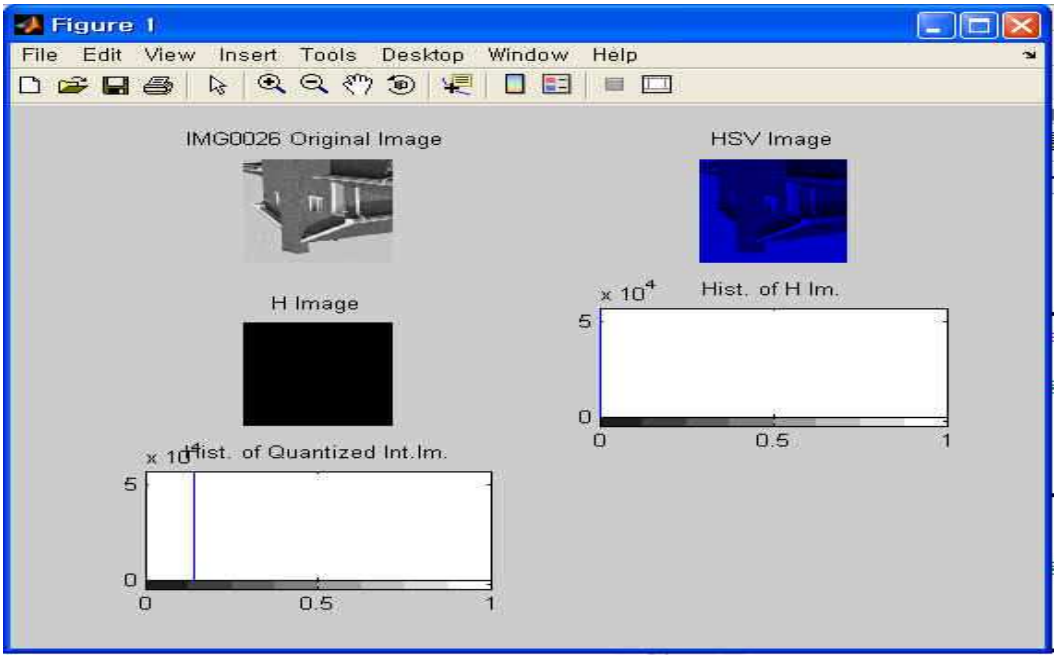


그림 5-16 정제된 컬러 영상 IMG0026

아래 표 5-17 IMG0027의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은0개, 인접 평균은64768, 비인접 평균은 0으로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 0%, 인접 그룹은 0개, 비인접 그룹은 0개, 인접 평균은 0 비인접 평균은 0개로 나타낸다.

표 5-17 IMG0025가 좌우반전(IMG0027) 정제된 컬러 영상 기술자 값

IMG0027	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	0	0	0
빈 1	100	0	1	0	64768	0
빈 2	0	0	0	0	0	0
빈 3	0	0	0	0	0	0
빈 4	0	0	0	0	0	0
빈 5	0	0	0	0	0	0
빈 6	0	0	0	0	0	0
빈 7	0	0	0	0	0	0

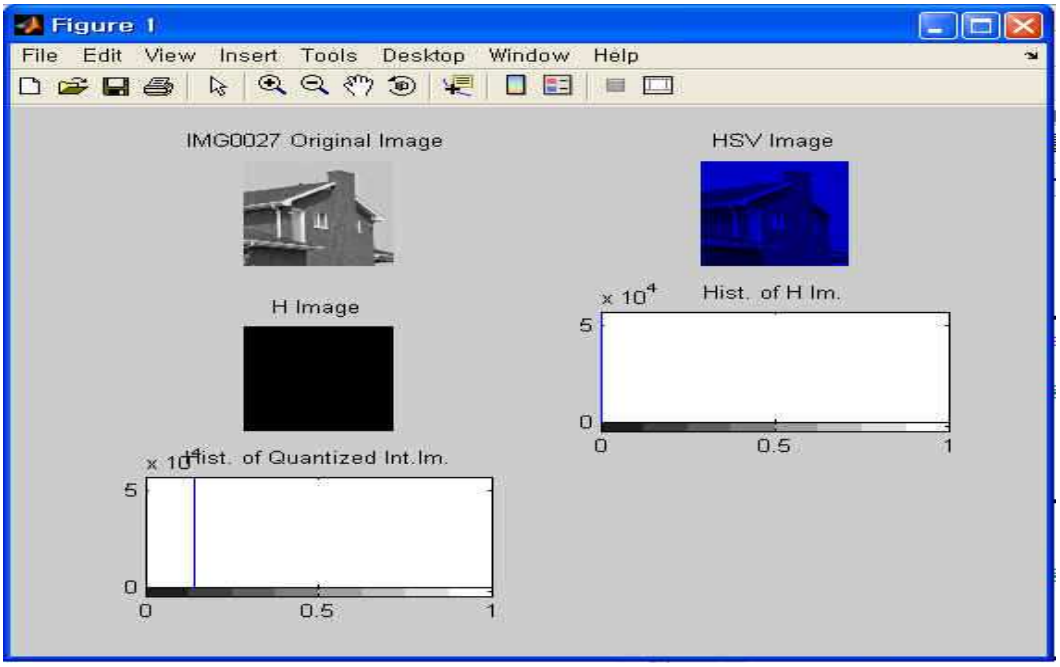


그림 5-17 정제된 컬러 영상 IMG0027

아래 표 5-18 IMG0028의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은0개, 인접 평균은130048, 비인접 평균은 0으로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 0%, 인접 그룹은 0개, 비인접 그룹은 0개, 인접 평균은 0 비인접 평균은 0개로 나타낸다.

표 5-18 IMG0025가 좌우늘임(IMG0028) 정제된 컬러 영상 기술자 값

IMG0028	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	0	0	0
빈 1	100	0	1	0	130048	0
빈 2	0	0	0	0	0	0
빈 3	0	0	0	0	0	0
빈 4	0	0	0	0	0	0
빈 5	0	0	0	0	0	0
빈 6	0	0	0	0	0	0
빈 7	0	0	0	0	0	0

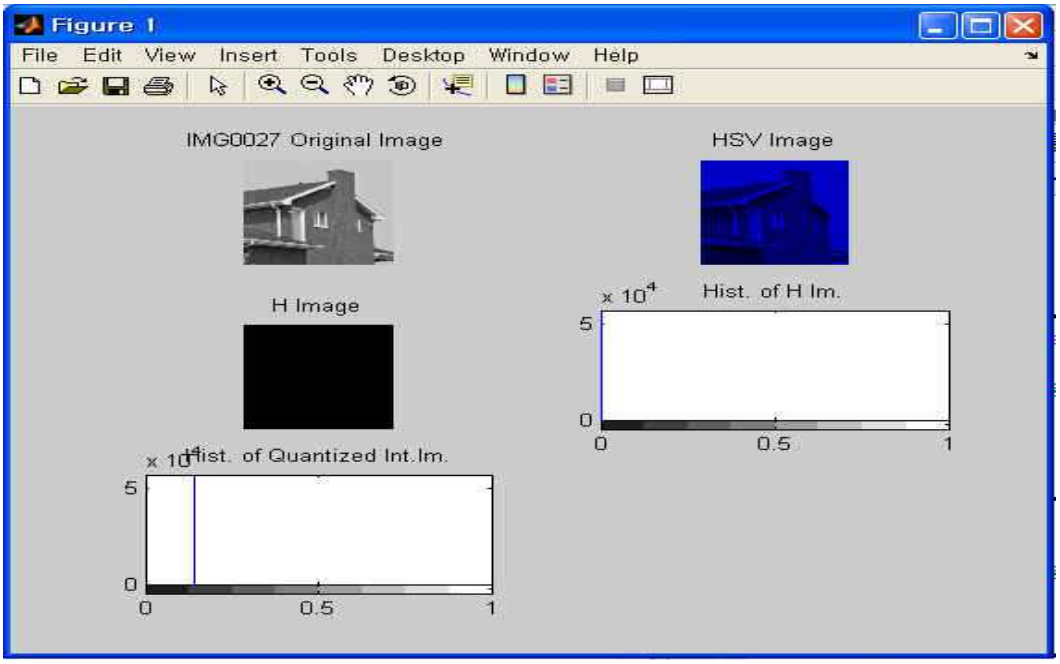


그림 5-18 정제된 컬러영상 IMG0028

아래 표 5-19 IMG0032의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은0개, 인접 평균은130048, 비인접 평균은 0으로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 0%, 인접 그룹은 0개, 비인접 그룹은 0개, 인접 평균은 0 비인접 평균은 0개로 나타낸다.

표 5-19 IMG0025가 상하반전(IMG0032) 정제된 컬러 영상 기술자 값

IMG0032	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	100	0	1	0	64768	0
빈 2	0	0	0	1	0	0
빈 3	0	0	0	1	0	0
빈 4	0	0	0	1	0	0
빈 5	0	0	0	1	0	0
빈 6	0	0	0	1	0	0
빈 7	0	0	0	1	0	0

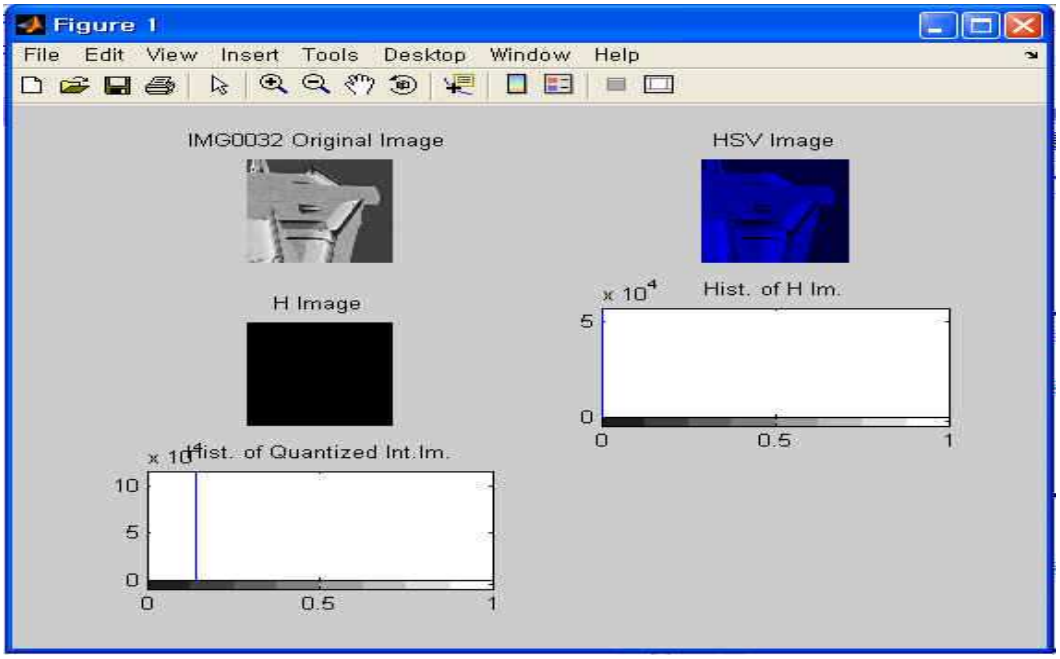


그림 5-19 정제된 컬러영상 IMG0032

아래 표 5-20 IMG0038의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 98%, 비인접 백분율은 2%, 인접 그룹은 1개, 비인접 그룹은62개, 인접 평균은22585, 비인접 평균은 7로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 43개, 인접 평균은 0 비인접 평균은 2개로 나타낸다.

표 5-20 정제된 컬러 영상 IMG0038 기술자 값

IMG0038	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	98	2	1	62	22585	7
빈 2	78	22	1	163	7983	14
빈 3	34	66	2	110	419	15
빈 4	0	100	0	65	0	2
빈 5	0	100	0	41	0	2
빈 6	0	100	0	43	0	2
빈 7	0	100	0	43	0	3

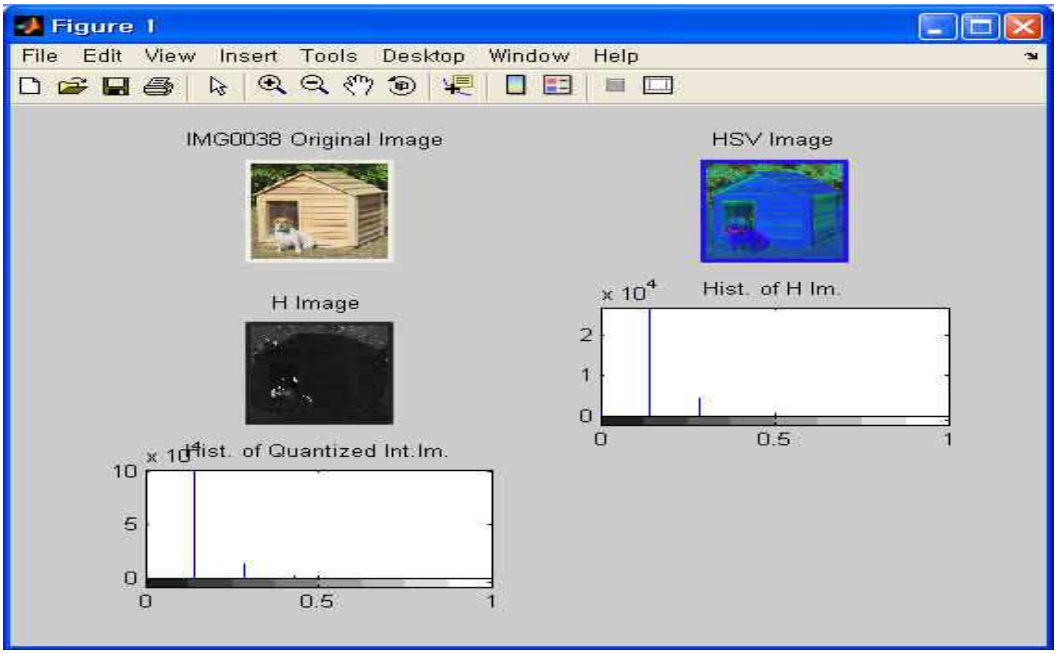


그림 5-20 정제된 컬러영상 IMG0038

아래 표 5-21 IMG0039의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 97%, 비인접 백분율은 3%, 인접 그룹은 1개, 비인접 그룹은37개, 인접 평균은23128, 비인접 평균은 20으로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 38개, 인접 평균은 0 비인접 평균은 5개로 나타낸다.

표 5-21 IMG0038가 좌우반전(IMG0039) 정제된 컬러 영상 기술자 값

IMG0039	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	97	3	1	37	23128	20
빈 2	88	12	2	67	4174	17
빈 3	28	72	1	76	551	19
빈 4	0	100	0	26	0	2
빈 5	0	100	0	13	0	3
빈 6	0	100	0	14	0	5
빈 7	0	100	0	33	0	8

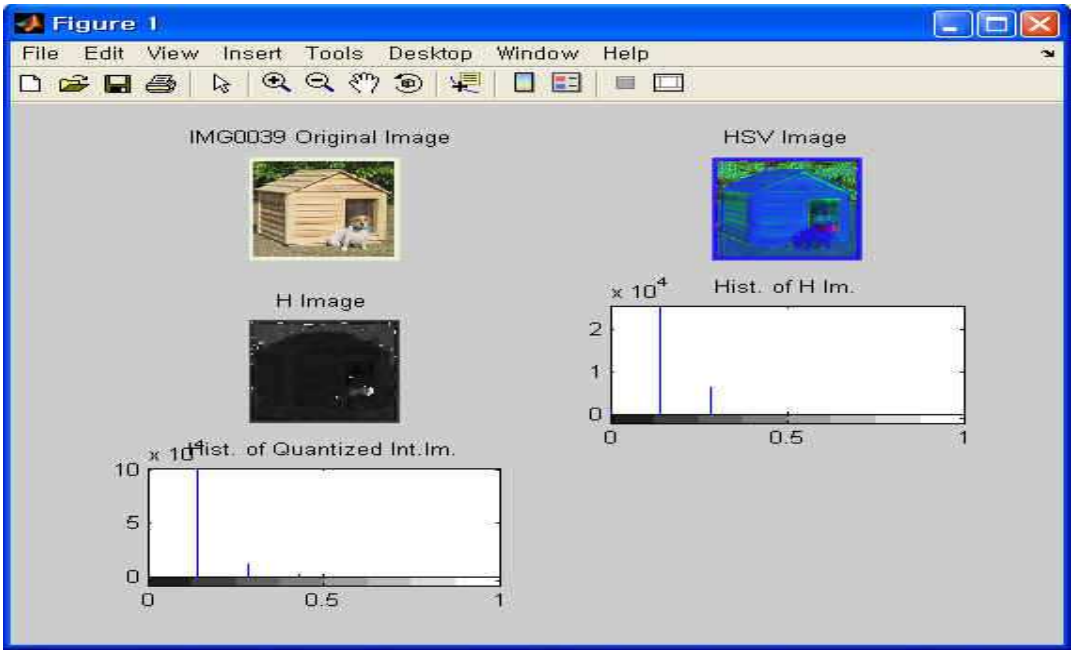


그림 5-21 정제된 컬러영상 IMG0039

아래 표 5-22 IMG0040의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 98%, 비인접 백분율은 2%, 인접 그룹은 1개, 비인접 그룹은27개, 인접 평균은23633, 비인접 평균은 14로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 9개, 인접 평균은 0 비인접 평균은 6개로 나타낸다.

표 5-22 IMG0038가 상하반전(IMG0040) 정제된 컬러 영상 기술자 값

IMG0040	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	98	2	1	27	23633	14
빈 2	83	17	1	63	7948	26
빈 3	33	67	1	90	593	13
빈 4	0	100	0	22	0	4
빈 5	0	100	0	10	0	5
빈 6	0	100	0	9	0	6
빈 77	0	100	0	22	0	8

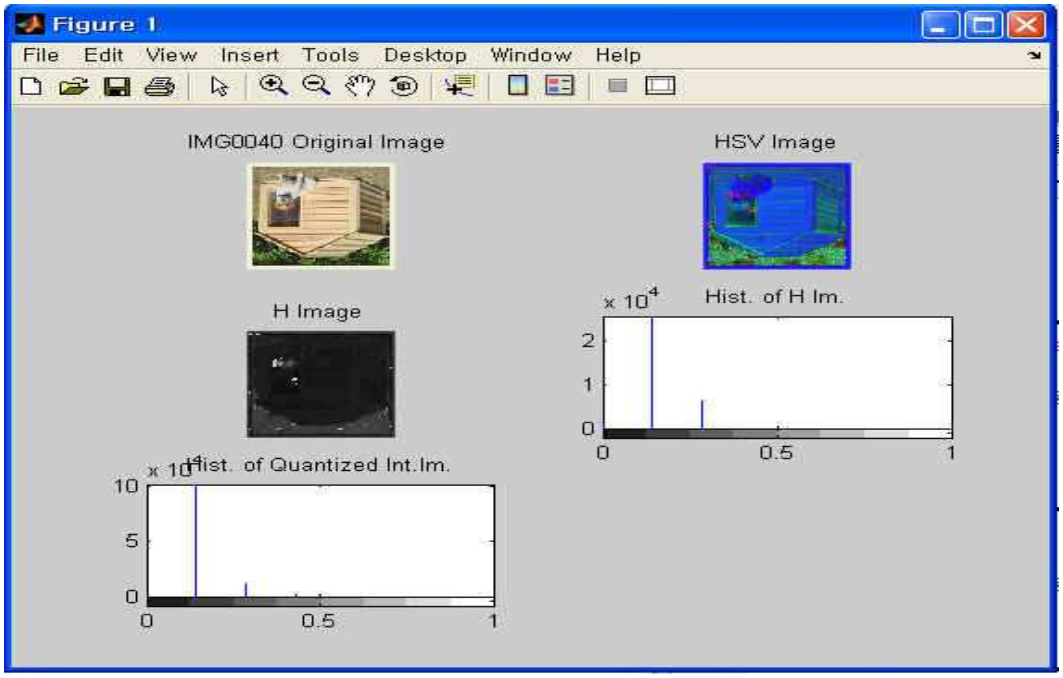


그림 5-22 정제된 컬러영상 IMG0040

아래 표 5-23 IMG0041의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 98%, 비인접 백분율은 2%, 인접 그룹은 1개, 비인접 그룹은27개, 인접 평균은23633, 비인접 평균은 14로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 9개, 인접 평균은 0 비인접 평균은 6개로 나타낸다.

표 5-23 IMG0038가 90° 회전(IMG0041) 정제된 컬러 영상 기술자 값

IMG0041	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	97	3	1	21	23397	29
빈 2	85	15	2	52	3959	26
빈 3	30	70	1	72	625	21
빈 4	0	100	0	23	0	4
빈 5	0	100	0	12	0	4
빈 6	0	100	0	10	0	5
빈 7	0	100	0	23	0	9

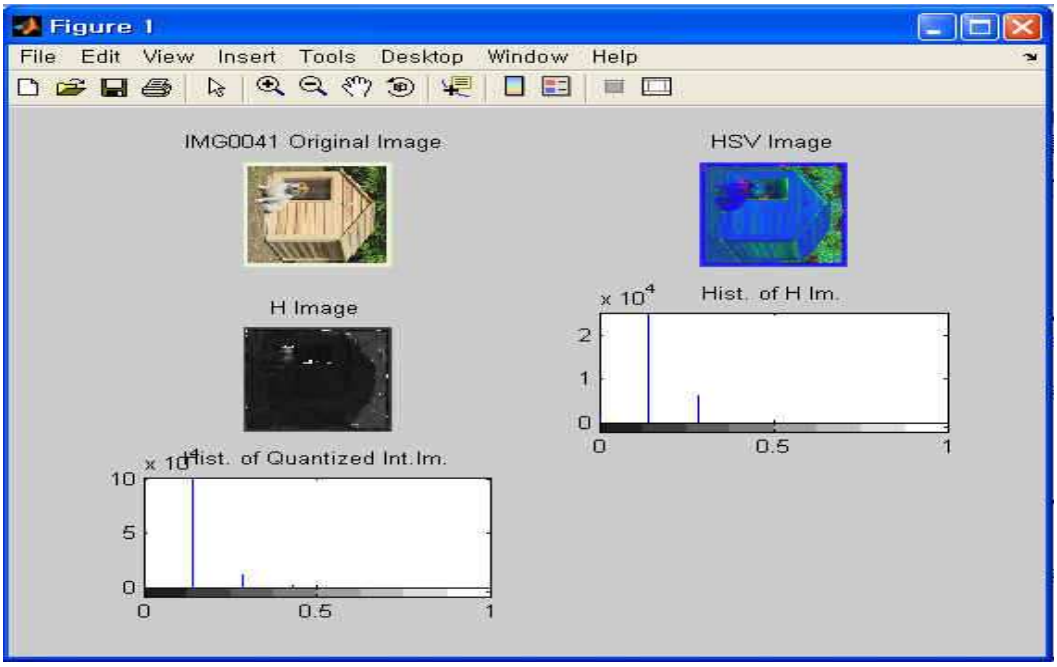


그림 5-23 정제된 컬러영상 IMG0041

아래 표 5-24 IMG0043의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 98%, 비인접 백분율은 2%, 인접 그룹은 1개, 비인접 그룹은27개, 인접 평균은23633, 비인접 평균은 14로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 9개, 인접 평균은 0 비인접 평균은 6개로 나타낸다.

표 5-24 IMG0038가 30° 비틀림(IMG0043) 정제된 컬러 영상 기술자 값

IMG0043	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	100	0	1	12	21027	2
빈 2	89	11	1	45	5789	16
빈 3	51	49	2	39	350	17
빈 4	0	100	0	9	0	5
빈 5	0	100	0	6	0	7
빈 6	0	100	0	7	0	7
빈 7	0	100	0	13	0	8

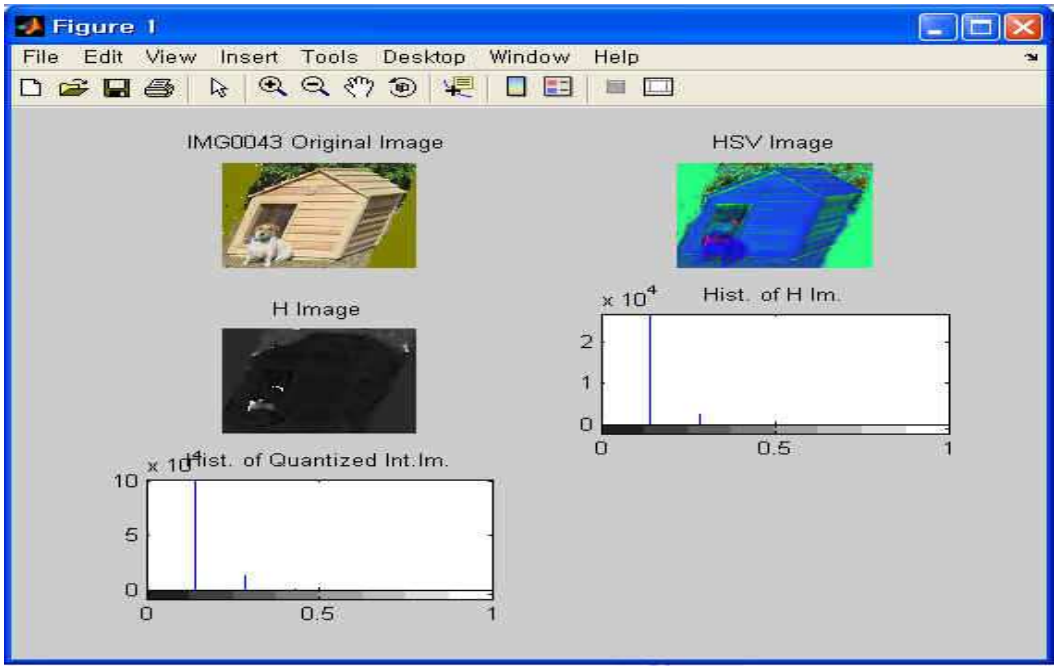


그림 5-24 정제된 컬러영상 IMG0043

아래 표 5-25 IMG0045의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 98%, 비인접 백분율은 2%, 인접 그룹은 1개, 비인접 그룹은27개, 인접 평균은23633, 비인접 평균은 14로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 9개, 인접 평균은 0 비인접 평균은 6개로 나타낸다.

표 5-25 IMG0038가 120° (IMG0045) 정제된 컬러 영상 기술자 값

IMG0045	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	0	0	0
빈 1	100	0	1	148	273757	6
빈 2	0	100	0	187	0	17
빈 3	0	100	0	151	0	5
빈 4	0	100	0	174	0	11
빈 5	0	100	0	149	0	18
빈 6	0	100	0	173	0	7
빈 7	90	10	2	115	12154	23

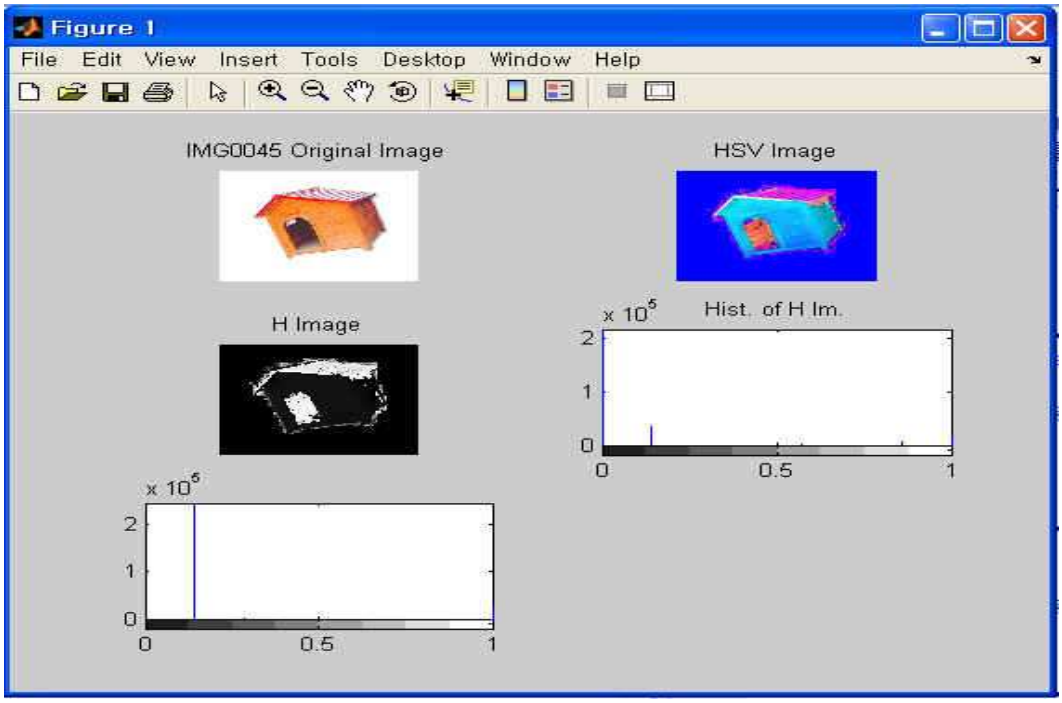


그림 5-25 정제된 컬러영상 IMG0045

아래 표 5-26 IMG0046의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은51개, 인접 평균은104638, 비인접 평균은 5로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 71개, 인접 평균은 0 비인접 평균은 4개로 나타낸다.

표 5-26 IMG0038가 150° 비틀림(IMG0046) 정제된 컬러 영상 기술자 값

IMG0046	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	0	0	0
빈 1	100	0	1	51	104638	5
빈 2	0	100	0	113	0	8
빈 3	0	100	0	62	0	4
빈 4	0	100	0	80	0	7
빈 5	0	100	0	68	0	19
빈 6	0	100	0	71	0	4
빈 7	89	11	2	78	3082	10

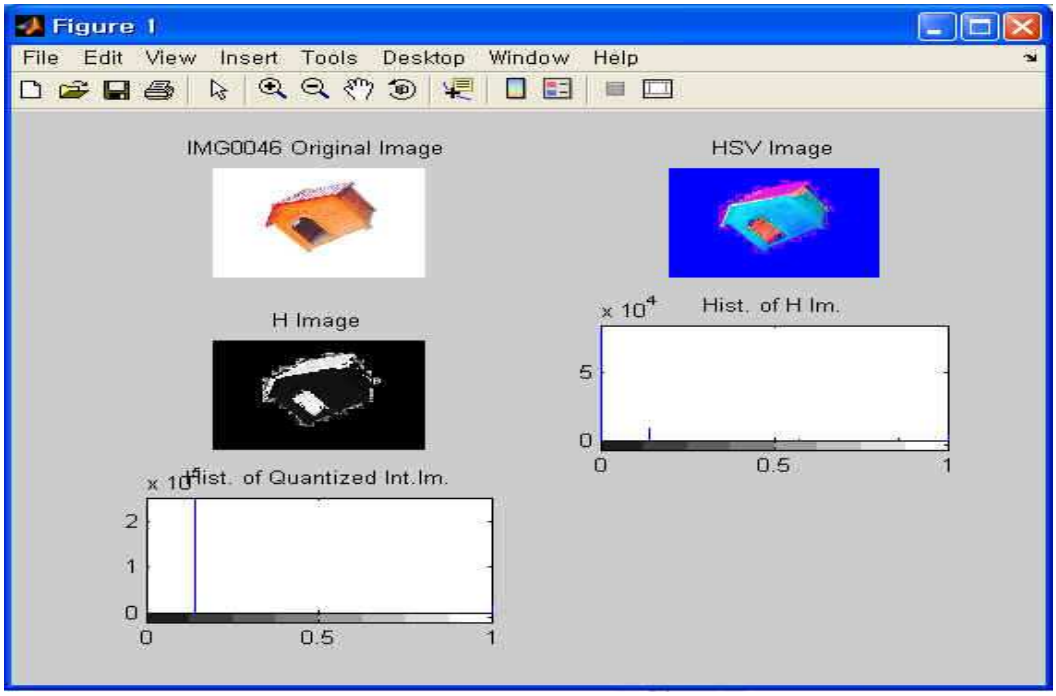


그림 5-26 정제된 컬러영상 IMG0046

아래 표 5-27 IMG0047의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은63개, 인접 평균은10696, 비인접 평균은 4로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 64개, 인접 평균은 0 비인접 평균은 3개로 나타낸다.

표 5-27 IMG0038가 60° 비틀림(IMG0047) 정제된 컬러 영상 기술자 값

IMG0047	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	100	0	1	63	10696	4
빈 2	0	100	0	104	0	9
빈 3	0	100	0	81	0	4
빈 4	0	100	0	99	0	8
빈 5	0	100	0	69	0	18
빈 6	0	100	0	64	0	3
빈 7	86	14	2	53	3082	20

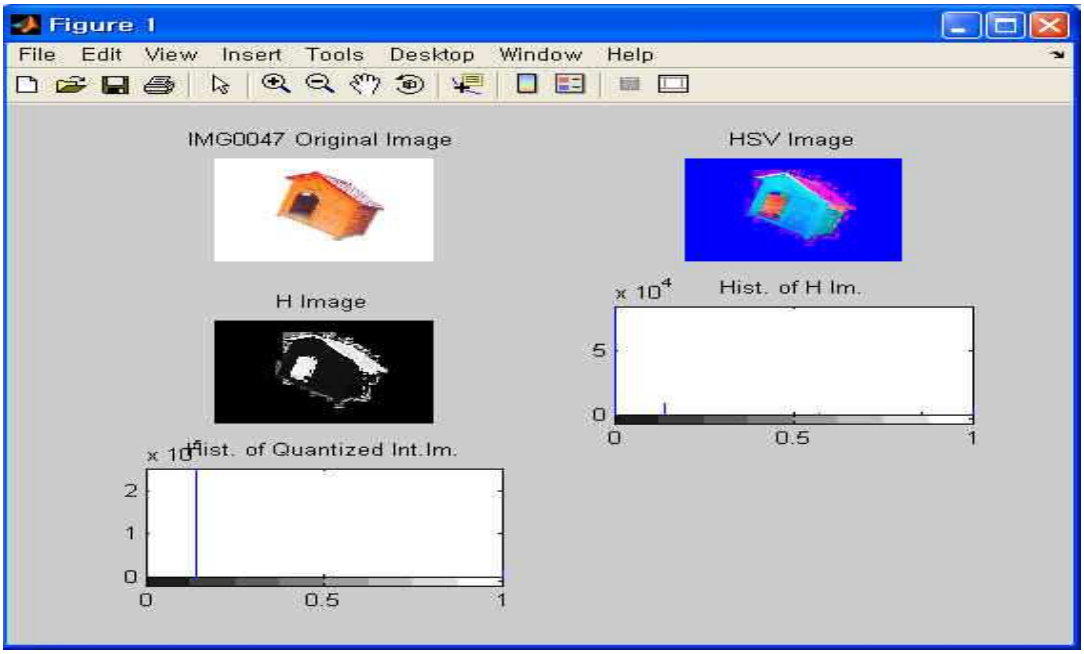


그림 5-27 정제된 컬러영상 IMG0047

아래 표 5-28 IMG0048의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은60개, 인접 평균은103714, 비인접 평균은 4로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 96개, 인접 평균은 0 비인접 평균은 2개로 나타낸다.

표 5-28 IMG0038가 45° 비틀림(IMG0048) 정제된 컬러 영상 기술자 값

IMG0048	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	100	0	1	60	103714	4
빈 2	0	100	0	133	0	7
빈 3	0	100	0	70	0	3
빈 4	0	100	0	96	0	9
빈 5	0	100	0	69	0	23
빈 6	0	100	0	96	0	2
빈 7	88	12	2	91	3185	9

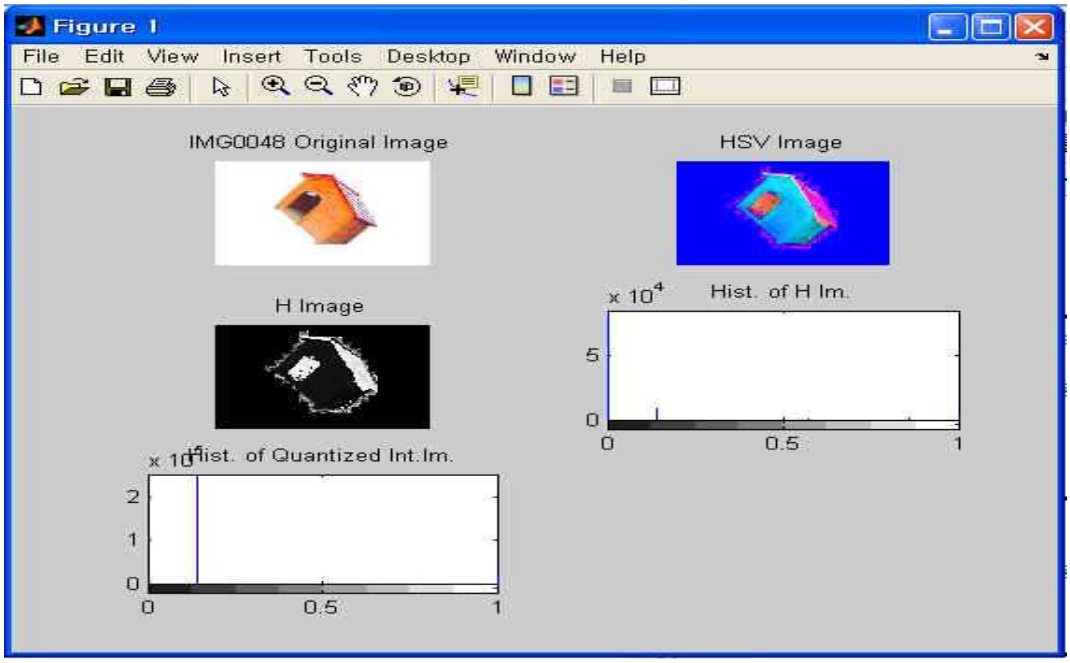


그림 5-28 정제된 컬러영상 IMG0048

아래 표 5-29 IMG0049의 입력영상을 HSV로 변환하여 H 영상만을 적용하여 8빈의 히스토그램으로 하고, 이것을 양자화시킴 으로서 빈0 에서는 정확하게 정제되었으며, 빈1에서 인접 백분율은 100%, 비인접 백분율은 0%, 인접 그룹은 1개, 비인접 그룹은60개, 인접 평균은103714, 비인접 평균은 4로 나타나고, 빈6에서는 인접에서는 0%, 비인접에서는 100%, 인접 그룹은 0개, 비인접 그룹은 96개, 인접 평균은 0 비인접 평균은 2개로 나타낸다.

표 5-29 IMG0038가 좌우늘임(IMG0049) 정제된 컬러 영상 기술자 값

IMG0049	αj	βj	$C\alpha j$	$C\beta j$	$\mu\alpha j$	$\mu\beta j$
빈 0	0	0	0	1	0	0
빈 1	98	2	1	91	37485	7
빈 2	0	100	0	113	0	5
빈 3	0	100	0	72	0	2
빈 4	0	100	0	94	0	5
빈 5	0	100	0	99	0	8
빈 6	0	100	0	94	0	4
빈 7	86	14	2	76	3101	13

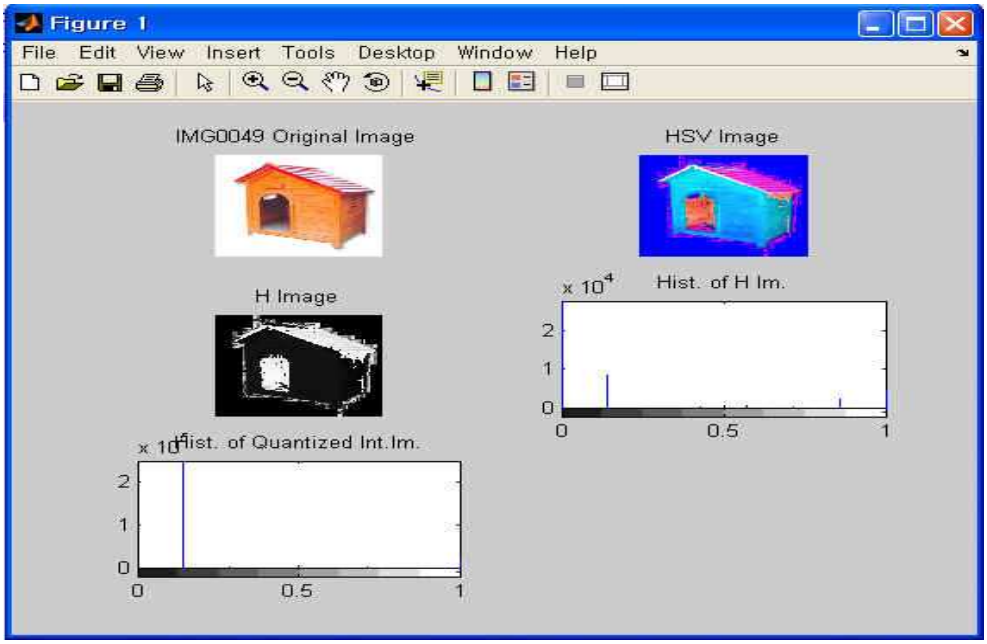


그림 5-29 정제된 컬러영상 IMG0049

IMG0001부터 IMG0004번까지 영상은 자연영상을 컬러 정제방법으로 분석한 결과 재현율은 기존 알고리즘에서는 0.76이였으나 제안된 영상 정제에서는 0.86으로 0.1%향상되었고, 정확도에서는 기존 알고리즘에서는 0.08이고 영상 정제에서는 0.09로서 0.01정도 정확한 검색 결과를 보여 주였고, 컬러 에지 모두 사용 했을 때는 재현율은 0.9와 정확도는 0.09로 컬러 정보만 사용 했을 때보다 향상된 검색 결과를 보였다. 또 IMG0005부터 IMG0006번까지 영상은 컬러 정제방법으로 분석한 결과 재현율은 기존 알고리즘에서는 0.9였으나 제안된 영상 정제에서는 0.9로 동일하였으나 정확도에서는 기존 알고리즘에서는 0.12이나 영상 정제에서는 0.12로서 같은 결과로 나타냈다, IMG0007부터 IMG0015번까지 영상은 컬러 정제방법으로 분석한 결과 재현율은 기존 알고리즘에서는 0.9였으나 제안된 영상정제에서는 0.1%향상되었고 재현율에서는 기존 및 제안된 알고리즘 모두 0.1로 동일하였다, IMG0025부터 IMG0049번까지 영상은 회전과 크기 변화로서의 컬러 정제방법으로 분석한 결과 재현율은 기존알고리즘에서는 0.9이였으나 제안된 영상정제에서는 0.9로 동일한 검색 결과를 보여 주었고, 재현율에서는 기존 및 제안된 알고리즘 모두 0.1로 동일하였으며, 컬러 에지 모두를 사용한 결과 재현율에서는 1과 정확도 에서는 0.01로 모두를 사용 했을 때 탁월한 검색 결과를 보았다.

2. 에지 정제에 따른 시뮬레이션 결과 및 분석

아래 표 5-30은 IMG001의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 775개, 수평 그룹 픽셀 수 110개 평균 픽셀 수 7개 직선 에지 클러스터 수는 39개, 비직선 에지 클러스터 수는 16개로 나타나고, 수직에지인 총 픽셀 수 791개, 수평그룹픽셀 수 170개, 평균 픽셀 수 5개, 직선에지 클러스터 수는 52개, 비직선 에지 클러스터 수는 30개로 나타낸다.

표 5-30 정제된 에지 영상 IMG0001 기술자 값

Image 0001	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	775	110	7	39	16
빈 2: Vertical Edges	791	170	5	52	30

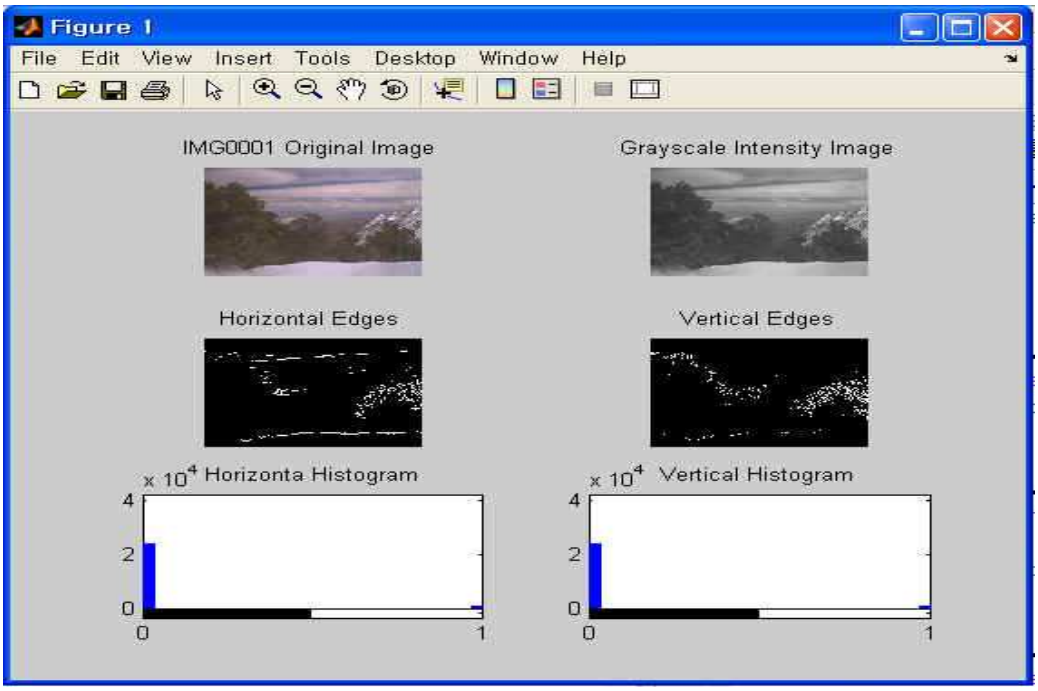


그림 5-30 정제된 에지영상 IMG0001

아래 표 5-31은 IMG002의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 852개, 수평 그룹 픽셀 수 120개 평균 픽셀 수 7개 직선 에지 클러스터 수는 42개, 비직선 에지 클러스터 수는 18개로 나타나고, 수직에지인 총 픽셀 수 819개, 수평그룹픽셀 수 147개, 평균 픽셀 수 6개, 직선에지 클러스터 수는 42개, 비직선 에지 클러스터 수는 27개로 나타낸다.

표 5-31 정제된 에지 영상 IMG0002 기술자 값

Image 0002	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	852	120	7	42	18
빈 2: Vertical Edges	819	147	6	42	27

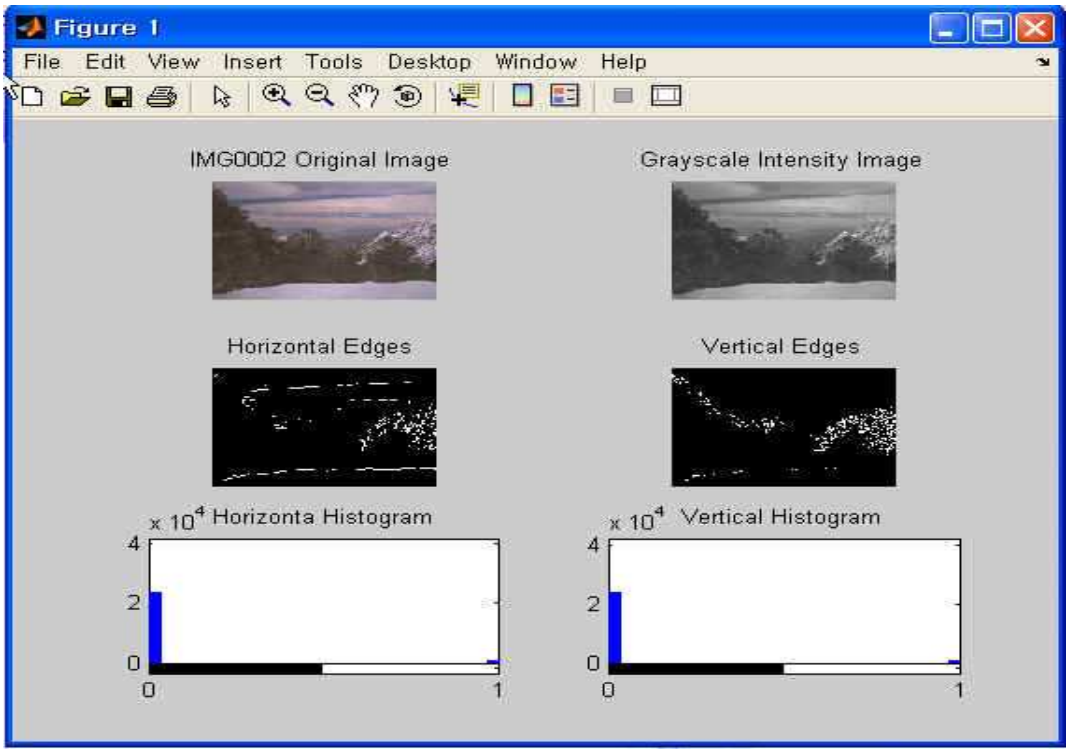


그림 5-31 정제된 에지영상 IMG0002

아래 표 5-32는 IMG003의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 1174개, 수평 그룹 픽셀 수 203개 평균 픽셀 수 6개 직선 에지 클러스터 수는 80개, 비직선 에지 클러스터 수는 41개로 나타나고, 수직에지인 총 픽셀 수 1013개, 수평그룹픽셀 수 215개, 평균 픽셀 수 5개, 직선에지 클러스터 수는 80개, 비직선 에지 클러스터 수는 51개로 나타낸다.

표 5-32 정제된 에지 영상 IMG0003 기술자 값

Image 0003	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	1174	203	6	80	41
빈 2: Vertical Edges	1013	215	5	80	51

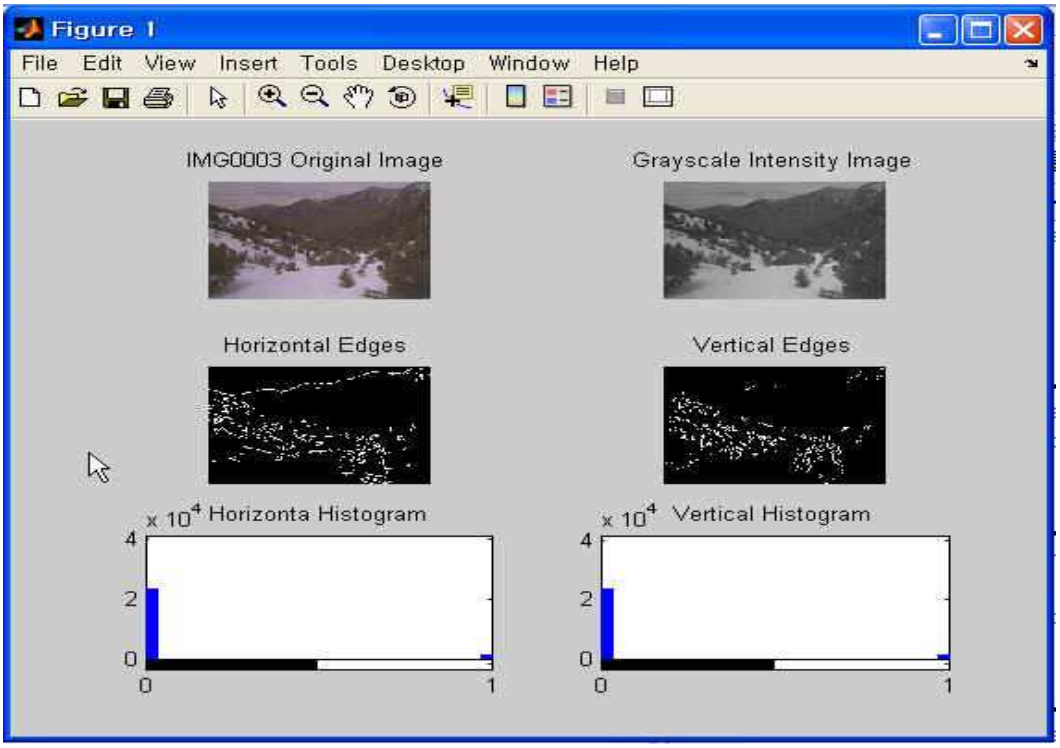


그림 5-32 정제된 에지영상 IMG0003

아래 표 5-33은 IMG005의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 824개, 수평 그룹 픽셀 수 166개 평균 픽셀 수 5개 직선 에지 클러스터 수는 64개, 비직선 에지 클러스터 수는 18개로 나타나고, 수직에지인 총 픽셀 수 954개, 수평그룹픽셀 수 226개, 평균 픽셀 수 4개, 직선에지 클러스터 수는 75개, 비직선 에지 클러스터 수는 29개로 나타낸다.

표 5-33 정제된 에지 영상 IMG0005 기술자 값

Image 0005	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	824	166	5	64	18
빈 2: Vertical Edges	954	226	4	75	29

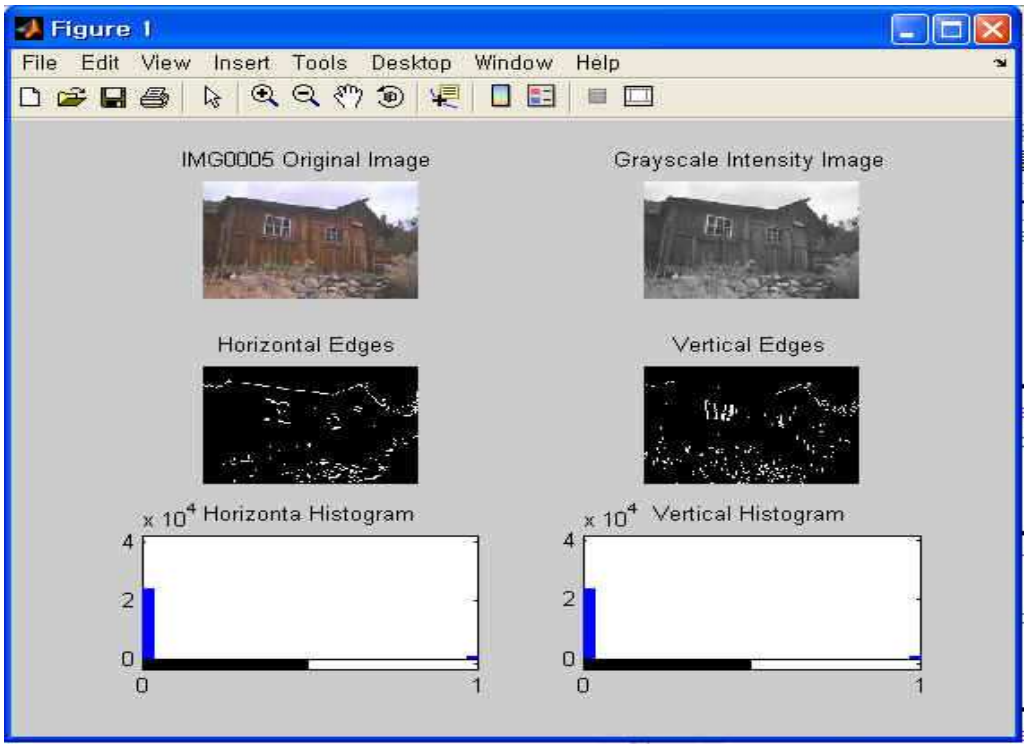


그림 5-33 정제된 에지영상 IMG0005

아래 표 5-34는 IMG0011의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 1063개, 수평 그룹 픽셀 수 204개 평균 픽셀 수 5개 직선 에지 클러스터 수는 70개, 비직선 에지 클러스터 수는 28개로 나타나고, 수직에지인 총 픽셀 수 961개, 수평그룹픽셀 수 251개, 평균 픽셀 수 4개, 직선에지 클러스터 수는 54개, 비직선 에지 클러스터 수는 30개로 나타낸다.

표 5-34 정제된 에지 영상 IMG0011 기술자 값

Image 0011	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	1063	204	5	70	28
빈 2: Vertical Edges	961	251	4	54	30

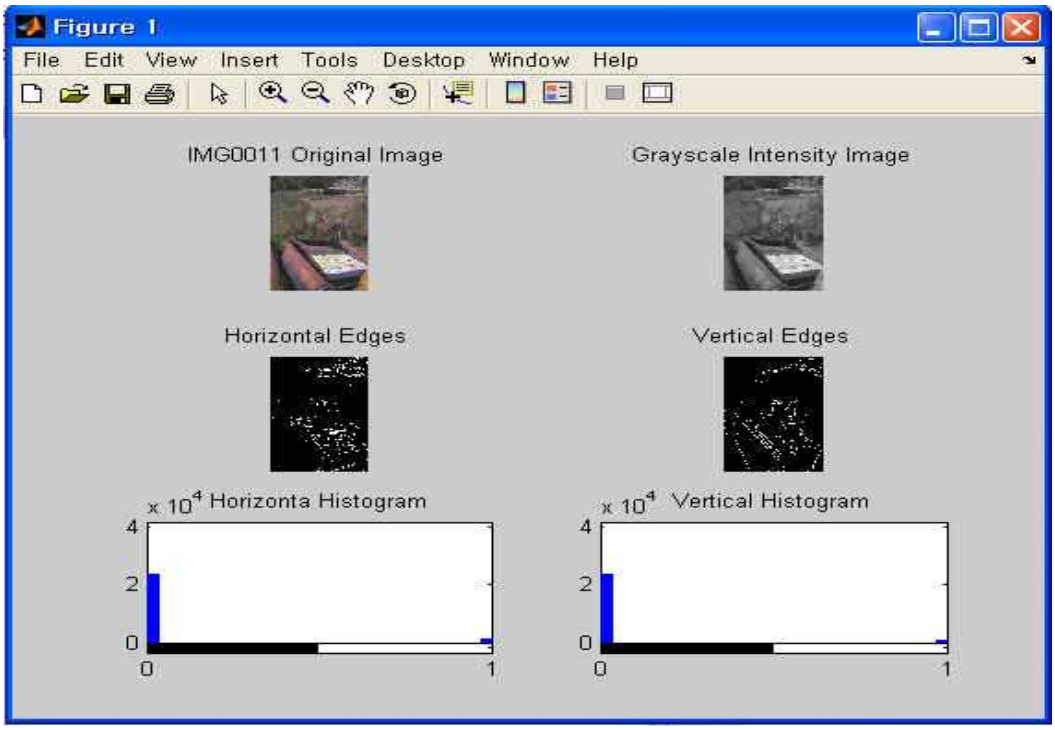


그림 5-34 정제된 에지영상 IMG0011

아래 표 5-35는 IMG0025의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 2040개, 수평 그룹 픽셀 수 187개 평균 픽셀 수 11개 직선 에지 클러스터 수는 85개, 비직선 에지 클러스터 수는 35개로 나타나고, 수직에지인 총 픽셀 수 1617개, 수평그룹픽셀 수 103개, 평균 픽셀 수 16개, 직선에지 클러스터 수는 42개, 비직선 에지 클러스터 수는 31개로 나타낸다.

표 5-35 정제된 에지 영상 IMG0025 기술자 값

Image 0025	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	2040	187	11	85	35
빈 2: Vertical Edges	1617	103	16	42	31

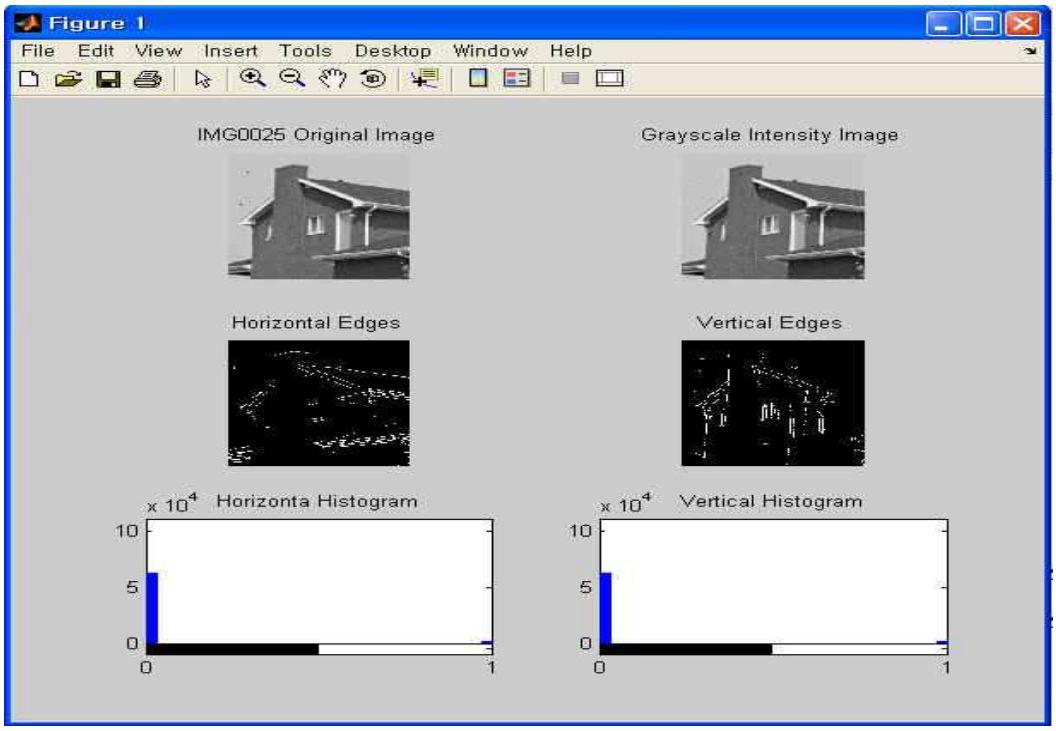


그림 5-35 정제된 에지영상 IMG0025

아래 표 5-36은 IMG0026의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 2039개, 수평 그룹 픽셀 수 188개 평균 픽셀 수 11개 직선 에지 클러스터 수는 82개, 비직선 에지 클러스터 수는 38개로 나타나고, 수직에지인 총 픽셀 수 1640개, 수평그룹픽셀 수 121개, 평균 픽셀 수 14개, 직선에지 클러스터 수는 46개, 비직선 에지 클러스터 수는 37개로 나타낸다.

표 5-36 정제된 에지 영상 IMG0026 기술자 값

Image 0026	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	2039	188	11	82	38
빈 2: Vertical Edges	1640	121	14	46	37

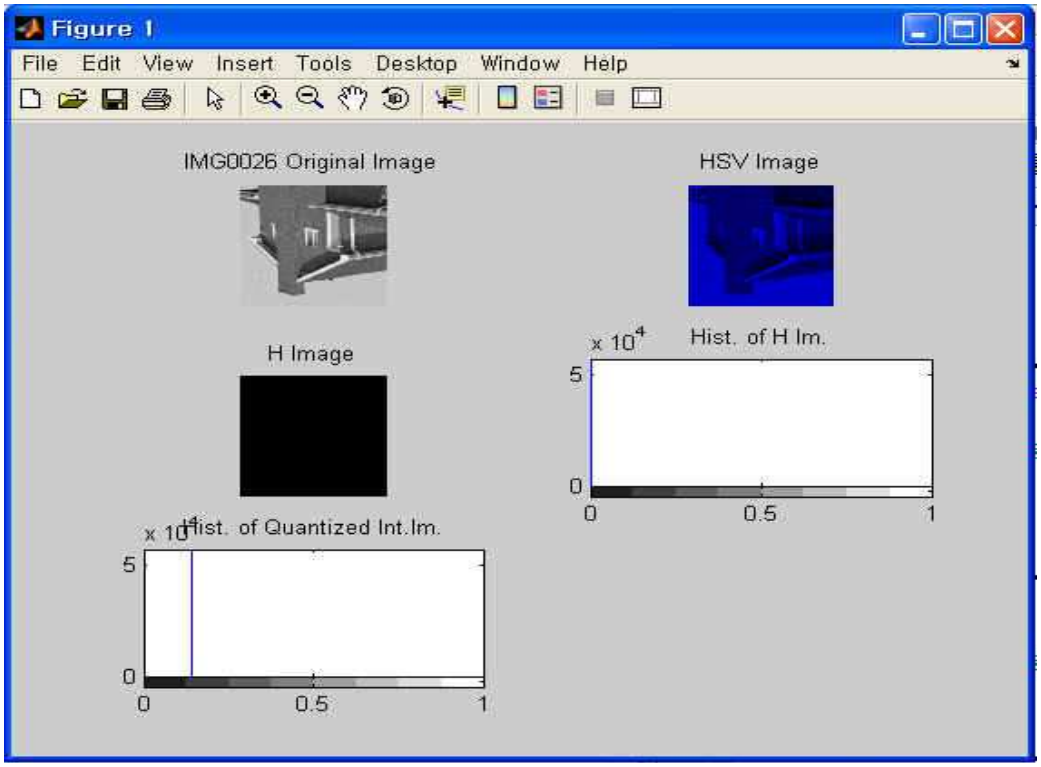


그림 5-36 정제된 에지영상 IMG0026

아래 표 5-37은 IMG0027의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 2041개, 수평 그룹 픽셀 수 182개 평균 픽셀 수 11개 직선 에지 클러스터 수는 86개, 비직선 에지 클러스터 수는 33개로 나타나고, 수직에지인 총 픽셀 수 1626개, 수평그룹픽셀 수 100개, 평균 픽셀 수 16개, 직선에지 클러스터 수는 38개, 비직선 에지 클러스터 수는 31개로 나타낸다.

표 5-37 정제된 에지 영상 IMG0027 기술자 값

Image 0027	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	2041	182	11	86	33
빈 2: Vertical Edges	1626	100	16	38	31

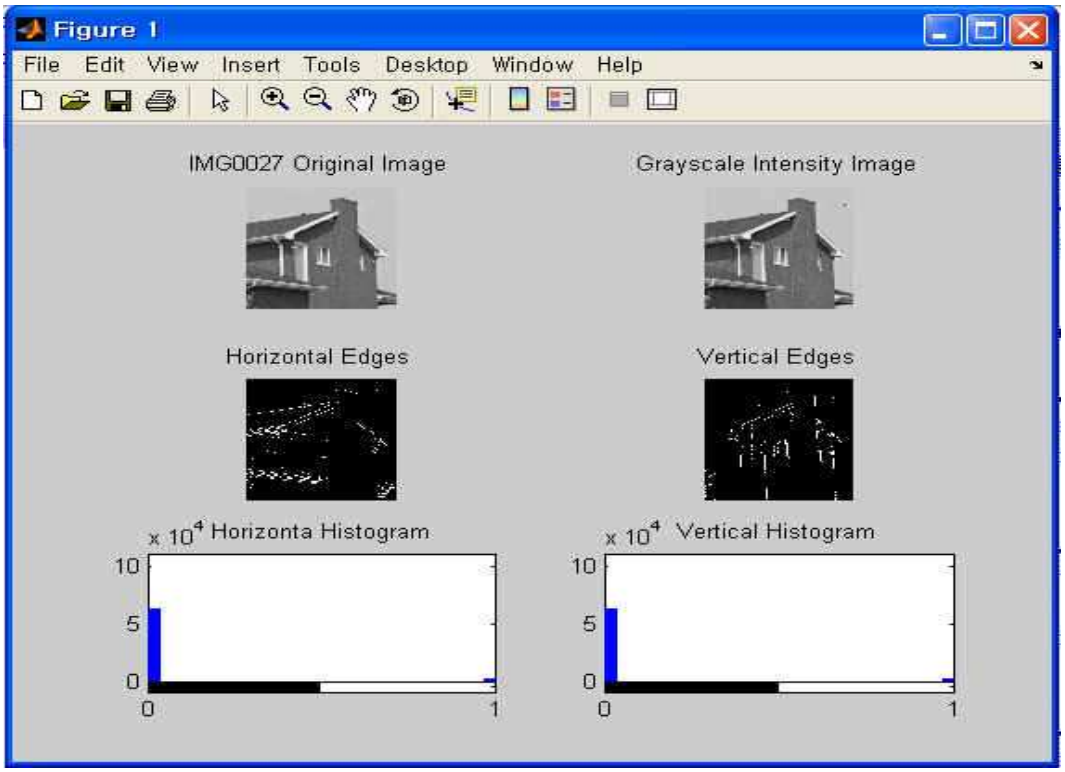


그림 5-37 정제된 에지영상 IMG0027

아래 표 5-38은 IMG0028의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 1990개, 수평 그룹 픽셀 수 185개 평균 픽셀 수 11개 직선 에지 클러스터 수는 77개, 비직선 에지 클러스터 수는 34개로 나타나고, 수직에지인 총 픽셀 수 1621개, 수평그룹픽셀 수 93개, 평균 픽셀 수 17개, 직선에지 클러스터 수는 37개, 비직선 에지 클러스터 수는 32개로 나타낸다.

표 5-38 정제된 에지 영상 IMG0028 기술자 값

Image 0028	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	1990	185	11	77	34
빈 2: Vertical Edges	1621	93	17	37	32

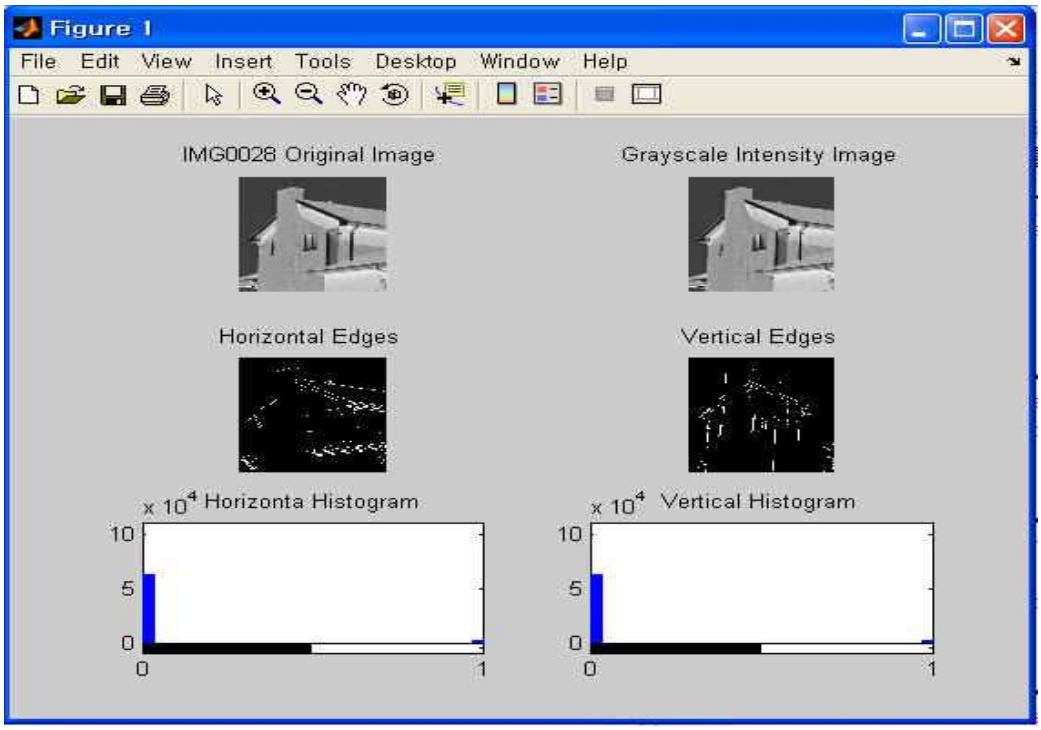


그림 5-38 정제된 에지영상 IMG0028

아래 표 5-39는 IMG0029의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 2642개, 수평 그룹 픽셀 수 407개 평균 픽셀 수 6개 직선 에지 클러스터 수는 107개, 비직선 에지 클러스터 수는 70개로 나타나고, 수직에지인 총 픽셀 수 1835개, 수평그룹픽셀 수 150개, 평균 픽셀 수 12개, 직선에지 클러스터 수는 30개, 비직선 에지 클러스터 수는 63개로 나타낸다.

표 5-39 정제된 에지 영상 IMG0029 기술자 값

Image 0029	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	2642	407	6	107	70
빈 2: Vertical Edges	1835	150	12	30	63

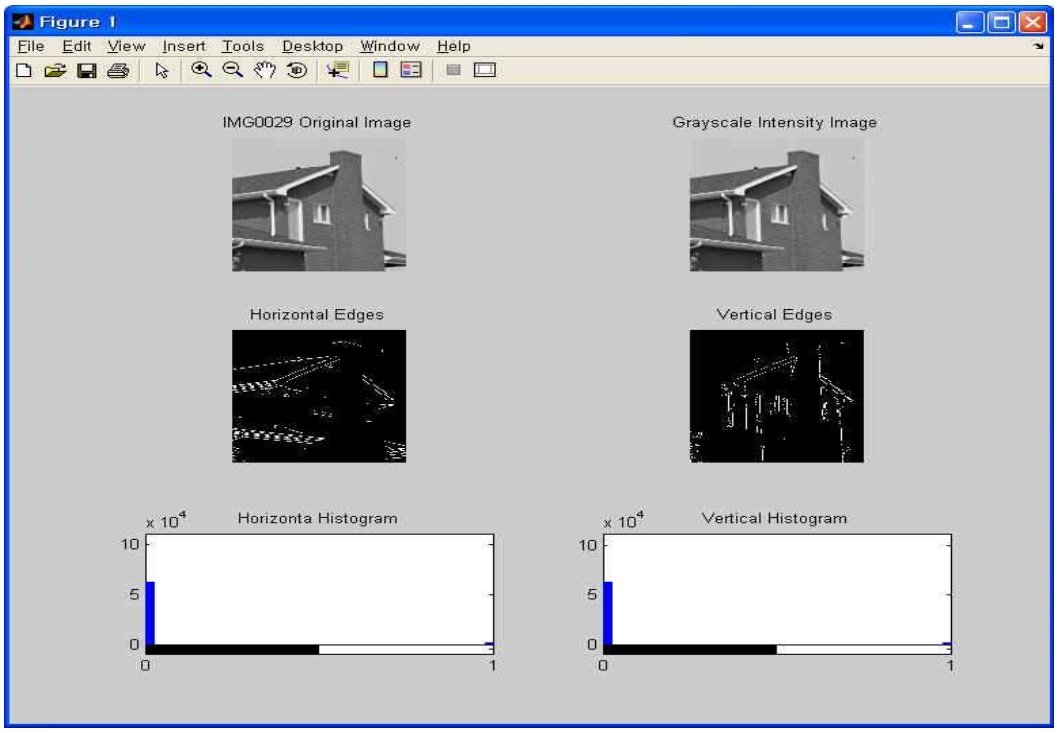


그림 5-39 정제된 에지영상 IMG0029

아래 표 5-40은 IMG0038의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 2038개, 수평 그룹 픽셀 수 1527개 평균 픽셀 수 13개 직선 에지 클러스터 수는 72개, 비직선 에지 클러스터 수는 17개로 나타나고, 수직에지인 총 픽셀 수 1637개, 수평그룹픽셀 수 203개, 평균 픽셀 수 8개, 직선에지 클러스터 수는 81개, 비직선 에지 클러스터 수는 11개로 나타낸다.

표 5-40 정제된 에지 영상 IMG0038 기술자 값

Image 0038	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	2038	152	13	72	17
빈 2: Vertical Edges	1637	203	8	81	11

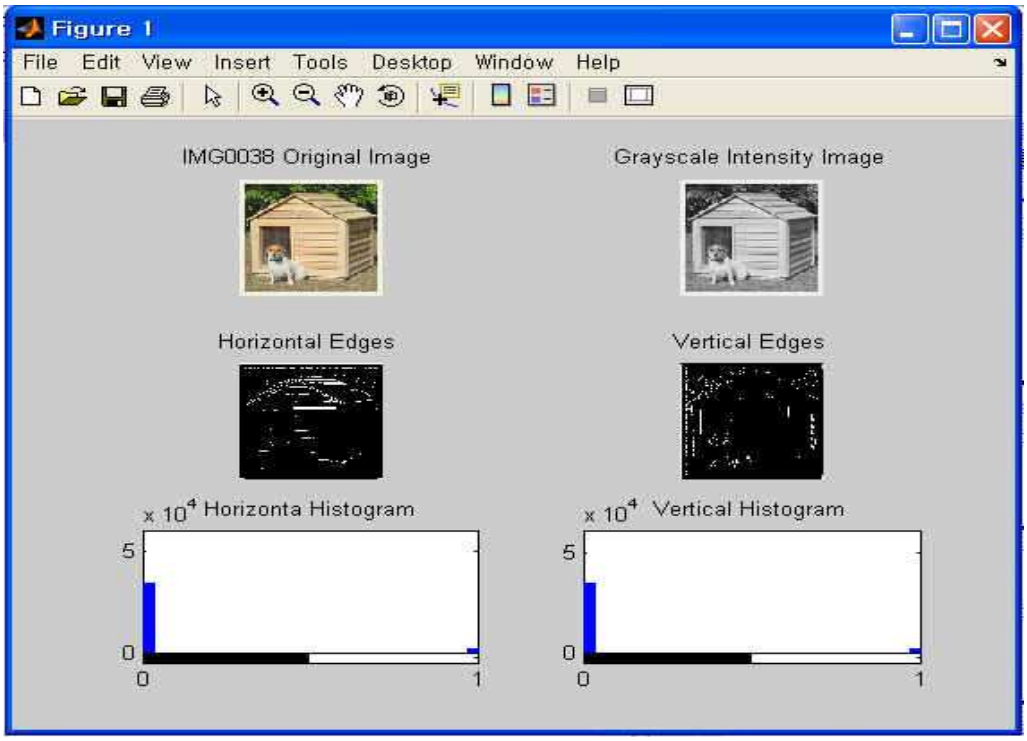


그림 5-40 정제된 에지영상 IMG0038

아래 표 5-41은 IMG0039의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 2043개, 수평 그룹 픽셀 수 166개 평균 픽셀 수 12개 직선 에지 클러스터 수는 82개, 비직선 에지 클러스터 수는 18개로 나타나고, 수직에지인 총 픽셀 수 1660개, 수평그룹픽셀 수 207개, 평균 픽셀 수 8개, 직선에지 클러스터 수는 80개, 비직선 에지 클러스터 수는 12개로 나타낸다.

표 5-41 IMG0038이 좌우반전(IMG0039) 정제된 에지 영상 기술자 값

Image 0039	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	2043	166	12	82	18
빈 2: Vertical Edges	1660	207	8	80	12

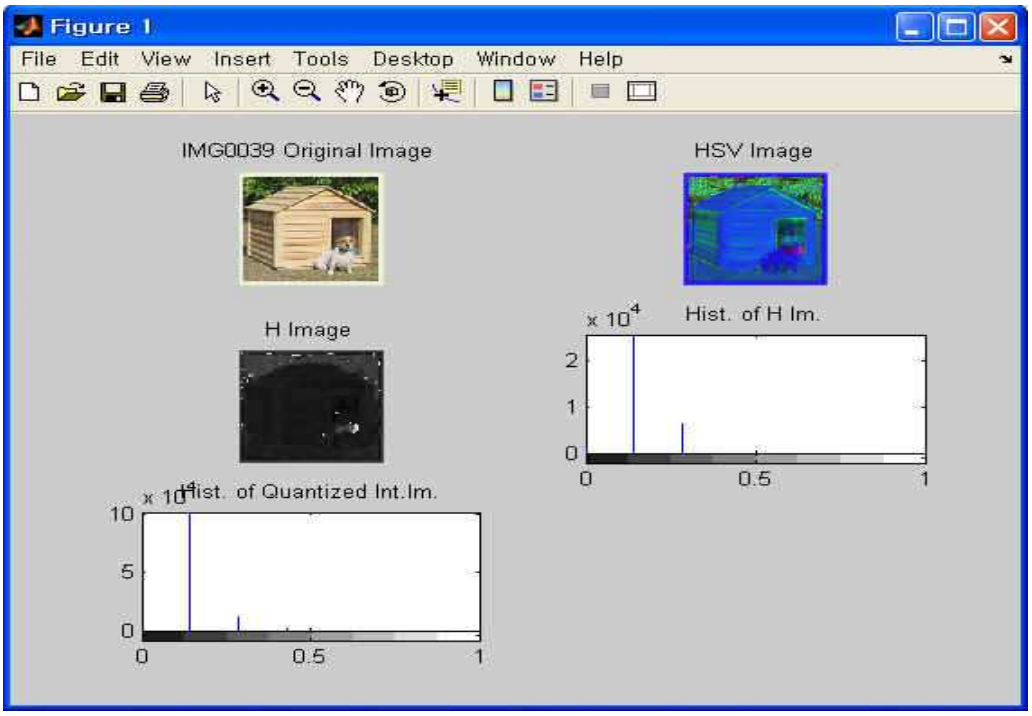


그림 5-41 정제된 에지영상 IMG0039

아래 표 5-42는 IMG0040의 입력영상을 그레이로 변환하여 소벨을 적용 수직과 수평으로 나누어 에지를 정제하기 위해 히스토그램의 값으로 변환 수평에지인 총 픽셀 수 2103개, 수평 그룹 픽셀 수 156개 평균 픽셀 수 13개 직선 에지 클러스터 수는 75개, 비직선 에지 클러스터 수는 17개로 나타나고, 수직에지인 총 픽셀 수 1599개, 수평그룹픽셀 수 199개, 평균 픽셀 수 8개, 직선에지 클러스터 수는 72개, 비직선 에지 클러스터 수는 12개로 나타낸다.

표 5-42 IMG0038이 상하반전(IMG0040) 정제된 에지 영상 기술자 값

Image 0040	αj	Cj	μj	τj	Γj
빈 1: Horizontal Edges	2103	156	13	75	17
빈 2: Vertical Edges	1599	199	8	72	12

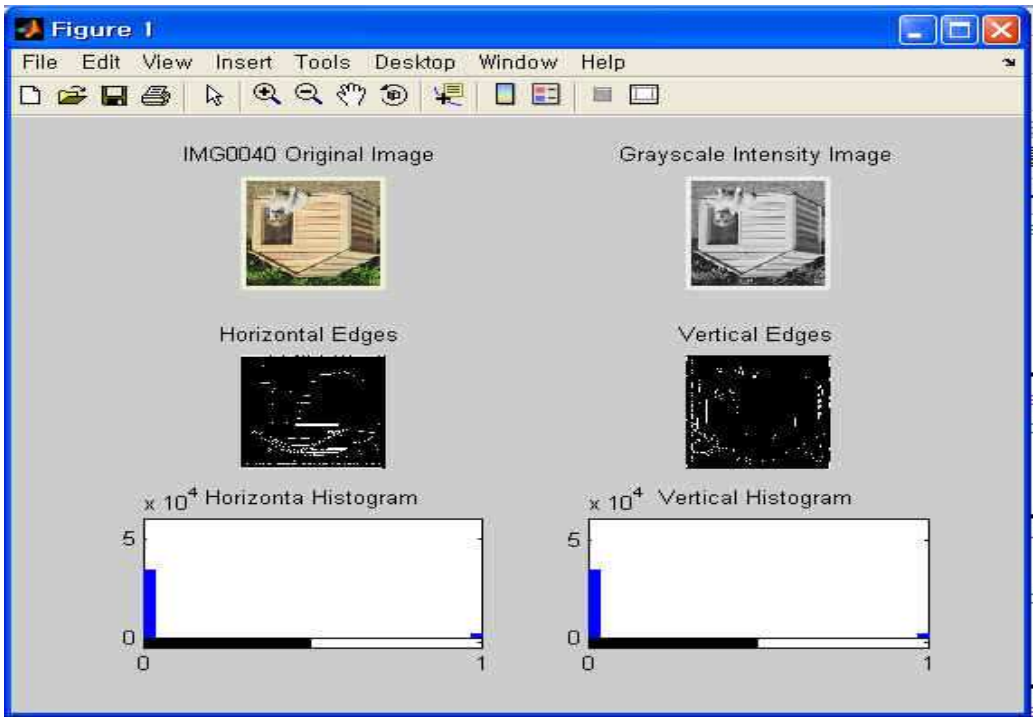


그림 5-42 정제된 에지영상 IMG0040

IMG0001부터 IMG0003번까지 영상은 자연영상으로 에지을 사용하여 수직과 수평을 가지고, 검색 결과 재현율은 기존알고리즘에서는 0.7이였으나, 제안된 알고리즘에서는 재현율은 0.9로 0.2%향상 되었고 정확도에서는 기존과 제안 모두 0.1로 동일하였다, 컬러에지정제 모두를 사용 했을 때는 재현율은 0.9와 정확도는 0.09로 모두 사용 했을 때 탁월한 검색능력을 보였다. IMG0005에서 IMG0011영상에서는 기존알고리즘에서는 0.88이였으나 제안된 알고리즘에서는 0.97로서 0.09정도 향상되었고 정확도에서는 모두 0.1로서 동일하였고, 모두를 사용 했을 때는 재현율은 1, 정확도는 0.1로 더욱더 향상된 검색 결과를 확인 했다. IMG025부터 IMG0040번에서는 상하 좌우 회전 및 영상을 늘려 수직과 수평을 가지고 검색한 결과 기존 및 제안알고리즘 모두재현율은 0.9이고 정확도에서는 0.1로 모두 동일하였다 컬러 에지 정제를 모두 사용 시는 재현율은 0.97과 정확도는 0.11로서 모두를 사용한 더욱더 좋은 결과를 보였다

3. 컬러 및 에지 정제 방법에 따른 시뮬레이션 결과 및 성능 분석

표 5-43 컬러 및 에지 정제 방법에 따른 시뮬레이션 및 성능 분석

Method / Simulation	제안한 알고리즘						기존알고리즘	
	컬러 정제 방법		에지 정제 방법		컬러 에지 정제 방법			
	재현율	정확도	재현율	정확도	재현율	정확도	재현율	정확도
IMG001	0.8	0.08	0.8	0.08	0.8	0.08	0.6	0.06
IMG002	0.8	0.08	1.0	0.10	1.0	0.10	0.8	0.08
IMG003	0.8	0.10	0.8	0.10	0.8	0.10	0.8	0.10
IMG004	1.0	0.10	1.0	0.10	1.0	0.10	0.8	0.08
IMG005	1.0	0.08	1.0	0.08	1.0	0.08	1.0	0.08
IMG006	0.8	0.16	0.9	0.18	1.0	0.20	0.8	0.16
IMG007	0.9	0.16	1.0	0.18	1.0	0.18	0.8	0.14
IMG008	1.0	0.10	1.0	0.10	1.0	0.10	0.8	0.08
IMG009	1.0	0.08	1.0	0.08	1.0	0.08	1.0	0.08
IMG0010	0.9	0.14	0.9	0.14	1.0	0.16	0.75	0.12
IMG0011	1.0	0.06	1.0	0.06	1.0	0.06	1.0	0.06
IMG0012	1.0	0.08	1.0	0.08	1.0	0.08	1.0	0.08
IMG0013	1.0	0.12	1.0	0.12	1.0	0.12	0.8	0.10
IMG0014	1.0	0.08	1.0	0.08	1.0	0.08	1.0	0.08
IMG0015	0.9	0.12	0.9	0.12	1.0	0.14	0.9	0.12
IMG0016	0.9	0.14	0.9	0.14	1.0	0.14	0.75	0.12
IMG0017	1.0	0.06	1.0	0.06	1.0	0.06	1.0	0.06
IMG0018	0.9	0.16	0.9	0.16	1.0	0.18	0.9	0.16
IMG0019	0.8	0.18	0.9	0.20	0.9	0.20	0.8	0.18
IMG0020	1.0	0.08	1.0	0.08	1.0	0.08	1.0	0.08
평 균	0.92	0.11	0.95	0.11	0.98	0.12	0.87	0.10

표 5-43은 본문에서 제안한 컬러 에지 정제법과 기존 알고리즘의 성능비교 결과이다. 공간 정보를 이용한 컬러 및 정제 방법은 거리에 따라 성능 값이 다르지만 거리를 변경하여 실험하여 본 결과 2로 하였을 때 좋은 검색 성능을 보이는 것으로 확인 되었다.

표 5-44 기존 알고리즘과 제안한 알고리즘의 비교결과

Method \ Measure		재현율	정확도
기존알고리즘		0.87	0.1
제안된알고리즘	컬러정제	0.92	0.11
	에지정제	0.95	0.11
	컬러에지정제	0.98	0.12
	평균	0.95	0.11

표 5-44에서는 기존의 알고리즘의 재현율 & 정확도의 평균값과 제안한 알고리즘의 재현율 & 정확도 평균값을 비교해 보았다. 제안한 알고리즘이 더 좋은 성능을 보인 것이 확인 되었다.

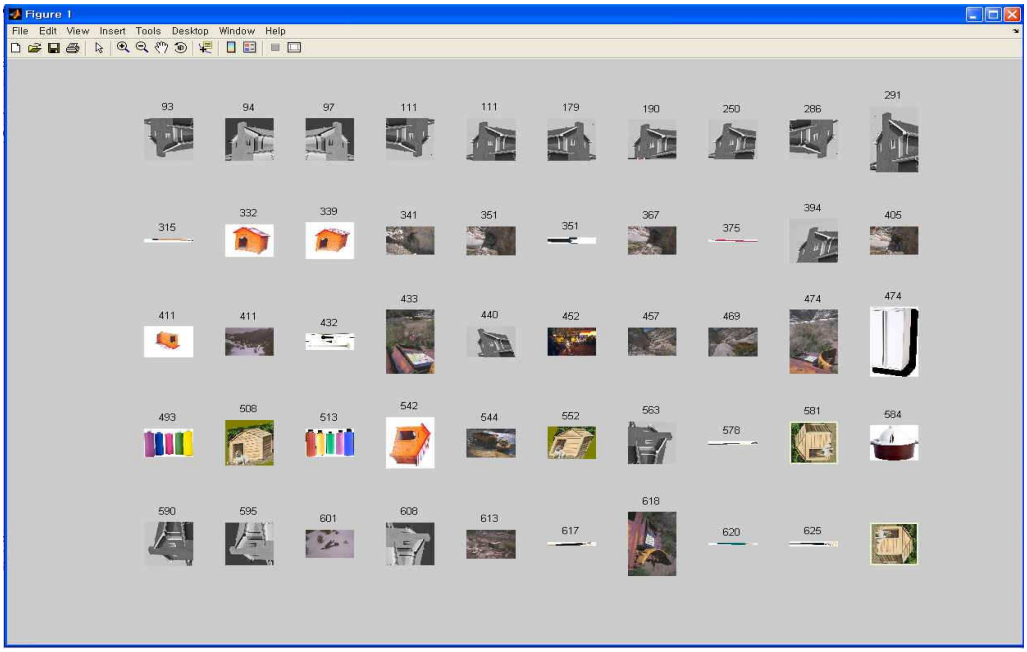


그림 5-43 재현율 & 정확도 성능 분석 이미지 1

그림 5-43은 입력 영상을 데이터베이스에 삽입하여 제안한 컬러에지정제에 기반한 알고리즘 사용 재현율과 정확도를 분석한 영상이다.

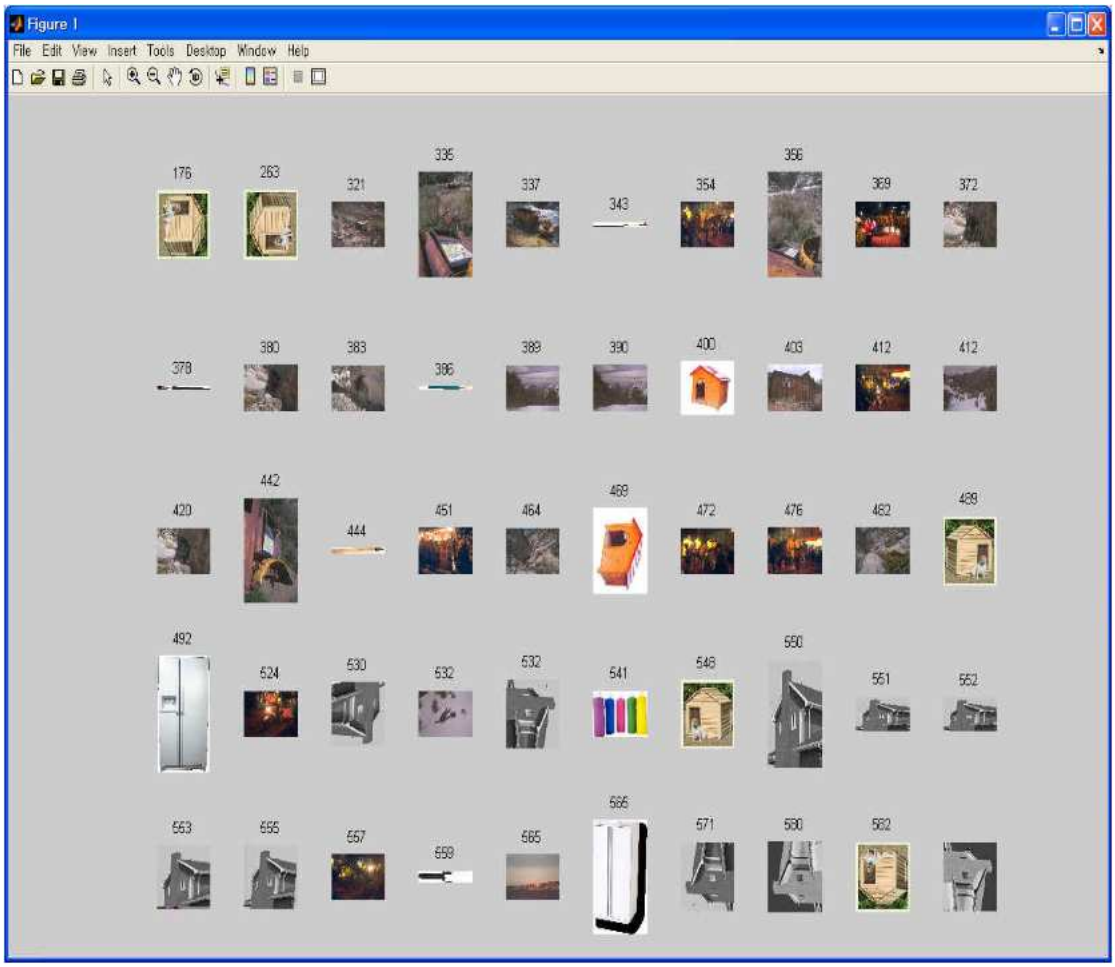


그림 5-44 재현율 & 정확도 성능 분석 이미지 2

그림 5-44은 입력 영상을 회전하여 데이터베이스에 삽입하여 제안한 컬러에지 정제에 기반한 알고리즘 사용 재현율과 정확도를 분석한 영상이다.

Ⅵ. 결 론

본 논문에서는 개선된 컬러와 에지 정제에 기반한 영상 검색 알고리즘을 설계하였고 시뮬레이션을 통하여 그 타당성을 입증하였다. 시뮬레이션에서는 데이터베이스 3419개의 영상을 사용하였으며 쿼리 영상으로 입력영상은 자연 컬러 영상 24개 그레이 하우스영상 13개 강아지 영상7개 회전된 컬러 하우스 영상 7개를 선정하여 시험하였으며 그 결과, 다음과 같은 타당성과 효율성을 확인 하였다.

-. 개선된 컬러 및 에지 정제에 기반한 영상 검색 융합형 알고리즘을 구현함으로서 기존의 영상 검색 성능 보다 효율성이 높은 검색 시스템을 구현하였다.

-. 컬러와 에지 기반을 융합형 검색 알고리즘을 구현하여 융합 가중치를 조정함으로서 회전, 밝기, 이동, 크기 등 잡음에 강인한 성능을 얻을 수 있었다.

-. 컬러 및 에지 정제에 기반한 알고리즘을 이용하여 입력 영상과 데이터베이스 3419개의 영상을 비교하였을 때 동일 영상에서는 완벽하게 검색하였으며 유사 영상순으로 추출됨을 확인하였다.

-. 주어진 동일 환경에서 기존의 컬러/에지 정제 기반 알고리즘의 재현율은 87% 이였으며, 정확도의 값은 0.1 인 반면, 제안된 컬러 및 에지 정제 기반 알고리즘의 재현율은 95% 이고, 정확도는 0.11 로서 그 우수성을 확인하였다.

-. 대부분의 내용기반 영상 검색 시스템들은 단일한 특징만을 검색에 사용하지 않고 컬러, 에지, 질감, 모양 등의 여러 가지 특징 벡터들을 결합한 형태를 사용하고 있다, 그런 면에서 제안한 방법의 정량적 수치의 우수성과 함께 내용기반 영상 검색 분야에서 컬러와 에지 정제가 효과적으로 사용될 수 있는 방법을 제시했다는 점에서 충분히 연구의 가치를 찾을 수 있다고 하겠다.

참고문헌

[1] M. Flickner et al. "Query by image and video content: The QBIC system", IEEE computer, Vol. 28(9), pp. 23-32, 1995.

[2] Virginia Ogle and Michael Stonebraker, "Chabot : Retrieval from a relational database of images", IEEE computer, Vol. 28(9), pp. 40-48, 1995.

[3] Greg Pass and Ramin Zabih, "Histogram Refinement for content - based image retrieval", In IEEE Workshop on Applications of Computer Vision, pp. 96-102, 1996.

[4] M. J. Swain and D. H. Ballard, "Indexing via color histograms, in Image Understanding Workshop", Pittsburgh, PA, pp. 623-630, 1990.

[5] Nagasaka A. and Tanaka. Y, "Automatic video indexing and full video search for object appearances" In Knuth, Wegner (Eds.), IFIP Trans., Visual Database Systems II, pp. 113-127, 1992.

[6] Grosky W and Mehrotra R., "Image Database Management, in Advances in Computers", Academic Press, New York, Vol. 34, pp. 237-291, 1992

[7] James Z. Wang, Gio Wiederhold, Oscar Firschein and Sha Xin Wei, ``Content-based image indexing and searching using Daubechies' wavelets'' International Journal of Digital Libraries(IJODL), Springer-Verlag, Vol. 1(4), pp. 311-328, 1998.

[8] John R. Smith and Shih-Fu Chang, "VisualSEEK: a fully automated content-based image query system", Boston, Massachusetts, United States ACM Multimedia'96, pp. 87-98, 1996.

[9] Markus Stricker and Alexander Dima. "Color indexing with weak spatial constraints", SPIE proceedings, pp. 2670:29-40, 1996.

[10] Haijing Wang, Peihua Li and Tianwen Zhang, "Proposal of Novel Histogram Features for Face Detection", ICAPR (2) 2005, pp. 334-343, 2005.

[11] Mohan S. Kankanhalli, Babu M. Mehtre and Ran Kang Wu, "Cluster-based color matching for image retrieval", Pattern Recognition, Vol. 29(4), pp. 701-708, 1996.

[12] Jing Huang, S. Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih, "Image indexing using color correlograms", In IEEE Conference on Computer Vision and Pattern Recognition, pp. 762-768, 1997.

[13] Mrinal K. Mandal, Tyseer Aboulnasr and Sethuraman Panchanathan, "Fast Wavelet Histogram Techniques for Image Indexing", Computer Vision and Image Understanding, Vol. 75(1-2), pp. 99-110, 1999.

[14] V. Buzuloiu, M. Ciuc, R.M. Ranayya and C. Vertan, "Adaptive-neighborhood histogram equalization of color images", J. Electron. Imaging, Vol. 10(2), pp. 445-459, 2001.

[15] Rick Rickman and John Stonham, "Content based image retrieval using color tuple histograms", SPIE proceedings, pp. 2670:2-7, 1996.

[16] Jianping Fan and David. K.Y.Yan, "Automatic Image Segmentation by Intergrating Color-Edge Extraction and Seeded Region Growing", IEEE Trans. on Image Processing, Vol.10, No.10, pp. 1454-1466, 2001.

[17] R Albero Salinas, C. Richardson, M. A. Abidi and R. C. Gonzalez, "Data Fusion : Color Edge Detection and Suface Reconstruction through Regularization", IEEE Tans. Ind Elec., Vol.43, No.3, pp. 355-363, 1996.

[18] S.Di Zenzo, "A Note on the Gradient of a Multi-Image", CVGIP, Vol.33, No.1, pp. 116-125, 1986.

[19] P.E.Tranhanias and A.N.Venetsanopoulos, "Vector Order-statistics Operators

as Color Edge Detectors", IEEE Trans on Sys. Man and Cyb-B, Vol.26, No.1, pp. 135-143, 1996.

[20] M. Safar, C. Shahabi and X. Sun, "Image retrieval by shape: a comparative study", ICME 2000, Vol. 1, pp. 141-144, 2000.

[21] A. Khotanzad and Y.H. Hong, "Invariants Image Recognition by Zernike Moments ", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.12, No.5, pp. 489-497, 1990.

[22] A.K. Jain and A. Vailaya, "Image Retrieval using Color and Shape", Pattern Recognition, Vol. 29, No. 8, pp. 1233-1244, 1996.

[23] C. Harris and M. Stephens, "A combined corner and edge detector", Alvey Vision Conference, pp. 147-151, 1988.

[24] R. Deriche and G. Giraudon, "A Computational Approach for Corner and Vertex Detection", Int'l Journal of Computer Vision, Vol. 10:2, pp. 101- 124, 1993.

[25] C. Schmid, R. Mohr and C. Bauckh age, "Comparing and evaluating inter est points", International Conference on Computer Vision, pp. 230-235, 1998.

[26] D. Androutsos , K.N. Plataniotis, and A.N. Venet sanopoulos, "A Novel Vector-Based Approach to Color Image Retrieval Using a Vector An gular-Based Distance Measure", Computer Vision and Image Under standing, Vol. 75, NO. 1/2, pp. 46-58, 1999.

[27] C. Schmid and R. Mohr. "Matching by Local Invariants, "Technical report N2644, INRIA, 1995.

[28] C. Schmid and R. Mohr. "Local Grayvalue Invariants for Image Retrieval, "IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 5, pp. 530-535, 1997.

[29] J . Canny, "A Computational Approach to Edge Detection", IEEE Trans. on Pattern Recognition and Machine Intelligence, Vol. 8, pp. 679-698, 1986.

[30] Wynne Hsu, T. S. Chua and H. K. Pung. "An integrated color-spatial approach to content based image retrieval", In ACM Multimedia Conference, pp. 305-313, 1995.

[31] Michael Swain and Dana Ballard. "Color indexing. International Journal of Computer Vision", Vol. 7(1), pp. 11-32, 1991.

[32] Muhammad Bilal Ahmad, Jong-An Park and Tae-Sun Choi, "Modified chain code and statistical invariant based shape matching for MPEG-7", IEEE International Symposium on Consumer Electronics, Erfurt Germany, pp. 23-26, 2002.

[33] C.K.Yang and W.H.Tsai, "Reduction of Color Space Dimensionality by Moment-preserving Thresholding and its application for edge-detection in Color Images", Pattern Recognition Letters, Vol. 17, No. 5, pp. 481-490, 1996.

[34] Sami Brandt, Jorma Laaksonen, and Erkki Oja, "Statistical Shape Features in Content-based Image Retrieval", Proc. of 15th Internation Conference on Pattern Recognition, Barcelona, Spain, Vol. 2, pp. 6062, 2000.

[35] A. Gupta and R. Jain, "Visual Information Retrieval", Comm.Assoc. Comp. Mach. Vol. 40, No. 55, pp. 70-79, 1997.

[36] ISO/IEC JTC1/SC29/WG11 : "Information Technology-Multimedia Content Description Interface-Part 5 : Multimedia Description Schemes", MPEG document N4005, Singapore, 2001.

[37] ISO/IEC JTC1/SC29/WG11/p186 : "Proposal for MPEG-7 evaluation : Still Image DS", Lancaster, 1999.

[38] Jong-An Park, Min-Hyuk Chang, Tae Sun Choi and Muhammad Bilal Ahmad, "Histogram based chain codes for shape description", IEICE Trans. On Communications, Vol. E86-B, no.12, pp. 3662-3665, 2003.

[39] I. K. Park, I. D. Yun and S. U. Lee, "Color Image Retrieval Using Hybrid Graph Representation", 제9회 영상처리 및 이해에 관한 워크샵, pp. 197-292, 1997.

[40] H. Y. Jin, H. Y. Lee, H. S. Kim and Y. H. Ha, "Content-based Image Retrieval Using Color Adjacency and Gradient", 대한전자공학회, pp. 157-160, 2000.

[41] T. H. Kim, T. H. Hwang, J. W. Shin and D. S. Jeong, "Image Retrieval Technique Using Entropy and Color Features", 제10회 신호처리종합학술대회, pp. 1073-1076, 1997.

감사의 글

망설임과 설레는 마음을 멀리하고 불혹의 나이에 대학원생활을 시작한 학문이였기에 더욱더 고심하고 아파했던 순간들과 숱한 일들이 주마등처럼 지나갑니다

어려운 삶 속에서도 항상 정직해야 한다는 마음을 간직하면서도 여전히 불효로 남아 있는 소생에게 오늘의 건강한 삶이 있게 해주시고 세상을 떠나 계신 아버지, 어머니께 이 논문을 바칩니다.

학문의 길을 열어 좋은 결실을 맺을 수 있도록 깊은 가르침을 주신 박종안 지도 교수님께 고개 숙여 감사를 드립니다. 그리고 학위과정 중 학문적 지도와 세심한 보살핌을 주신 전자정보통신공학부 교수님들께 진심으로 감사드립니다.

어려움이 있을 때마다 충고와 격려를 아끼지 않으셨던 광주인력개발원 박홍주 원장님, 홍천인력개발원 이성호 원장님 김대균박사, 문광명, 장성렬 교수님께 감사 드립니다. 또 항상 가까운 곳에서 따뜻한 배려와 힘을 주었던 많은 분들께 항상 고마움을 기억하고자 합니다.

논문이 완성될 때까지 따뜻한 애정으로 동고동락을 같이 하며 많은 협조와 성원을 주었던 DSP Lab의 선・후배님께 진심으로 고마움을 전합니다.

오늘이 있기까지 늘 건강을 염려하시어 사랑으로 보살펴 주신 형제분들 그리고 물심양면으로 용기를 북돋아 주시고 많은 정성과 사랑을 주신 처가 형제분들께도 마음속 깊이 감사 드립니다.

어려운 생활 속에서도 늘 사랑과 격려로 뒷바라지를 해준 사랑하는 아내와 씩씩한 장한 두 아들 준영, 동현이와 기쁨을 함께 하고자 합니다.

끝으로 이 작은 결실을 학자의 길로 가는 첫걸음이라 생각하고 더욱 정진하여 연구 노력할 것을 다짐합니다.

2008 년 6 월

박 태 수 드림.

저작물 이용 허락서

학 과	정보통신	학 번	20047371	과 정	박사
성 명	한글 : 박 태 수 한문 : 朴 泰 洙 영문 : PARK TAE SOO				
주 소	경기도 고양시 일산서구 대화동2081성저건영빌라711동302호				
연락처	E-MAIL : cyberbus@korcham.net				
논문제목	한글 : 컬러와 에지 정제에 기반한 잡음불변 영상 검색 영문 : Noise-invariant Image Retrieval based on Color and Edge Refinement				

본인이 저작한 위의 저작물에 대하여 다음과 같은 조건아래 조선대학교가 저작물을 이용할 수 있도록 허락하고 동의합니다.

- 다 음 -

1. 저작물의 DB구축 및 인터넷을 포함한 정보통신망에의 공개를 위한 저작물의 복제, 기억장치에의 저장, 전송 등을 허락함.
2. 위의 목적을 위하여 필요한 범위 내에서의 편집·형식상의 변경을 허락함. 다만, 저작물의 내용변경은 금지함.
3. 배포·전송된 저작물의 영리적 목적을 위한 복제, 저장, 전송 등은 금지함.
4. 저작물에 대한 이용기간은 5년으로 하고, 기간종료 3개월 이내에 별도의 의사표시가 없을 경우에는 저작물의 이용기간을 계속 연장함.
5. 해당 저작물의 저작권을 타인에게 양도하거나 또는 출판을 허락을 하였을 경우에는 1개월 이내에 대학에 이를 통보함.
6. 조선대학교는 저작물의 이용허락 이후 해당 저작물로 인하여 발생하는 타인에 의한 권리 침해에 대하여 일체의 법적 책임을 지지 않음.
7. 소속대학의 협정기관에 저작물의 제공 및 인터넷 등 정보통신망을 이용한 저작물의 전송·출력을 허락함.

2008 년 6 월 25일

저작자 : 박 태 수 (서명 또는 인)

조선대학교 총장 귀하