

칼라 히스토그램 기반 영상 검색을 위한 효율적인 칼라 특징 정보 추출 기법

정희원 이호영*, 김영태**, 김희수*, 배태면*, 하영호*

Efficient Color Feature Information Extraction Method for Color Histogram-based Image Retrieval

Ho-Young Lee*, Young-Tae Kim**, Hee-Soo Kim*, Tae-Meon Bae* and Yeong-Ho Ha*

Regular Members

요약

칼라 히스토그램 방법은 영상 획득 시 광원의 위치나 광원색의 변화에 따라서 칼라 분포가 변하게 되어, 동일한 내용을 담고 있는 영상이라 할 지라도 검색 효율이 떨어지는 단점을 가진다. 그래서 본 논문에서는 이러한 문제점을 해결하기 위해 조명의 변화가 배제된 칼라 정보를 추출한 후, 이에 대한 칼라 히스토그램을 이용한 영상 검색 방법을 제안한다. 제안한 방법은 생동폭 조절(dynamic range control)을 통하여 광원의 위치 변화 때문에 발생하는 음영(shadow)으로 인한 색 분별력(color discrimination) 저하를 방지하였으며, 물체의 확산 반사(diffuse reflection) 성분만을 이용한 gray world 가정 기반의 광원색 제거 방법으로 조명 조건 변화가 배제된 칼라 정보를 추출하였다. 실험 결과, 추출된 칼라 정보를 이용한 히스토그램 방법이 조명의 영향이 배제되지 않은 입력 영상의 칼라 정보를 이용한 히스토그램 방법보다 높은 검색 효율을 보임을 확인할 수 있었다.

ABSTRACT

Color distribution is changed according to the variation of illumination position and illumination color. Therefore, even if images are relevant each other, retrieval accuracy is degraded. In this paper, we propose the image retrieval method using color information excluded illumination component. The proposed dynamic range control method removes the shadow region generated by change of illumination position to increase the color discrimination power. To exclude the illuminant color, we use the diffuse reflection component of object and gray world assumption. The experimental results show that the color histogram method using color information excluded illuminant has higher retrieval accuracy than conventional color histogram using the color information of input image.

1. 서론

급격한 컴퓨터의 성능 향상과 고속 정보 통신망의 보급으로 인하여 멀티미디어 정보의 증가는 기하급수적으로 증가하고 있다. 문자, 음성, 영상, 비

디오 등의 여러 멀티미디어 정보 가운데에서 영상과 비디오 정보가 차지하는 정보량은 다른 미디어와는 비교가 되지 않을 정도로 방대하여, 디지털 영상 및 비디오 데이터를 효율적으로 전송, 저장, 처리하기 위한 연구는 지금까지도 중요한 관심사가

* 경북대학교 전자전기공학부 칼라 및 영상처리 연구실(hylee@m80.knu.ac.kr),

** 한국전자통신연구원 방송미디어 연구부

논문번호 : 00183-0523, 접수일자 : 2000년 5월 23일

되고 있다. 현재까지는 주로 디지털 영상 및 비디오 데이터 양을 효율적으로 전송, 저장하기 위한 영상 압축 기술에 대한 연구가 집중되었으나, 지금에 와서는 디지털 형태로 압축된 영상 및 비디오 데이터를 효율적으로 활용하는 방법 즉, 압축 저장된 수많은 영상 및 비디오 데이터 중에서 사용자가 원하는 특별한 형태의 비디오 및 영상들을 빠른 시간 내에 찾아내는 방법 등에 대한 연구가 큰 관심을 끌고 있다. 이러한 최근의 관심 변화와 시장 요구를 바탕으로 하여 국제 표준화 기구인 ISO와 IEC의 연합 기술위원회 산하의 MPEG은 영상의 압축에 대한 표준화인 MPEG-1, MPEG-2, 그리고 MPEG-4에 대한 표준화 완료 후에 멀티미디어 내용기술 인터페이스(multimedia content description interface)라는 이름으로 멀티미디어 데이터의 내용기반 검색을 위한 내용 표현 방식에 관한 국제 표준화 작업인 MPEG-7이 진행 중에 있다¹⁷⁾.

멀티미디어 데이터는 다양한 형태로 구성되어 있고, 데이터의 크기가 방대하기 때문에 MPEG-7에서는 효율적인 멀티미디어 데이터의 내용 기술을 위해 의미론적인(semantic) 내용을 기술하는 상위 레벨 방법과 칼라, 형태, 위치 등 하위 레벨의 내용 표현 기술을 이용한 방법을 통한 두 가지 방향의 접근 방법을 이용하는데, 하위 레벨 방법의 경우 사용자 질의 구성이 손쉬운 뿐만 아니라 자동적인 검색이 손쉬운 반면, 상위 레벨의 방법의 경우 데이터의 내용 표현에 조금 더 자연스러운 장점이 있지만 자동 검색의 한계가 있어서 사용자의 개입을 필요로 하는 단점을 가진다¹⁶⁾. 본 논문에서는 하위 레벨의 내용 기술 방법 중에서 영상 및 비디오 데이터의 관측 위치, 크기 변화 등에 무관하며, 영상 전체의 특징을 비교적 잘 표현하는 칼라 정보를 이용한 방법에 대해 연구하였다. 칼라를 이용한 방법은 영상내의 물체(object)의 이동, 회전등에 영향을 적게 받을 뿐만 아니라 방법이 간단하여 적은 연산량으로 정확한 영상 검색이 가능하므로 가장 많이 사용되어지는 방법이다. 대표적인 칼라 특징(feature)을 이용한 방법으로는 칼라 히스토그램(color histogram)¹¹⁾이 있는데, 이는 각 칼라 채널의 값을 일정한 비율로 나누어 히스토그램 구간(bin)을 결정한 후, 유사한 칼라끼리 군집화 하는 방법으로서 대부분의 영상에 대해 비교적 정확한 검색 결과를 보여주고 있다.

그러나, 컬러 히스토그램을 이용한 방법은 영상내의 조명 변화에 민감할 뿐만 아니라, 칼라 분포의

공간적인 정보를 이용하지 못한다는 단점 또한 가지고 있다. 두 가지 단점 가운데에서 조명 조건(illumination condition) 변화에 의한 단점은 서로 다른 광원 위치(illumination position)와 서로 다른 광원색(illumination color)에 의해 영상 전체에 대한 칼라 분포가 변하게 되어 동일한 내용을 가지는 영상이라 할지라도 서로 다른 영상으로 검색되는 경우이며, 칼라 분포의 공간적인 정보를 이용하지 못하는 경우는 서로 다른 내용을 담고 있는 영상임에도 불구하고 칼라 분포가 동일하여 서로 유사성이 있는 영상(relevant image)으로 검색되는 경우이다. 조명 조건 변화 가운데에서 광원의 위치가 크게 다른 경우에는 동일한 내용을 담고 있는 영상이라 할지라도 서로 다른 음영(shading) 영역이 존재하게 되어 영상 전체의 칼라 분포를 크게 변화시키는 경우이며, 서로 다른 광원색으로 인한 영향은 주로 실 내에서 획득된 영상에 주로 나타나는데, 실내의 조명 조건에 따라 동일한 물체이지만 광원의 서로 다른 색성분에 의해 영상 전체의 칼라 분포가 크게 변하게 되는 경우이다. 따라서, 본 논문에서는 칼라 히스토그램 기반의 영상 검색에서 광원의 영향이 최대한으로 배제된 칼라 정보를 추출하는 알고리즘을 제안하였다. 기존의 광원의 영향을 제거하기 위한 알고리즘들은 주로 광원 위치가 동일하다는 가정 하에서 서로 다른 유색 광원의 영향을 배제하는데 중점을 두고 있을 뿐만 아니라, 그 연산 과정 또한 매우 복잡하여 수많은 영상 및 비디오 데이터에서 사용자가 원하는 영상 데이터에 대해 빠른 검색 속도를 요구하는 응용 분야에 적합하지 않은 것들이 대부분이다^{2-9, 10, 11, 12, 13)}. 그래서 본 논문에서 조명 조건 변화로 인한 칼라 분포 변화를 줄이기 위한 방법에 있어서의 주안점은 기존 칼라 히스토그램의 장점인 빠른 수행 속도와 물체의 이동, 회전에 대한 강인성을 최대한 유지하면서 적은 연산으로 효과적으로 서로 다른 조명 조건의 영향을 줄이는데 역점을 두었다. 또한 서로 다른 유색광원으로 인한 칼라 분포뿐만 아니라, 실제 영상에서 자주 발생할 수 있는 광원의 위치 변화로 인한 칼라 분포의 변화도 고려하여 보다 나은 검색 효율을 가질 수 있도록 하였다.

본 논문의 구성은 II장과 III장에서 광원의 영향과 제안한 칼라 정보 추출 알고리즘에 대해 각각 설명한 후, IV장에서 제안한 칼라 정보 추출 알고리즘을 이용한 실험 결과를 살펴보고, V장에서 결론을 맺는다.

II. 광원의 영향

2.1 광원 조사 방향 변화의 영향

일반적으로 카메라로 획득되어진 영상에서의 각 채널 밝기값은 아래 식과 같이 표현되어진다.

$$a_k^p = \int_{\lambda} (n_s \cdot n^p) e(\lambda) s^p(\lambda) Q_k(\lambda) d\lambda \quad k=R, G, B \quad (1)$$

여기서 $e(\lambda)$ 는 광원의 분광 조성(spectrum)을, $s^p(\lambda)$ 는 위치 p 에서의 표면 반사 계수를, Q_k 와 a_k^p 는 각각 카메라 RGB 채널에서의 스펙트럼 반응과 카메라의 출력, 그리고, n_s 와 n_p 는 광원의 방향 벡터(illumination direction vector)와 표면 법선 벡터(surface normal vector)를 각각 나타낸다. 위치 p 에서 카메라 채널의 스펙트럼 반응, 표면 반사 계수, 광원 스펙트럼이 모두 동일하다 할지라도 광원의 방향 벡터가 서로 다른 경우에는 RGB 각 채널의 밝기값 변화가 발생함을 식(1)에서 확인할 수 있다. 즉, 광원의 위치가 변하게 되면, 광원 방향 벡터가 변하게 되어 동일한 물체라 할지라도 밝기값 변화가 크게 발생할 수 있다. 광원의 방향 변화는 주로 영상의 무채색 영역인 음영 영역을 결정하게 되는데, 광원 방향의 변화가 아주 크게 되면 영상 전체에서 음영 영역이 차지하는 비율이 크게 차이가 나게 되어 영상 검색 효율이 떨어지게 되는 것이다. 그림 1은 광원 방향 변화가 발생한 예를 나타낸다. 그림 1에서 볼 수 있듯이 동일한 물체를 가진 영상들 가운데에서 영상 획득 과정에서 광원 위치 변화로 인해 음영(shading) 정도 차이가 큰 영상들이 자주 발생할 수 있음을 확인할 수 있다.

2.2 광원색 변화의 영향

광원의 위치 변화이외에 영상의 칼라 분포 변화에 큰 영향을 미치는 또 다른 요인으로는 유색 광원의 영향을 들 수 있다. 인간이 지각하는 색은 물체에 조사된 광원(illumination) 성분과 물체 고유의 색과 특성을 나타내는 표면 반사(surface reflectance)성분이 서로 곱하여져 눈의 망막에 전달되어 색을 느끼고 구분하게 된다. 그러므로 동일한 물체라 할지라도 조사되는 광원 성분이 변하게 되면 기존의 RGB 좌표계로 표현할 경우 서로 다른 색으로 표현되게 된다. 이러한 영상들에 대하여 기존의 영상 검색 방법을 이용할 경우에는 동일한 내용(content)을 담고 있는 영상이라 할지라도 서로 다



그림 1. 광원 방향 변화

른 영상으로 판단되게 되어 검색 효율이 떨어지게 된다. 그러나 인간은 주관적인 색채지각으로 인한 색 항상성(color constancy)을 가지고 있어서, 광원의 종류 및 광원에 관련된 주변환경이 달라져도 물체색의 변화를 거의 느끼지 못한다. 즉, 인간은 자연광, 백열등 또는 형광등의 여러 가지 광원 아래에서도 흰색을 항상 회색 느끼는 항상성을 가진다고 할 수 있다. 이러한 인간 시각의 주요 특징인 색 항상성을 기반으로 하여 광원의 색을 결정짓는 광원 스펙트럼에 관계없이 물체 고유의 색을 복원하기 위한 방법들이 많이 연구되어져 왔다^{2-9, 10, 11, 12, 13}. 그러나 기존의 색 항상성 알고리즘들은 정확한 색 재현이 목적이므로 많은 연산량을 수반하는 경우가 많다. 특히 Maloney¹²의 방법은 파장 영역에서 기본 함수의 선형 조합으로 광원 성분을 추정하여 광원 성분을 제거하기 때문에 밝기값을 기반으로 하는 대부분의 영상 검색 방법에는 적합하지 않다. 영상의 밝기값을 기반으로 하는 Jobson¹⁷ 방법의 경우 주위 함수(surround function)의 상수값(C) 조절을 통해 색 항상성뿐만 아니라 색표현 및 생동 폭까지도 조절할 수 있지만, 영상 검색의 경우에는 광원 성분의 제거뿐만 아니라 정확한 색표현이 중요하므로 검색에 쓰이기 위해서는 주위 함수의 상수값이 60이상은 되어야 한다. 가우시안 함수

(gaussian function)와의 상승 적분(convolution)을 공간 영역에서 수행하기 위해서는 창크기(window size)가 가우시안 함수 표준 편차의 $6\sqrt{2}$ 배 이상 되어야하는데^[9], 가우시안 주위 함수(gaussian surround function)의 경우 표준 편차가 $C/\sqrt{2}$ 이므로 실제적인 창 크기는 $6C$ 이상이어야 한다. 즉, 한 화소의 위치에서 결과값을 얻기 위해 사용되어지는 곱셈의 수가 $36C^2$ 가 되고, 영상의 크기가 234×234 일 경우 총 $54,756 \times 36 \times C^2$ 번 이상의 곱셈 연산이 필요로 하게 된다. 그러므로, Jobson의 방법을 영상 검색에 이용하는 것은 수행 속도 면에서 매우 어렵다고 할 수 있다. 또 다른 밝기값 기반의 방법인 Von Kries 가설^[6]을 이용한 색순응(chromatic adaptation)의 경우에도 정확한 색순응이 이루어지기 위해서는 다른 조영 조건인 2개의 영상이 필요하므로 영상마다의 독립적인 멀티미디어 내용을 기술할 수 없다는 단점이 있다. 실제로 칼라를 특징값으로 하여 영상 검색을 수행할 때, 광원 스펙트럼의 차이로 인해 동일한 내용을 담고 있는 영상들간에 칼라 분포의 차이가 발생하는 경우가 많은데 그림 2는 이러한 영상들의 예를 보여준다. 그림에서 볼 수 있듯이 두 영상 모두 동일한 실내공간을 배경으로 획득된 영상이나 서로 다른 분광 조성을 가지는 광원의 영향으로 흰 바다부분의 색이 서로 다를 수 있다.

III. 제안한 칼라 정보 추출

3.1 생동폭 조절(dynamic range control)을 통한 광원 조사 방향의 영향 제거

앞절에서 살펴보았듯이 광원의 위치 변화로 인해 나타나는 가장 큰 영향은 무채색 계열인 음영 영역의 발생으로 볼 수 있다. 무채색(achromatic color) 계열인 음영 영역이 광원의 위치 변화로 인해 많이 발생하는 경우에는 동일한 물체를 가진 서로 관련성이 깊은 두 영상간에 정확한 색차의 표현이 불가능하게 되어 검색 효율이 떨어지게 된다. 가장 많이 사용되어지고 있는 칼라 히스토그램의 경우, 유채색(chromatic color) 계열인 경우에는 RGB 채널별로 일정한 히스토그램 구간(bin)으로 나누어 유사한 칼라별로 군집화하기 때문에 색차를 비교적 정확하게 표현할 수 있지만, 무채색 계열인 경우에는 정확한 색차 표현의 불가능으로 인한 색 분별력(color discrimination) 저하로 검색 효율이 떨어지게 된다.

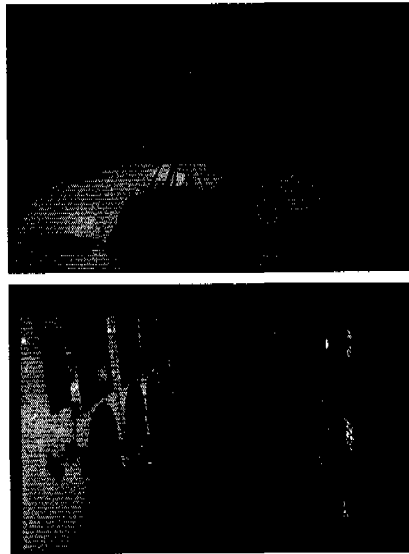


그림 2. 서로 다른 유색 광원으로 인한 칼라 분포 변화

이러한 광원의 위치 변화로 인한 검색 효율의 저하를 방지하기 위해서는 광원의 위치 변화로 인한 음영 영역을 제거하여 무채색 계열의 화소의 수를 낮추어 주는 것이 필요하다.

그래서 본 연구에서는 영상의 생동폭 조절을 통해 광원의 위치 변화로 인한 음영 영역을 효과적으로 제거하는 알고리즘을 제안하였다. 아주 어두운 무채색 계열인 음영 영역이 많이 발생하게 되면, 일반적으로 영상의 생동폭(dynamic range)은 넓어지고 영상의 가시도(visibility)는 떨어지게 된다. 그러므로 영상에서 어두운 영역의 밝기값을 높여서 생동폭을 줄이는 것이 음영 영역에 의한 영향을 제거할 수 있을 뿐만 아니라 영상의 전체 가시도를 향상시킬 수 있다. 따라서 본 연구에서는 각 RGB채널의 밝기값을 μ -law 스케일링 함수^[10]를 기반으로 조절하여 어두운 무채색 계열인 음영 영역의 밝기값을 높여 영상의 생동폭을 줄이면서 가시도를 향상시켰다. μ -law 함수에 대한 수식적 표현은 다음과 같다.

$$y = y_{max} \frac{\ln[1 + \mu(x/x_{max})]}{\ln(1 + \mu)} \quad (2)$$

여기서 y 는 스케일링된 출력값을 x 는 입력값을 나타내며, y_{max}, x_{max} 는 최대 출력과 입력값을 각각 나타낸다. 그림 3은 μ 값의 변화에 따라 최대 입력값으로 정규화된 μ -law 함수의 곡선을 나타낸다. 그림 3에서 볼 수 있듯이 μ 값이 증가할수록 어두운 영역의 밝기값이 크게 증가하고, 상대적으로

밝은 영역에 비해 어두운 영역의 밝기값 변화가 큰 것을 확인할 수 있는데, 이러한 성질은 음영 영역을 효과적으로 제거할 뿐만 아니라 영상의 가시도(visibility)를 증가시키는 역할을 하게 된다. 그러나, 음영 영역을 제거하기 위해 아주 큰 μ 값을 적용하는 경우에는 영상의 색 표현이 부자연스러울 뿐만 아니라, 색의 왜곡(color distortion)이 발생하는 경우가 많이 발생하게 된다. 실제로 weber 법칙에 따르면 인간 시각의 JND(just noticeable difference)는 밝기값에 비례하여 어두운 영역의 대조 민감도($1/JND$) 값이 높아지게 되는데, 이러한 성질로 인해 어두운 영역의 작은 밝기값 변화는 오히려 인간 시각에 거슬리는 역효과를 나타내게 된다. 그러므로, 효과적인 음영 영역의 제거를 위해서는 각 영역의 특성에 따라 적용적으로 μ 값을 설정하는 것이 바람직하다.

이를 위하여 본 연구에서는 RGB 채널의 밝기값의 평균($(R+G+B)/3 \leq Th$)으로 영상의 임의의 위치에서의 휘도(luminance)를 구한 후, 영역의 밝고 어두움에 따라 주위 화소와의 상관 관계를 고려하여 μ 값을 결정하였다. $(R+G+B)/3 \leq Th$ 인 경우에는 어두운 음영 영역의 화소로, $(R+G+B)/3 > Th$ 인 경우에는 밝은 영역의 화소로 간주하였다. 어두운 음영 영역의 화소의 경우에는 인간 시각이 대조 민감도가 높으므로 색 왜곡을 방지하기 위해 주변 8개의 이웃 화소의 평균 휘도와와의 상관 관계를 고려하여 식(3)과 같이 μ 값을 결정하였다.

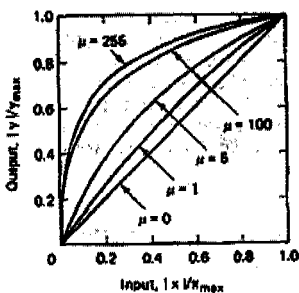


그림 3. μ -law 함수

$$\begin{aligned} &\text{if}(\text{Brightness} < \text{Brightness}_{\text{local mean}}) \mu = \mu_1 \\ &\text{else} \mu = \mu_2 \end{aligned} \quad (3)$$

여기서 $\text{Brightness}_{\text{local mean}}$ 은 주위 8 이웃화소의 평균 휘도를 나타낸다.

그리고, $(R+G+B)/3 > Th$ 인 밝은 화소의 경우에

는 인간 시각의 대조 민감도가 어두운 영역보다 낮으므로 주위 화소와의 상관 관계 고려 없이 $\mu = \mu_3$ 으로 고정하였으며, 각 μ 값의 대소 관계는 $\mu_1 > \mu_2 > \mu_3$ 이다. μ_1, μ_2, μ_3 값은 실험적으로 결정하였으며, 색의 왜곡을 방지하기 위해 RGB 채널에 대해 동일한 μ 값을 적용하였다. 그림 4는 제한한 생동폭 조절 알고리즘의 적용 결과를 나타내는데 원영상에 존재하는 많은 음영 영역들이 생동폭 조절 알고리즘에 의해 가시도(visibility)가 향상되어 색분별력(color discrimination)이 향상됨을 확인할 수 있다.

3.2 광원색 영향의 제거

1) 색 항상성(color constancy)

인간 시각의 주관적인 색채지각인 색항상성은 다양한 광원 하에서의 색을 표준 광원이 조사되었을 때의 색으로 변화시키는 과정이라고 볼 수 있는데, 광원색의 영향을 제거하여 물체 고유의 색만으로 표현하기 위해서는 표준 광원의 정의가 필수적이다. 국제 조명 학회(CIE)에서는 표준 광원을 일반 태양 광원 D65로 정의하고 있으며, 이를 주파수 영역에서의 분광 조성으로는 정확하게 표현이 가능하지만 일반적인 영상 검색에서 사용되어지는 RGB좌표계에서 표준 광원을 각 채널의 밝기값으로 표현하는 것은 어렵다. 그래서 광원이나 표면 반사 계수의 정확한 분광 조성(spectral composition)에 대한 정보 없이 광원에 독립적인(illumination invariant) 물체색을 인식하기 위한 방법으로 물체색의 통계적인 분포에 대한 가정을 이용한 방법들이 색 항상성 분야에 많이 연구되어지고 있다^[12]. 광원색에 의한 칼라 분포 변화를 방지하기 위해 광원 변화로 인한 각 채널 밝기값에 대한 영향을 살펴보면 다음과 같다. 채널 밝기값에 대한 모델링인 식(1)에서 각 채널의 스펙트럼 반응이 협대역(narrow band)이라면 각 채널의 스펙트럼 반응은

$$Q_k(\lambda) \approx Q_k(\lambda_k) \delta(\lambda - \lambda_k) \quad k=R, G, B \quad (4)$$

이므로, 카메라 출력은 다음과 같다.

$$\begin{aligned} q_k^c &= (n_s \cdot n^p) e(\lambda_k) s^p(\lambda_k) Q_k(\lambda_k) \\ &= \tau_k s^p(\lambda_k) Q_k(\lambda_k) \end{aligned} \quad (5)$$

여기서 $\tau_k = (n_s \cdot n^p) e(\lambda_k)$ 이다.

광원의 방향 벡터와 광원의 스펙트럼이 각각 n_s

와 $e'(\lambda)$ 로 변화된 동일 물체에 대한 카메라의 출력력을 ρ'_k 이라 하면

$$\rho'_k = \frac{\tau_k}{\tau_k} \rho_k^0 \quad (6)$$

이다. 즉, 광원의 변화는 RGB 각 채널의 값을 독립적으로 변화시켜 식(7)과 같이 표현할 수 있다.

$$\begin{bmatrix} \rho'_R \\ \rho'_G \\ \rho'_B \end{bmatrix} = \begin{bmatrix} \tau_R/\tau_R & 0 & 0 \\ 0 & \tau_G/\tau_G & 0 \\ 0 & 0 & \tau_B/\tau_B \end{bmatrix} \begin{bmatrix} \rho_R \\ \rho_G \\ \rho_B \end{bmatrix} \quad (7)$$



그림 4. 제한한 생동폭 조절 결과

2) 제한한 광원색 제거 알고리즘

일반적으로 각 채널에서 밝기값은 아래와 같이 광원과 표면 반사 계수의 곱으로 표현되어진다.

$$I(x, y) = S(x, y)r(x, y) \quad (8)$$

광원의 변화가 발생하게 되면 입력으로 주어지는 밝기값 $I(x, y)$ 는 동일한 표면 반사 계수를 가진다 할지라도 다른 값을 가지게 되므로, 광원의 영향을 받지 않기 위해서는 광원 성분 또는 표면 반사 계

수를 추정하여야 한다. 광원 성분을 추정하는 방법으로는 대표적으로 Maloney^[2]의 방법을 들 수 있으며, 대부분의 방법이 광원 성분을 광원 주변 함수에 대한 기본 함수들의 선형 조합으로 나타내는 것이다. 그러나, 광원 성분의 분광 조성을 찾아낸다는 것은 채널 밝기값만을 이용하여서는 정확한 광원 성분을 찾아내기는 불가능할 뿐만 아니라, 광원 성분의 분광 조성을 표준 광원의 분광 조성으로 변환하는데도 많은 어려움이 있다. 표면 반사 계수를 추정하는 방법으로 물체색의 통계적인 분포에 대한 가정을 이용한 이론들이 많이 있으며, 그 가운데 Buchsbaum^[13]은 영상 전체에서의 평균 표면 반사 계수는 그레이(gray)이며, 그 값은 중간 그레이(middle gray)값들의 평균 표면 반사 계수와 같다(gray world assumption)라는 가정을 발표하였다. 즉, 영상에서의 평균 표면 계수는 각 채널 화소값의 평균이며, 그 값은 대부분의 영상들에서 전체 그레이값 범위의 중간값에 해당한다는 것이다.

본 연구에서도 직접적인 광원 성분의 추정 없이 gray world 가정을 이용하여 표면 반사 계수의 정규화(normalization)를 통해 광원의 영향을 제거하였다. 표면 반사 계수 $r(x, y)$ 는 동일하지만 광원 성분 $S(x, y)$ 가 다른 두 영상에서 각 채널 평균값은 영상의 평균 표면 반사 계수를 나타내게 된다. 그러나, 두 영상에서 각 채널 평균값은 실제 표면 반사 계수는 동일하지만 광원의 영향에 의해 밝기값이 달라지므로 평균값은 달라지게 된다. 그래서 본 연구에서는 광원의 영향을 받은 표면 반사 계수의 보정을 위해 gray world 가정인 영상 평균 표면 반사 계수는 중간 그레이값의 평균 반사 계수와 같다는 것을 이용하였다. 중간 그레이값의 평균 반사 계수를 128로 보고, 중간 gray값의 평균 반사 계수와 채널 평균값과의 비를 이용하여 각 채널의 밝기값 변환 상수($\tau_R/\tau_R, \tau_G/\tau_G, \tau_B/\tau_B$)를 결정하였다. 즉, 영상의 평균 반사 계수가 중간 그레이 값의 평균 반사 계수를 갖도록 화소값을 변화시키는 것이다. 이는 복잡한 광원 추정의 과정을 거치지 않으므로 연산이 아주 간단할 뿐만 아니라 효과적으로 광원 성분을 제거할 수 있다. 아래 수식은 제한한 광원색 제거 알고리즘을 나타낸다.

$$\begin{bmatrix} \rho'_R \\ \rho'_G \\ \rho'_B \end{bmatrix} = \begin{bmatrix} 128/R_{avg} & 0 & 0 \\ 0 & 128/G_{avg} & 0 \\ 0 & 0 & 128/B_{avg} \end{bmatrix} \begin{bmatrix} \rho_R \\ \rho_G \\ \rho_B \end{bmatrix} \quad (9)$$

여기서 $R_{avg}, G_{avg}, B_{avg}$ 는 각 채널에서의 화소 평균값이다.

3) 적응적 변환 상수 결정

영상의 각 채널에서 평균 표면 반사 계수를 구하기 위해 채널의 평균값을 구할 때, 전체 영상이 아주 어두운 경우에는 평균값의 크기가 작으므로 화소값 변환 상수의 크기가 너무 커져서 실제 물체색을 제대로 표현하지 못하게 된다. 마찬가지로 영상이 아주 밝은 경우도 평균값의 크기가 너무 커서 결과 영상의 밝기가 너무 어둡게 되는 문제를 발생시키게 된다. 영상의 아주 어두운 부분의 화소는 표면 반사가 거의 없는 무채색 계열의 색이며, 아주 밝은 부분은 거울 반사(specular reflection)가 일어났다고 볼 수 있다. 앞서 언급한 문제점들을 제거하기 위해서는 무채색 계열의 화소들과 거울 반사가 일어난 부분은 채널 평균의 계산에서 제외되어야 한다. 이를 위하여 본 연구에서는 shafer^[14]의 이색성 반사 모델(dichromatic reflection model)을 이용하였다. Shafer는 표면 반사 계수를 확산 반사(diffuse reflection)와 거울 반사(specular reflection)로 나타내었는데, 그림 5는 확산 반사와 거울 반사의 예를 보여준다.

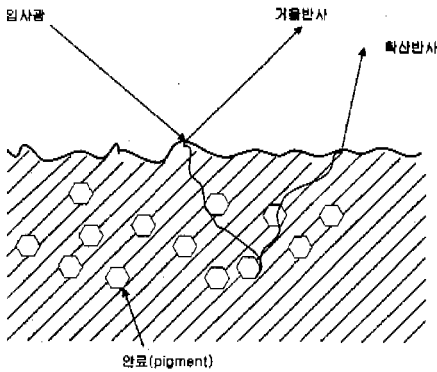


그림 5. 확산 반사와 거울 반사

확산 반사는 물체의 표면을 통과하여 물체의 내부에서 반사와 회절을 거친 후, 나타나는 반사 성분으로서 물체색을 결정하며, 거울 반사는 물체의 표면에서 바로 반사되어 나타나는 성분으로 반사가 모든 파장 성분에 대해 동일하게 나타나 대부분 흰색(white) 계열의 무채색을 가지게 된다. 그러므로 물체색을 결정하는 물체의 표면 계수를 구하기 위해서는 확산 반사 성분만을 고려하여야 된다. 이색성

반사 모델을 수식으로 표현하면 아래와 같이 주어진다^[15].

$$L(\lambda, a, \theta) = L_s(\lambda, a, \theta) + L_d(\lambda, a, \theta) \tag{10}$$

$$= m_s(a, \theta)C_s(\lambda) + m_d(a, \theta)C_d(\lambda)$$

여기서 θ 와 a 는 그림 6에서처럼 표면 법선 벡터와 광원 방향 벡터사이의 각과 시각 위치와 반사 성분사이의 각을 나타낸다.

카메라 출력 RGB 밝기값은 모든 가시광선 파장에 걸쳐 식(10)을 적분하게 되는데, 적분 결과는 식(11)과 같이 확산 반사와 거울 반사 성분으로 나누어 나타낼 수 있다. 일반적으로 거울 반사 성분은 대부분 무채색인 밝은 흰색에 가까우므로 식(11)에서 거울 반사 성분을 상수로 가정하면, 이색성 반사 모델의 수식인 식(10)은 식(12)와 같이 표현할 수 있다.

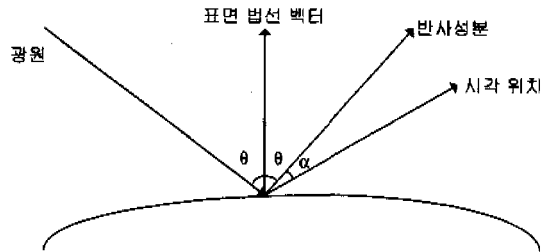


그림 6. 이색성 반사 모델

$$C_L = m_d C_d + m_s C_s \tag{11}$$

$$= m_d \begin{pmatrix} r_d \\ g_d \\ b_d \end{pmatrix} + m_s \begin{pmatrix} r_s \\ g_s \\ b_s \end{pmatrix}$$

$$= \begin{pmatrix} m_d r_d + m_s r_s \\ m_d g_d + m_s g_s \\ m_d b_d + m_s b_s \end{pmatrix}$$

$$C_L = \begin{pmatrix} m_d r_d + a \\ m_d g_d + a \\ m_d b_d + a \end{pmatrix} \tag{12}$$

식 (9)에서 각 채널의 밝기값 변환 상수($128/R_{avg}$, $128/G_{avg}$, $128/B_{avg}$)를 결정하는데 있어서 문제가 되는 무채색과 거울 반사 성분을 제거하기 위해 본 연구에서는 각 화소 위치에서의 최대 밝기값과 최소 밝기값의 차이를 이용하였다.

$$\begin{aligned}
 I &= \text{MAX}(r, g, b) - \text{MIN}(r, g, b) \\
 &= \text{MAX} \begin{pmatrix} m_d r_d + \alpha \\ m_d g_d + \alpha \\ m_d b_d + \alpha \end{pmatrix} - \text{MIN} \begin{pmatrix} m_d r_d + \alpha \\ m_d g_d + \alpha \\ m_d b_d + \alpha \end{pmatrix} \quad (13) \\
 &= \text{MAX} \begin{pmatrix} m_d r_d \\ m_d g_d \\ m_d b_d \end{pmatrix} - \text{MIN} \begin{pmatrix} m_d r_d \\ m_d g_d \\ m_d b_d \end{pmatrix}
 \end{aligned}$$

일반적으로 무채색의 경우에는 각 채널 밝기값의 차이가 아주 작으므로, 최대 밝기값과 최소 밝기값과의 차이를 구하여 무채색과 유채색을 구분하였으며, 거울 반사 성분은 흰색의 무채색에 가까우므로 차이값을 구하는 과정에서 서로 상쇄되어 확산 반사 성분만이 남게 된다. 확산 반사 성분의 유채색과 무채색 구별을 위해 본 연구에서는 *KTH* 이면 무채색의 확산 반사 성분이라 가정하고, 평균 반사 계수를 구하기 위한 채널 평균의 계산에서 제외하였다.

제안한 광원색 제거 알고리즘의 성능을 평가하기 위해 동일 물체에 대해 붉은 계열의 광원과 푸른 계열의 광원을 조사한 영상에 대해 실험을 수행하였다. 그림 7은 실험에 사용된 영상을 나타내는 데, 동일한 물체임에도 불구하고 시각적으로 느끼는 칼라 분포 차이가 많이 나는 것을 확인할 수 있다. 그림 8은 유채색의 확산 반사 성분만을 고려하지 않고, 각 채널의 모든 밝기값을 이용하여 광원 성분을 제거한 영상이다. 영상이 어두워서 평균값의 크기가 작아 영상의 물체색이 제대로 복원되지 않았음을 확인할 수 있다. 그림 9는 유채색의 확산 반사 성분만을 고려하였을 경우의 결과 영상으로 광원 성분이 완전히 제거되어 칼라 분포가 거의 동일함을 확인할 수 있다. 그림 10은 서로 다른 광원 성분 가진 그림 4에 대한 결과 영상을 나타내는데, 제안한 알고리즘을 적용시킨 결과 광원 성분이 제거되어 광원 성분에 관계없이 실내 흰색의 바다 부분이 제대로 복원됨을 확인할 수 있다.

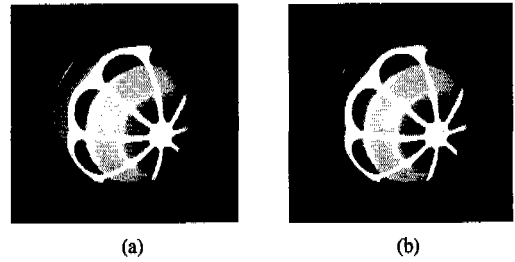


그림 8. 모든 화소를 고려한 경우의 결과 영상
(a) 붉은 광원의 영상 (b) 푸른 광원의 영상

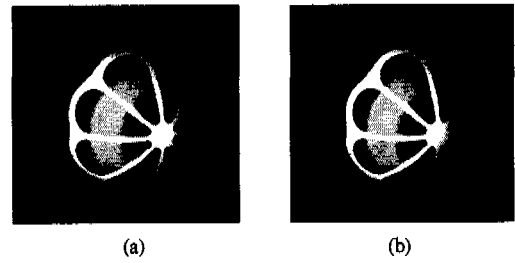


그림 9. 유채색 확산 반사만을 고려한 경우의 결과영상
(a) 붉은 광원의 영상 (b) 푸른 광원의 영상

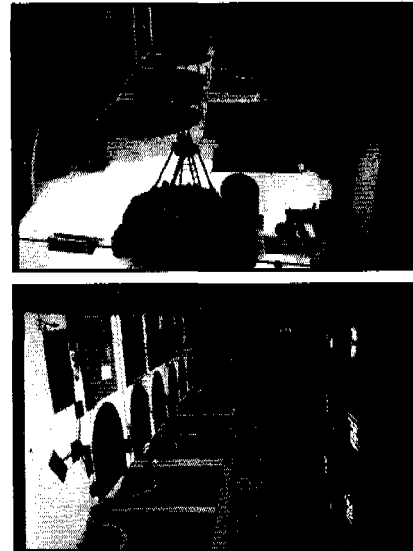


그림 10. 그림 4의 광원색 제거 결과



그림 7. 광원색 제거 실험 영상 (a) 붉은 광원의 영상 (b) 푸른 광원의 영상

제안한 광원에 독립적인 색 정보 추출 알고리즘의 순서도는 그림 11과 같다. 그림 11에서 $Diff_{RG}$, $Diff_{RB}$, $Diff_{GB}$ 는 각 채널간의 평균 밝기값의 차이의 절대값을 나타낸다. 본 연구에서 제안한 생동폭 조절 알고리즘과 광원색 제거 알고리즘은 각 채널 별로 독립적인 밝기값 조절에 의해 광원 조건의 변

화를 최소화하게 된다. 광원의 위치 변화가 심한 경우에는 생동폭 조절과 광원색 제거 알고리즘을 함께 적용할 경우에는 별 문제가 발생하지 않으나, 광원의 위치 변화가 작은 어느 정도 밝은 영상의 경우에는 두 가지 알고리즘을 함께 적용할 경우, 영상에 따라

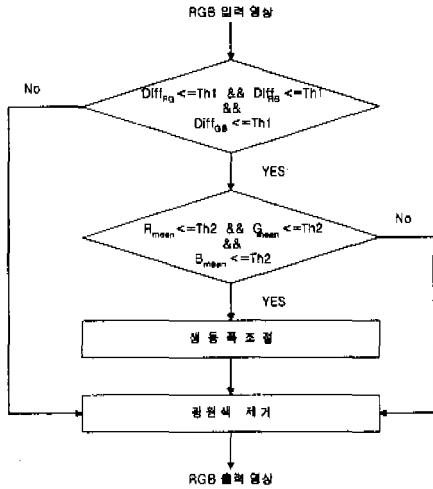


그림 11. 제안한 색 정보 추출 알고리즘의 흐름도

생동폭 조절 단계에서의 과도한 밝기값 조절로 인해 부분적으로 밝기값이 너무 밝게 되어 오히려 색 분별력이 떨어지는 경우가 생길 수 있다. 이러한 문제점을 제거하기 위해서는 생동폭 조절이 필요한 영상을 판단하는 것이 필요하게 되는데, 일반적으로 광원 위치의 변화가 심한 경우에는 영상 전체 영역에서 음영 영역이 차지하는 부분이 증가하게 되어 영상 전체의 평균 밝기값이 떨어지는 성질을 가지므로, 영상이 입력되면 각 채널의 평균값과 채널간의 평균값 차이의 절대값을 이용하여 생동폭 조절 여부를 결정하도록 하였다.

IV. 실험 및 고찰

제안한 색 정보 추출 알고리즘의 성능 평가를 위한 실험을 위해 본 논문에서는 자연, 인물, 동식물 등 다양한 종류의 5000 여 개의 영상으로 데이터 베이스를 구성하였다. 성능 평가를 위해 사용한 검색 정확도 측정 방법(retrieval accuracy measure)으로는 현재 MPEG-7의 표준화에서 사용하는 ANMRR^[17]을 이용하였다. 기존의 가장 널리 쓰이는 검색 정확도를 측정하는 방법인 precision 과

recall 에 비해 객관적인 비교가 가능해서 MPEG-7 표준화에 제안된 여러 방법들의 성능 비교에 쓰이고 있는 객관적인 성능 평가 방법이다. ANMRR를 구하기 위해서는 먼저 질의 영상을 결정하고, 각각의 질의 영상에 대해 관련성이 있는 영상들의 set인 ground-truth 영상을 결정하여야 하는데, 본 실험에서는 총 30개의 질의 영상을 결정하고, 각 질의 영상에 대한 ground-truth 영상들을 결정하였다. 최종적인 성능 평가값인 ANMRR를 구하기 위한 과정은 아래와 같으며, 평균 rank값을 나타내는 ANMRR은 그 값이 낮을 수록 검색 결과가 우수함을 나타낸다.

- NG(q) : 질의 영상 q에 대한 ground-truth 영상 수
- NR(q) : 검색된 영상 중에서 최상위 rank K 안에 검색된 ground-truth 영상 수

$$K = \min(4 \times NG(q), 2 \times GTM) \quad (14)$$

$$GTM = \max\{NG(q)\} \text{ for all } qs \text{ of a dataset}$$

- MR(q) : ground-truth 영상 가운데 검색되지 않은 영상 수

$$MR(q) = NG(q) - NR(q) \quad (15)$$

- ① AVR(Average Rank) :

$$AVR(q) = \frac{\sum_{k=1}^{NG(q)} Rank(k)}{NG(q)} \quad (16)$$

- ② MRR(Modified Retrieval rank) :

$$MRR(q) = AVR(q) - 0.5 - \frac{NG(q)}{2} \quad (17)$$

- ③ RR(Retrieval Rate) :

$$RR(q) = \frac{NR(q)}{NG(q)} \quad (18)$$

- ④ NMRR(Normalized Modified Retrieval rank) :

$$NMRR(q) = \frac{MRR(q)}{K + 0.5 - 0.5 \times NG(q)} \quad (19)$$

- ⑤ ANMRR(Average Normalized Modified Retrieval rank) :

$$ANMRR = \frac{1}{Q} \sum_{q=1}^Q NMRR(q) \quad (20)$$

여기서 Q 는 전체 질의 영상 수를 나타낸다. 실험에서 사용한 영상간의 유사도 측정(similarity measure) 방법은 제안한 칼라 정보 추출 알고리즘의 성능을 정확히 평가하기 위해 단순하게 아래와 같이 계산하였다.

$$\text{Similarity} = \sum_{i=0}^{\text{bin size}-1} |H_q(i) - H_r(i)| \quad (13)$$

여기서 H_q 와 H_r 는 질의 영상과 데이터베이스 영상을 각각 나타내며, 실험에서는 bin의 수를 252(6×6×7)개, 히스토그램 bin의 비트 수를 8로 하여 실험을 수행하였다.

표 1은 30개의 질의 영상에 대하여 기존의 입력 영상의 칼라 정보를 수정 없이 사용한 칼라 히스토그램의 결과와 제안한 칼라 정보 추출 알고리즘에 의해 추출된 칼라 정보를 이용한 칼라 히스토그램을 결과를 나타낸다. 제안한 방법으로 추출된 칼라 정보를 이용한 방법이 입력 영상의 칼라 정보를 그대로 이용하는 것 보다 더 나은 ANMRR값을 가짐을 확인할 수 있다.

표 1. 기존 방법과 제안한 방법의 ANMRR 비교

	ANMRR
기존 칼라 히스토그램	0.26
제안한 칼라 정보를 이용한 칼라 히스토그램	0.15



(a)



(b)



(c)

그림 12. 검색 결과 화면

(a) 질의 영상 (b) 기존 방법 (c) 제안한 방법

그림 12는 동일한 질의 영상에 대해 입력 영상의 칼라 정보를 이용한 경우의 칼라 히스토그램과 제안한 알고리즘으로 추출된 칼라 정보를 이용한 경우의 칼라 히스토그램 결과 가운데서 최상위 4개의 rank 값을 가지는 검색 결과를 나타낸다. 질의 영상에 대해 검색된 결과 영상들의 rank는 왼쪽에서 오른쪽으로 갈수록 증가하며, 질의 영상의 ground-truth 영상의 수는 4개이다. 제안한 방법으로 추출된 칼라 정보를 이용하여 칼라 히스토그램을 적용시킬 경우에는 동일한 내용을 담고 있는 영상 대부분을 검색하였으나, 입력 영상의 칼라 정보를 그대로 이용할 경우에는 관련 영상들을 제대로 검색하지 못함을 확인할 수 있다.

V. 결론

현재 표준화가 한창 진행중인 MPEG-7에서 고려되고 있는 여러 가지 특징들 가운데 칼라 특징은 표준화 이후에 영상 검색에 가장 대중적으로 사용되어질 수 있는 것으로 알려진다. 칼라 특징값이 다른 특징값들에 비해 정확한 영상 검색이 가능할 뿐만 아니라, 그 연산량이 다른 특징값을 이용할 때 보다 아주 적고, 영상의 부분적인 변화, 즉 회전, 이동, 폐색 등에 관계없이 영상의 내용을 잘 표현한다는 장점을 가지고 있다. 그러나 칼라 정보는 광원 조건이 변화하면 칼라 분포가 변하게 되어, 동일한 물체를 가진 영상이라 할 지라도 서로 다른 영상으로 검색되어지는 치명적인 결함을 가지고 있다.

그래서 본 논문에서는 광원의 위치 변화와 광원색의 변화에 영향을 받지 않는 칼라 정보 추출을 위한 연구를 수행하였다. 기존에는 광원색의 변화만을 고려하는데 비해, 본 연구에서는 광원의 위치 변화까지 고려하여 보다 정확한 물체색 복원이 가능하도록 하였다. 제안한 방법은 기존의 칼라 히스토그램 방법이 가지는 빠른 수행 속도의 장점을 해치지 않을 뿐만 아니라, 모든 연산이 RGB 채널의 밝기값을 가지고 수행되므로, 다른 과도한 연산량을 가지는 색 항상성 알고리즘들보다 내용 기반 영상 검색이라는 MPEG-7의 목적에 부합한다고 할 수 있다. 더 나은 검색 효율을 가지기 위해서는 영상마다 부분적으로 발생하는 밝기값의 부적절한 증가로 인한 색 분별력의 감소를 막을 수 있는 계조 재생(tone reproduction) 기법에 대한 개발이 필요하다.

