

논문-00-5-1-06

컬러와 모양 정보를 이용한 캐릭터 이미지 검색

이동호*, 유광석*, 김희율*

Charactor Image Retrieval Using Color and Shape Information

Dong-Ho Rhee*, Kwang-Seok Ryoo*, and Whoi-Yul Kim*

요 약

본 논문에서는 캐릭터 이미지 검색에 가장 적합한 컬러와 모양 정보를 동시에 표현하는 새로운 복합 특징량을 제안한다. 제안된 방법은 YCbCr 컬러 모델상에서 얻어진 Y 이미지에 대해서는 회전, 이동, 크기 변화에 불변한 Zernike 모멘트를 사용하여 모양 정보를 추출하고, Cb 및 Cr 이미지로부터 DCT 계수를 사용하여 색상 정보를 추출하여 캐릭터 이미지를 검색한다. 이 방법은 인간의 시각적인 특성을 잘 표현하는 방법으로서 36개의 적은 특징량으로 높은 검색효율을 나타내기 때문에 대용량 데이터베이스와 같은 웹 검색과 애니메이션 검색에 적합한 방법이다. 캐릭터 이미지 3,834개를 대상으로 실험하였으며 MPEG-7 컬러/질감 기술자들의 성능 평가에 사용된 ANMRR(Average of Normalized Modified Retrieval Rank)과 모션/모양 기술자들의 성능 평가에 사용된 BEP(Bull's Eye Performance)를 사용하여 캐릭터 이미지 검색에서 우수한 성능을 나타냄을 실험으로 확인하였다.

Abstract

In this paper, we propose a new composite feature consists of both color and shape information that are suitable for the task of character image retrieval. This approach extracts shape-based information using Zernike moments from Y image in YCbCr color space. Zernike moments can extract shape-based features that are invariant to rotation, translation, and scaling. We also extract color-based information from the DCT coefficients of Cr and Cb image.

This approach is good method reflecting human visual property and is suitable for web application such as large image database system and animation because higher retrieval rate has been achieved using only 36 features. In experiment, this method is applied to 3,834 character images. We confirmed that this approach brought about excellent effects by ANMRR(Average of Normalized Modified Retrieval Rank), which is used in the evaluation measure of MPEG-7 color descriptor and BEP(Bull's Eye Performance), which is used in evaluation measure of shape descriptor in character image retrieval.

I. 서론

최근 컴퓨터와 통신 기술의 급속한 발달로 인해 인터넷을 통한 멀티미디어 데이터의 전송이 보편화 되면서 정

보 통신망 사용자들은 시간과 장소에 상관없이 다양한 정보에 접근할 수 있게 되었으며 이용할 수 있는 멀티미디어 정보의 양도 폭발적으로 증가하게 되었다. 하지만 정보량이 늘어날수록 원하는 정보를 찾기는 더욱 어려워져 멀티미디어 정보의 검색, 저장, 관리 기술에 대한 요구가 증가하게 되었다.

이미지의 경우 대규모 데이터베이스를 이용하여 원거

* 한양대학교
Dept. of Electronic Engineering

리 감지를 위한 항공/위성 사진, 의료 이미지 정보 시스템, 법의학에서의 지문인식, 범죄인식, 지리정보 시스템, 백과 사전, 박물관, 캐릭터 이미지, 상표와 로고 등록에 서와 같이 다양한 분야에서 널리 개발 및 연구되고 있다[1]. 따라서 대용량 데이터베이스에서 원하는 정보를 빠른 시간 내에 검색할 수 있는 검색 기법이 필요하게 되었다.

특히 상용 캐릭터 (클럽아트 포함) 이미지의 경우, 현재 디자인 산업에서 널리 이용되고 있으며, 제조업, 서비스업 등 다른 모든 업종에서 요구되고 있다. 따라서 캐릭터 이미지나 클럽아트가 상표의 등록 출원으로 지적 재산 권리를 인정 받고 실제 디자인 산업에서 효용성을 갖기 위해서는 대용량 데이터베이스의 구축 및 관리되어야만 한다.

과거의 텍스트에 기반한 검색 방법의 문제점을 극복하기 위해 이미지 자체의 특징인 색상, 질감, 모양 및 객체들의 상대적인 위치 같은 것을 이용하여 이미지를 검색하는 방법이 연구되었다. 일반적으로, 사람들은 두개의 이미지를 비교할 때 "이 두 이미지는 모양이 같다", "질감이 같다 또는 색깔이 같다" 라고 평가를 하게 된다. 이처럼, 일반적인 내용기반 이미지 검색(Content-based Image Retrieval) 방법은 많은 응용에서 문자기반 이미지 검색(Text-based Image Retrieval) 방법의 단점을 보완해 주는 방법으로서 이미지 그 자체에 있는 색상(Color), 질감(Texture), 모양(Shape) 및 구성(Composition)과 같은 인지적이고 의미 있는 특징벡터(Feature Vector)를 이용하여 사용자가 원하는 이미지를 검색하는 방법이다. 자연 이미지의 경우 컬러나 질감 정보 등의 복합 특징벡터를 이용하여 검색 하는 것이 더 높은 검색률을 나타내며, 특히, 캐릭터 이미지의 경우 같은 모양일지라도 다른 이미지가 많이 존재하며 색상과 모양을 같이 고려했을 때 더 정확한 이미지를 검색할 수 있다.

캐릭터 이미지는 표현되는 주된 대상이 전체적인 모양과 색상이 주된 특징이기 때문에 본 논문에서는 캐릭터 이미지 검색에 가장 적합한 컬러와 모양 정보를 동시에 표현하는 새로운 복합 특징량을 제안한다. YCbCr 컬러 모델상에서 얻어진 Y이미지에 대해서는 정규 모멘트를 기반으로 이미지 정규화 알고리즘을 적용하여 이미지를 정규화 한 후, Zernike 모멘트를 사용하여 회전, 이동, 크기, 비틀림의 변화에 불변한 모양 특징량을 추출한다. 또 Cb와 Cr 이미지로부터 DCT계수를 사용 색상 정보를 추출한 다음 이러한 복합 특징량으로 캐릭터 이미지를 검

색한다.

본 논문은 다음과 같이 구성되어 있다. 제2장에서는 색상 정보와 모양 정보를 이용한 내용기반 이미지 검색에 관련된 기존의 연구를 정리하고, 제3장에서는 본 논문에서 제안한 캐릭터 이미지 검색에 사용된 특징벡터의 추출 방법을 설명한다. 제4장에서는 이를 이용하여 이미지들간의 유사도를 결정하는 방법을 설명하고, 제5장에서는 개발된 알고리즘을 실제 캐릭터 이미지에 적용한 실험 결과를 제시한다. 마지막으로 제6장에서는 결론을 맺는다.

II. 관련 연구

초기의 내용기반 검색에는 색상, 질감 및 모양과 같은 단일 특징벡터를 사용하여 이미지를 검색하였으나 최근에는 보다 더 정확한 검색을 위해 모폴로지, 세그먼트이션 및 그룹화 등을 이용하여 색상이나 질감과 함께 공간 정보를 결합하거나, 이미지 내에 있는 객체 등을 이용하여 검색하는 방법들이 제안되었다^{[4][5][6][7]}. 이미지에 포함된 객체의 모양 정보에 의한 검색 방법은 다른 방법들에 비해 매우 복잡하고 어려운 문제를 많이 내포하고 있으나 높은 수준의 검색 환경을 제공하기 위해 중요한 연구대상이며, 또 애니메이션이나 대용량 데이터베이스에서 빠른 검색을 위해 보다 적은 특징량으로 높은 검색률을 나타내는 방법들이 현재 연구 대상으로 인식되고 있다.

1. 색상 정보를 이용한 내용기반 이미지 검색

내용기반 이미지 검색을 위한 방법으로 초기에는 색상의 분포를 이용하는 방법이 많이 사용되었다^[2]. 이는 전체적인 영상의 색상 분포를 의미하는데, 계산이 쉽고 물체의 회전이나 작은 이동 등과 같은 기하학적인 변화에 강건한 장점이 있다. 그러나 인지적으로 비슷하지 않은 이미지도 동일한 색상 분포를 가질 수 있고, 특징량이 많아 검색 시 많은 시간이 소요된다는 문제점이 제기되었다. 이를 해결하기 위한 방법으로 색상 정보와 공간 정보를 함께 적용한 여러 방법들이 제안되었다.

Stricker^[3]는 이미지를 5개의 부분적으로 겹쳐진 영역으로 나누고, 각 영역의 색 분포를 3차 모멘트까지 사용하여 색 정보와 공간 정보를 결합하였다. Pass^[4]는 이미지 내에

서 같은 색상의 응집 여부를 이용한 CCV(Color Coherence Vectors)를 제안하여 기존의 색상 히스토그램 방법을 개선하였으나, 공간 정보가 극히 적다는 단점이 지적 되었다. M. K. Mandal^[5]은 색 정보를 반영하기 위해서 색상 히스토그램에 대한 모멘트를 이용하였고, 공간 정보 반영을 위해 웨이브릿 부이이미지를 가우시안 분포로 근사한 후 그 계수를 사용하였다. 이 방법은 특징량의 수가 적어 색인에는 용이하지만, 물체나 배경의 큰 변화 등에 강인하지 못하다는 단점을 나타내었다. 또한 특징량 추출 시 색 정보와 공간 정보를 분리해 다루었기 때문에 이미지간 유사도 정의 시 실험에 의해 각 특징량의 가중치를 결정해야 하는 단점이 있었다.

J. Huang^[6]은 Color Correlogram이라는 새로운 방법을 제안하였다. 이 방법은 이미지 내의 화소로부터 특정 거리만큼 떨어진 점들에 대해 해당 화소의 빈도 수를 표시한 히스토그램으로 이미지를 비교하는 것으로, 색상 정보와 공간 정보를 효과적으로 결합하였으나, 다른 컬러와의 상관관계를 고려하지 않은 단점이 있다.

2. 모양 정보를 이용한 내용기반 검색

모양 정보를 이용한 검색 방법으로는 다음과 같은 방법들이 제안된 바 있다. H. V. Jagadish는 패턴을 커버하는 최소의 사각형의 조합으로 패턴을 계층 구조적으로 표현하고 이를 이용하여 패턴의 유사도를 판별하는 방법을 제안하였다^[7]. T. Kato는 주어진 패턴에서 몇 가지 특징들을 추출하여 패턴의 유사도를 판별하는 방법을 제안하였다^[8]. G. Cortelazzo 등은 스트링 매칭 기법을 이용해 상표 이미지의 모양을 기술하는 방법을 제안하였다^[9]. C. P. Lam 등은 STAR(System for Trademark Archive and Retrieval)라는 상표 검색 시스템을 제안했다^[10]. 이 시스템에서는 사용자가 지정한 상표의 중요 특징으로부터 추출한 불변 모멘트와 푸리에 기술자를 특징벡터로 사용하고, 상표들 사이의 유사도는 퍼지 사전에 의해 측정하였다.

J. Bign 등은 경계선 방향의 히스토그램과 비슷한 방위 라디오그램을 이용한 이미지 검색 시스템을 제안하였다^[11]. A. D. Bimbo 등은 스케일 스페이스 분석으로 유도된 곡선의 계층 모델을 이용한 이미지 검색 시스템을 제안하였다^[12]. F. Mokhtarian 등은 패턴의 곡률을 스케일 스페이스에서 표현하고, 이의 제로-크로싱(Zero-Crossing)의 극대값을 이용하여 패턴의 유사도를 판별하는 방법을 제안하였다^[13]. J. P. Eakins 등은 SAFARI (Shape Analysis

For Automatic Retrieval of Images)^[14]와 ARTISAN (Automatic Retrieval of Trademark Images by Shape ANalysis)^[15]라는 두 시스템에서 패턴의 경계선 정보를 이용하여 유사한 상표를 검색하는 방법을 제안하였다. A. K. Jain 등은 계층적 이미지 검색 시스템을 제안하고 상표 이미지 데이터베이스에 이를 적용하였다^{[16][17]}. 이 시스템은 경계선 방향의 히스토그램과 불변 모멘트를 특징벡터로 사용하는 빠른 1 단계 검색과 가변 템플릿 매칭을 사용하는 세부 매칭 단계라는 2단계 계층구조를 사용하였다.

기존의 방법들은 다음과 같은 문제점을 가지고 있다. 경계선은 패턴의 모양을 결정하는 중요한 요인이기는 하지만 패턴의 미세한 변화에도 쉽게 영향을 받을 수 있다는 단점을 가지고 있다. 또한 푸리에 기술자나 패턴의 곡률에 기초한 방법은 이미지 내에 오직 하나의 물체만 존재해야 적용할 수 있다. 이러한 문제의 해결을 위해 모폴로지나 근거한 전처리 과정의 도입도 생각해 볼 수 있으나, 이러한 전처리 과정은 그 결과가, 수축이나 팽창과 같은 연산의 수행 회수에 매우 민감하게 영향을 받기 때문에 데이터베이스 상의 수많은 이미지에 대해 일반적으로 적용하기는 어렵다. J. Bign 등과 A. K. Jain 등이 제안한 경계선의 히스토그램을 이용하는 방법은 경계선의 방향만 고려할 뿐, 경계선의 위치 정보를 전혀 고려하지 않기 때문에 패턴의 구별 능력이 현저히 떨어지게 된다^{[11][16][17]}. 즉, 사람의 눈으로 보기에는 전혀 다른 모양의 패턴들도 이 방법으로는 비슷한 패턴으로 분류되는 단점이 있다.

III. 캐릭터 이미지에 대한 특징량 추출 방법

컬러 이미지에 대한 특징량을 추출하는 과정에서 우선적으로 검토되어야 할 사항은 어떤 컬러 모델을 적용할 것인가에 대한 검토가 필요하다. 컬러 모델에는 RGB, HSV, YUV, YIQ 및 HMMD^[18] 등 다양한 모델이 사용되고 있다. 본 논문에서는 색상 정보와 모양 정보를 효율적으로 추출할 수 있고 애니메이션 검색하기 위해 JPEG이나 MPEG 과 같은 압축 표준에서 사용하는 YCbCr 컬러 모델을 사용하였다. 이 모델은 휘도 성분인 Y 이미지에 대해 모양 정보를 추출하고, 색차 성분인 Cb, Cr 이미지에 대해서는 색상 정보를 추출할 수 있게 한다. 그림 1은 본 논문에서 제안한 특징량 추출 방법에 대한 전체적인 알고리즘을 나타낸 것이다.

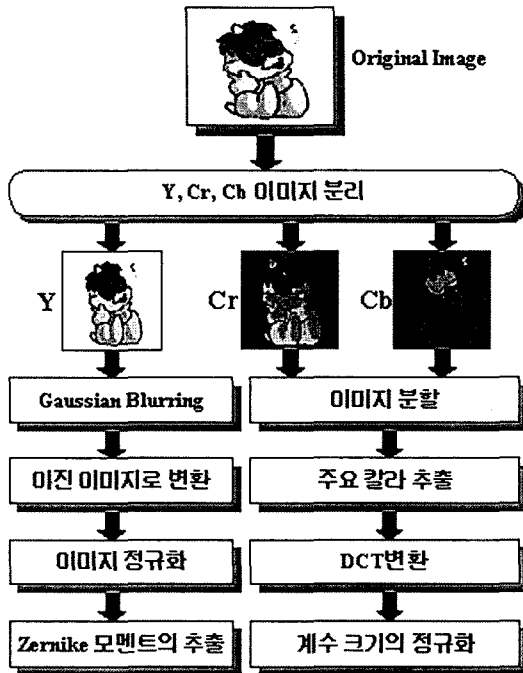


그림 1. 제안된 방법의 특징벡터 추출 과정
Fig. 1. Extraction process of feature vector by proposed algorithm

1. 모양 특징벡터의 추출

본 논문에서 제안하는 모양 정보의 추출 방법은 그림 1에서 보는 바와 같이 YCbCr 컬러 모델에서 얻어진 Y 이미지에 대해서 모양 정보를 추출 한다. 그림 6에서 보는 바와 같이 캐릭터 이미지 특성상 정적인 이미지보다 활동적인 캐릭터 이미지가 많기 때문에 전처리 과정으로 Gaussian blurring을 이용하여 캐릭터 이미지의 미세한 모양의 변화에 대한 윤곽을 부드럽게 하고, 패턴내의 미세한 구멍 같은 것을 채우게 한 다음, 캐릭터 이미지의 모양에 대한 영역을 얻기 위해 이진 이미지로 변환한다. 그 변환된 이미지가 이동, 크기 및 비틀림에 변화에 불변하기 위해 이미지를 정규화 한다.

정규화 알고리즘은 Zernike모멘트의 가장 취약점인 이미지의 비틀림을 보완해 주는 것으로 정규 모멘트 (Regular moment)를 토대로 이미지 정규화 하는 방법을 이용하였다^{[19][20]}. 이미지 정규화 알고리즘의 특징을 아래의 그림 2 ~ 4에서 살펴보면 크기, 이동 및 비틀림의 변화에 불변한 이미지를 얻을 수 있음을 알 수 있다. 각 그림에서 (a)~(c)는 크기, 이동 및 비틀림 등에 의해 변형된 이미지이고 (d)~(f)는 이미지 정규화 과정을 통

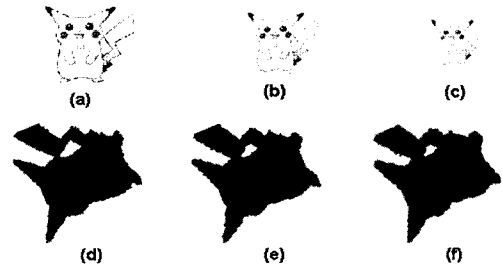


그림 2. 크기 불변 특징 (a)~(c) 원래 이미지, (d)~(f) 정규화된 이미지
Fig. 2. Property of scaling invariant (a)~(c) Original images (d)~(f) Normalized images

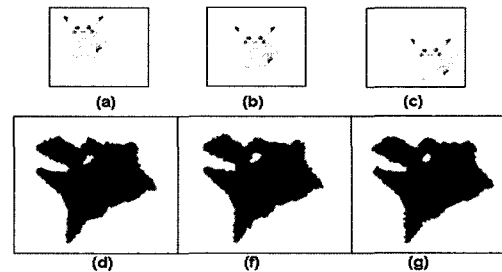


그림 3. 이동 불변 특징 (a)~(c) 원래 이미지 (d)~(f) 정규화된 이미지
Fig. 3. Property of translation invariant (a)~(c) Original images (d)~(f) Normalized images

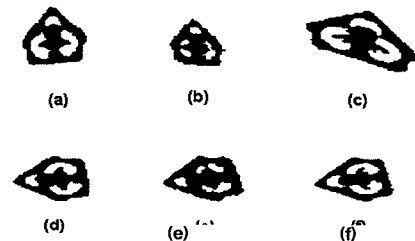


그림 4. 비틀림 불변 특징 (a)~(c) 원래 이미지 (d)~(f) 정규화된 이미지
Fig. 4. Property of skewness invariant (a)~(c) Original images (d)~(f) Normalized images

해서 얻어진 이미지들이다.

마지막으로, 전처리 과정에서 얻어진 정규화된 이미지에 대해서 Zernike 모멘트를 이용하여 특징벡터를 추출 한다. Zernike 모멘트는 패턴의 영역을 토대로 특징량을 추출하기 때문에 경계선 패턴의 모양으로 검색하는 방법들의 많은 단점을 극복할 수 있고, 회전 불변 및 잡음에 강인하기 때문에 패턴 인식 시스템에 널리 사용되어 왔다^[21]. 또 이미지 정규화 알고리즘에서의 약점인 회전에 불변하지 않은 특징을 보완해 준다. 본 논문에서는 정규

회전 이미지로부터 모양 정보를 추출하기 위한 특징벡터로는 DC 값인 0차 모멘트 값을 제외한 8차까지의 Zernike 모멘트의 절대값 24개를 사용하였다. 0차 모멘트는 객체의 면적으로서 크기에 불변한 특징량을 얻기 위해 사용된다. Zernike 모멘트에 대해 간단히 설명하면 다음과 같다.

2차원 이미지 $f(x, y)$ 의 Zernike 모멘트는 다음과 같이 정의되는 2차 복소 모멘트이다.

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(x, y) \quad (1)$$

여기에서 $n=0,1,2,\dots,\infty$ 의 범위를 가지고, m 은 다음 조건을 만족해야 한다.

$$n-|m|=even, \quad |m| \leq n. \quad (2)$$

Zernike 모멘트의 기저함수 $V_{nm}(x, y) = R_{nm}(\rho)e^{jm\theta}$ 는 $x^2 + y^2 \leq 1$ 인 단위원 내에서 정의되는 복소 함수이고, $R_{nm}(\rho)$ 는 $|\rho| \leq 1$ 에서 정의되는 직교 방사 다항식이다.

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

극좌표계에서 표현된 입력 이미지 $f(\rho, \theta)$ 이 α 만큼 회전된 패턴을 $f'(\rho, \theta)$ 라고 하자.

$$f'(\rho, \theta) = f(\rho, \alpha + \theta) \quad (4)$$

회전된 패턴의 Zernike 모멘트는 다음과 같이 주어진다.

$$A_{nm}^r = \frac{n+1}{\pi} \sum_x \sum_y f(\rho, \alpha + \theta) V_{nm}^*(\rho, \theta), \quad \text{s.t. } \rho \leq 1 \quad (5)$$

즉,

$$A_{nm}^r = A_{nm} \exp(-j m \alpha) \quad (6)$$

가 됨을 알 수 있다. 여기서 회전된 이미지의 Zernike 모멘트의 절대값은 기준 이미지의 Zernike 모멘트의 절대값과 동일한 식 (7)과 같다.

$$\|A_{nm}^r\| = \|A_{nm}\| \quad (7)$$

2. 컬러 특징벡터의 추출

색상 정보는 YCbCr 컬러 모델에서 얻어진 Cb, Cr 이미지에 대해서 특징벡터를 추출한다^[22]. Cb, Cr 이미지에 대해 특징벡터를 추출하는 과정은 이미지 분할, 주요 컬러 추출, DCT 변환, 특징량의 정규화 등 4단계의 과정으로 특징벡터를 추출한다. 처음, 이미지 분할은 데이터베이스 이미지들의 크기에 상관 없이 16×16으로 나눈다. 따라서 그 블록들의 가로와 세로의 크기는 각각 X_Width/16이고 Y_Height/16이 된다. 여기서 X_Width, Y_Height는 원 이미지의 가로 넓이와 세로 높이이다. 만일 원 이미지가 16의 배수가 아닐 경우 마지막 행과 열에 해당하는 블록은 원래 이미지에서 남은 부분에 대한 크기가 된다.

나누어진 블록에서 주요 색상 성분의 추출 방법은 블록 내의 모든 픽셀 값들의 평균 값으로 선택한다. Cb, Cr 이미지의 히스토그램 분포를 살펴보면 일정한 값들의 범위 내로 몰려있는 것을 볼 수 있는데 이러한 특징이 DCT 변환을 사용하여 몇 개의 계수만으로 효율적인 색

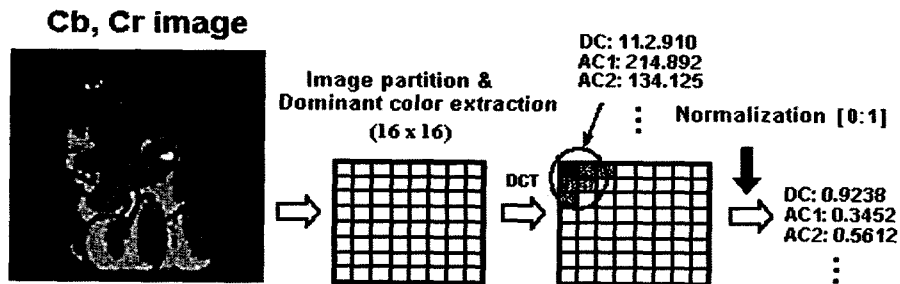


그림 5. 색상 이미지의 특징량 추출 방법

Fig. 5. Feature extraction method of chrominance components

상 정보를 추출하게 한다. JPEG압축 알고리즘에서 사용하는 지그재그 스캔 방식을 사용하여 2D-DCT 변환된 색상 이미지에 대해 6개의 특징벡터를 추출한다. 추출된 계수는 1개의 DC성분과 5개의 AC성분을 나타낸다. 5개의 AC성분은 색상 이미지에 대한 질감 정보를 나타낼 수 있는 정보량이기도 하다. 마지막으로 특징량 값들을 0과 1사이의 값을 갖기 위해 특징량들을 정규화 하는데 본 논문에서는 선형 정규화 방법으로 정규화 하였다. 그림 5는 앞에서 설명한 색상 이미지의 특징량 추출 과정을 그림으로 나타낸 것이다.

IV. 유사도 측정

캐릭터 이미지의 검색을 위해 질의 이미지와 데이터베이스내의 이미지들에 대해 각각 특징벡터를 추출하여 그 특징벡터 값을 이용하여 검색한다.

본 논문에서는 질의 이미지와 데이터베이스 내의 이미지와의 거리 D는 식(8)에 의해 구해지고 유사도S는 식(9)을 사용하여 유사도를 측정한다.

$$D = \omega_1 \sqrt{\sum_{eC1} (Q(1)_i - T(1)_i)^2} + \omega_2 \sqrt{\sum_{eC2} (Q(2)_i - T(2)_i)^2} + \omega_3 \sqrt{\sum_{eC3} (Q(3)_i - T(3)_i)^2} \quad (8)$$

$$S = \frac{1}{1+D} \quad (9)$$

여기서 $Q(1)_i, T(1)_i$ 는 각각 질의 이미지와 데이터베이스 이미지에 대한 Y 이미지의 Zernike모멘트 i번째 특징량을 나타내고, $Q(2)_i, T(2)_i$ 는 각각 질의 이미지와 데이터베이스 이미지에 대한 Cb이미지의 DCT 계수 i 번째를 나타낸다. 마지막으로 $Q(3)_i, T(3)_i$ 는 각각 질의 이미지와 데이터베이스 이미지에 대한 Cr이미지의 DCT 계수 i번째를 나타낸다. $\omega_1, \omega_2, \omega_3$ 는 각각의 이미지에 대한 가중치이고 그 가중치를 어떤 쪽에 많이 할당 하느냐에 따라 검색 되는 이미지에도 많은 차이가 있다. 유사도 S는 정규화 과정을 거쳐 0과 1사이의 값을 얻을 수 있다.

V. 실험 및 결과

본 논문에서 제안한 방법을 검증하기 위해 캐릭터 이미지 866개를 80개 그룹으로 나눈 경우와, 그 이미지를 그룹화 하지않은 2,968개의 이미지에 포함시켜 전체 3,834개의 실험 이미지로 구성한 경우 이 두 경우에 대해 실험을 하였다. 실험 환경은 펜티엄 III 450Mhz기종과 윈도우 98의 운영체제 상에서 진행하였으며, Visual C++ 6.0 컴파일러를 사용하였다.

그림 6은 제안된 알고리즘의 실험 전개를 위해 사용한 캐릭터 이미지들을 보여준 것이다. Y이미지를 정규화 할 때 스�কে링 팩터를 20으로 하여 120×120크기로 이미지를 정규화 하였으며 이진 영상으로 변환하는 과정에서 배경



그림 6. 실험에 사용된 캐릭터 이미지

Fig. 6. Character images used in the experiment

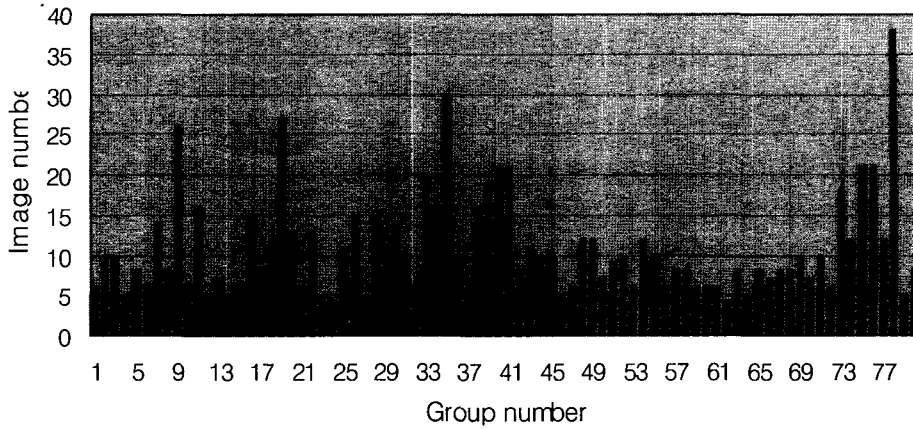


그림 7. 80개 그룹들에 대한 이미지 개수

Fig. 7. The number of images in each 80 ground truth

을 제거하기 위해 임계치를 230로 두고 실험하였다. 또 유사도 측정시 가중치를 $\omega_1=1, \omega_2=1, \omega_3=1$ 로 하였을 때 가장 좋은 결과를 얻을 수 있었다. 그림 7은 80 그룹에 속한 각각의 이미지 개수를 그래프로 나타낸 것이다.

성능 분석 지표로는 현재 MPEG-7 컬러/질감에 대한 성능평가 방법인 ANMRR(Average of Normalized Modified Retrieval Rank)^[23]과 모션/모양에 대한 성능평가 방법인 BEP(Bull's Eye Performance)^[24]로서 성능평가를 하였다. 다음 식은 ANMRR을 구하는 식이다.

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

$$\text{where } K = \min(4 * NG(q), 2 * GTM)$$

$$AVR(q) = \sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)} \quad (10)$$

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

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

여기서 $NG(q)$ 는 질의하는 이미지가 속해 있는 그룹내의 전체 이미지 개수이고, GTM 은 $NG(q)$ 중 가장 큰 값이다. K 가 결정되면 $Rank$ 는 어떤 데이터셋에 있는 이미지를 질의했을 때 검색순위가 K 값 이하일 경우, 검색의 순위와 같고 K 값보다 클 경우는 검색 순위에 상관없이 항상 $K+1$ 의 값을 가진다. $Rank$ 가 구해지면 $AVR(q)$,

표 1. Zernike 모멘트 12까지의 검색결과

Table 1. Retrieval performance of Zernike moments up to order 12th

차수	6	7	8	9	10	11	12
특징량 개수	16	20	25	30	36	42	49
BEP(%)	71.12	71.73	73.31	73.68	73.83	74.09	73.51
ANMRR	0.2704	0.2635	0.2483	0.2479	0.2419	0.2426	0.2450

$MRR(q)$ 등이 구해지고 $NMRR$ (Normalized Modified Retrieval Rank)은 항상 0에서 1의 값을 가지며, 낮은 값일수록 좋은 검색 결과를 나타낸다. 최종적으로 ANMRR은 $NMRR(q)$ 들에 대한 평균을 나타낸 값이 된다.

80개 그룹의 866개 데이터셋 이미지들에 대해서 다음과 같은 결과를 얻었다. 표 1은 Zernike 모멘트의 차수에 따른 특징량의 수와 검색률을 나타낸 것이다. 표 1에서 보는 바와 같이 높은 차수일 때 검색률이 약간 떨어지는 것을 볼 수 있는데 이것은 차수가 높아짐에 따라 모양의 미세한 변화를 표현하기 때문에 비슷한 모양을 검색하는 데에는 원하지 않은 것이 된다. 따라서 8차 이상은 검색률에 변화가 거의 없으면서 특징량이 작기 때문에 본 논문에서는 DC 값인 0차 항을 제외한 8차 24개를 특징량으로 사용하였다.

표 2는 DCT계수의 개수에 따른 검색률을 나타낸다. 표 2에서 보는 바와 같이 계수가 6개일 때 가장 좋은 검색률을 나타내고 계수의 수가 10이상일 때 검색률이 약간

표 2. DCT 계수에 따른 검색결과

Table 2. Retrieval performance according to different DCT coefficient

DCT 계수의 개수	3	4	5	6	10	15	21
BEP(%)	76.93	78.41	79.42	82.55	80.26	81.80	80.92
ANMRR	0.2209	0.2011	0.1947	0.1664	0.1847	0.1722	0.1793

표 3. 각 특징벡터 들에 대한 검색 결과

Table 3. Retrieval performance with regard to each feature vector

	컬러	모양	컬러와 모양
BEP(%)	82.55	73.66	89.19
ANMRR	0.1664	0.2483	0.098

감소하는 것을 볼 수 있는데 이것은 AC성분이 증가하므로 해서 컬러 이미지 내에 있는 질감 성분을 많이 포함하는 것으로 캐릭터 이미지를 검색하는데 원하지 않는 성분임을 나타낸다. 본 논문에서는 색상 성분 이미지에 대해 각각 6개씩 12개의 특징량을 사용하였다.

표 3은 컬러, 모양, 컬러와 모양 특징량을 통합했을 때의 검색률을 나타낸다. 컬러의 경우 DCT 계수만을 사용했을 때 나온 결과이고, 모양은 Zernike 모멘트만을 사용했을 때 검색률을 나타낸다. 컬러와 모양 정보를 통합한 것을 토대로 검색했을 때 더 좋은 검색률을 나타낼 수 있다. 이것은 사람이 이미지를 판단할 때와 같이 하나의 특징(컬러, 모양, 질감)만으로 판단하는 것이 아니라 그들의 조합으로 인식을 한다는 특징을 잘

표 4. 866개의 캐릭터 이미지에 대한 검색 결과

Table 4. Retrieval performance for the 866 character images

	컬러 히스토그램	제안된 방법
BEP(%)	86.75	89.19
ANMRR	0.1255	0.098

표 5. 3,834개의 캐릭터 이미지에 대한 검색 결과

Table 5. Retrieval performance for the 3,834 character images

	컬러 히스토그램	제안된 방법
BEP(%)	81.65	83.26
ANMRR	0.1853	0.1672

나타내준다.

그림 8과 표 4는 80개 그룹 866개 이미지에 대하여 NMRR과 실험 결과를 나타낸 것이다. 컬러 히스토그램의 경우 제안된 방법과 특성을 비교하기 위해 HSV 컬러 모델에서 36번을 사용하여 히스토그램을 구한 결과이다.

그림 9에서 (a)는 제안한 방법인 컬러와 모양 특징량을 동시에 사용했을 경우에 대한 검색 결과이고 (b)는 컬러 히스토그램으로 검색했을 경우의 검색 결과를 보여준다. 좌상단 첫번째 그림을 질의 했을 때의 검색 결과를 보여주는 그림으로서 제안된 방법인 (a)그림이 더 좋은 검색 결과를 보여준다. 컬러와 모양 특징을 동시에 고려하는 것이 인간의 시각적인 특성을 더 잘 반영해주는 것

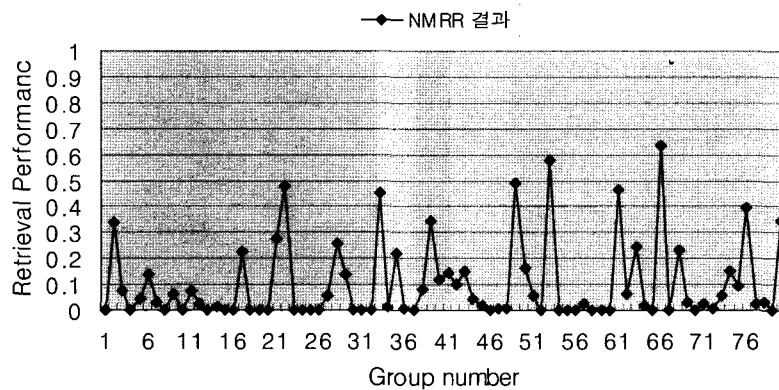


그림 8. 80개 그룹에 대한 NMRR결과

Fig. 8. NMRR results for each 80 ground truth

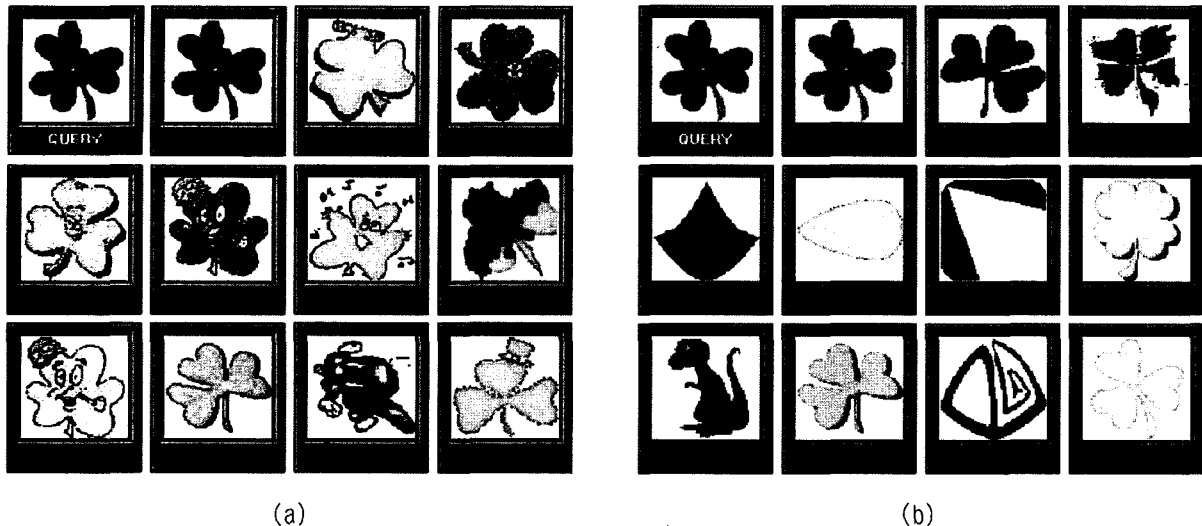


그림 9. 검색 결과 (a) 제안된 방법으로 검색 결과 (b) 컬러 히스토그램으로 검색 결과

Fig. 9. An example of retrieval result (a) Retrieval result by proposed method (b) Retrieval result by color histogram

을 알 수 있다. 표 5는 3,834개의 캐릭터 이미지에 대해 컬러 히스토그램과 비교 실험한 결과이다.

필요한 것으로 사료된다. 또한 대용량 데이터베이스에서 보다 빠른 검색을 위해 캐릭터 이미지들의 자동적인 클러스팅 기술이 요구된다.

VI. 결론

캐릭터 이미지의 경우 색상과 모양 정보가 특징벡터 추출의 주된 정보이기 때문에 본 논문에서는 모양 정보를 나타낼 수 있는 Zernike 모멘트와 색상 정보를 표현할 수 있는 DCT를 이용하여 모양과 컬러 정보가 결합된 새로운 내용기반 이미지 검색 알고리즘을 개발하고, 이를 3,834개의 캐릭터 이미지에 적용하여 그 성능을 평가하였다.

실험 결과를 통해, Zernike 모멘트와 DCT계수는 각각 모양 정보와 색상 정보를 기술하는 효과적인 특징벡터이며, 검색 이미지들의 크기에 불변한 특징벡터를 추출할 수 있는 큰 장점도 가지고 있다. 또 이 방법은 인간의 시각적인 특성을 잘 표현하는 방법으로서 36개의 매우 적은 특징량으로 검색하므로 대용량 데이터베이스와 같은 웹 검색 시스템과 애니메이션 검색에 적합한 방법을 제시하였다.

향후 연구과제로는 모양 정보와 컬러 정보를 결합하는데 사용되는 가중치에 대한 보다 정량적인 분석이 필요하고, 사용자로부터 검색 결과로 제시된 이미지들의 관련성을 피드백 받아 가중치를 재조정하는 방법에 대한 연구가

참고 문헌

- [1] V. N. Gudivada and V. V. Raghavan, "Content-based Image Retrieval Systems," *IEEE Computer*, pp. 18-22, Sep. 1995.
- [2] M. J. Swain and D. H. Ballard, "Color Indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11-32, 1991.
- [3] M. Stricker and A. Dimai, "Color Indexing with Weak Spatial Constraints," *Storage and Retrieval for Image and Video Databases IV*, SPIE Proceedings vol. 2670, 1996.
- [4] G. Pass and R. Zabih, "Histogram Refinement for Content-based Image Retrieval," *IEEE Workshop on Applications of Computer Vision*, pp. 96-102, 1996.
- [5] M. K. Mandal, T. Aboulnasr, and S. Panchanathan, "Image Indexing Using Moments and Wavelets," *IEEE Transactions on Consumer Electronics*, vol. 42,

- no. 3, pp. 557-565, Aug. 1996.
- [6] Jing Huang, S. Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih, "Image Indexing Using Color Correlogram", *International Conference on Computer Vision and Pattern Recognition, IEEE*, 1997.
- [7] H. V. Jagadish, "A Retrieval Technique for Similar Shapes", *Proc. of ACM SIGMOD*, pp.208-217, Jun, 1991.
- [8] T. Kato, "Database Architecture for Content-based Image Retrieval", *Proc. SPIE Conf. on Image Storage and Retrieval Systems*, vol. 1662, pp. 112-123, 1992.
- [9] G. Cortelazzo, G. A. Mian, G. Vezzi, and P. Zamperoni, "Trademark Shapes Description by String-matching Techniques", *Pattern Recognition*, vol. 27, no. 8, pp. 1005-1018, 1994.
- [10] C. P. Lam, J. K. Wu, and B. Mehtre, "STAR - a System for Trademark Archival and Retrieval," *Proceedings 2nd Asian Conf. on Computer Vision*, vol. 3, pp. 214-217, 1995.
- [11] J. Bign, S. K. Bhattacharjee, and S. Michel, "Orientation Radiograms for Image Retrieval: An Alternative to Segmentation," *Proc. of IEEE Int. Conf. on Pattern Recognition*, vol. 3, pp. 346-350, 1996.
- [12] A. D. Bimbo and P. Pala, "Image indexing using shape-based visual features," *Proc. of IEEE Int. Conf. on Pattern Recognition*, vol. 3, pp. 351-355, 1996.
- [13] F. Mokhtarian, S. Abbasi, and J. Kittler, "Efficient and Robust Retrieval by Shape Content Through Curvature Scale Space," *Proc. of Int. Workshop on Image Databases and Multimedia Search*, 1996.
- [14] J. P. Eakins, "Retrieval of Trademark Images by Shape Feature", *Proc. of ELVIRA Conf.*, May, 1994.
- [15] J. P. Eakins, K. Shields, and J. Boardman, "ARTISAN - a Shape Retrieval System Based on Boundary Family Indexing," *Proc. SPIE, Storage and Retrieval for Image and Video Database IV*, vol. 2670, pp. 17-28, Feb. 1996.
- [16] A. K. Jain and A. Vailaya, "Image Retrieval Using Color and Shape," *Pattern Recognition*, vol. 29, pp 1233-1244, Aug. 1996.
- [17] A. Vailaya, Yu Zhong and A. K. Jain, "A Hierarchical System for Efficient Image Retrieval," *Proc. of IEEE Int. Conf. on Pattern Recognition*, vol. 3, pp. 356-360, 1996.
- [18] Heon. Jun Kim, Sung Bae Jun, Jin Soo Lee, Jung Min Song and Hee Youn Lee, LG Corporate Institute of Technology, "Descriptor for Quantized Color Using HMMD Color Model and Color Histogram," *Proposal Doc. P669, MPEG-7 Ad-hoc Group Meeting, Lancaster*, Feb. 1999.
- [19] S. C. Pei and C. N. Lin, "Image Normalization for Pattern Recognition," *Image and Vision Computing*, vol. 13, no. 10, Dec. 1995.
- [20] D. Shen and H. H. S. Ip, "Generalized Affine Invariant Image Normalization," *IEEE Trans. Pattern Anal. And Machine Intell.*, vol. 19, no. 5, pp. 431-440, May. 1997.
- [21] Yong-Sung Kim and Whoi-Yul Kim, "Content-Based Trademark Retrieval System Using Visually Salient Feature," *Journal of Image and Vision Computing*, vol. 16/12-13, Aug. 1998.
- [22] A. Yamada et. al., "Results on Core Experiment of Compact Color Layout D," *ISO/IEC JTC1/SC29/WG11 MPEG99/M5356*, Oct. 1999.
- [23] Video Group, "Description of Core Experiments on MPEG-7 Color/Texture Descriptors," *ISO/IEC JTC1/SC29/WG11 MPEG99/ N2929*, Oct, 1999.
- [24] Video Group, "Description of Core Experiments for MPEG-7 Motion/Shape," *ISO/IEC JTC1/SC29/WG11 MPEG99/ N2690*, Mar. 1999.

 저 자 소개

**이 동 호**

1972년 8월 10일생

1998년 2월 : 인제대학교 전자공학과(공학사)

1998년 8월 ~ 현재 : 한양대학교 전자공학과 석사 과정

주관심분야 : 패턴인식, 영상처리, 비디오 인덱싱

**유 광 석**

1963년 10월 25일생

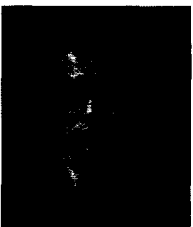
1986년 2월 : 한양대학교 전자공학과 졸업(공학사)

1990년 2월 : 한양대학교 전자공학과 졸업(공학석사)

1998년 2월 : 한양대학교 전자공학과 박사과정 수료

1990년 5월~1995년 4월 : 한국전기통신공사 정보통신사업본부 전임연구원

주관심분야 : 패턴인식, 영상처리, 비디오 인덱싱

**김 회 울**

1956년 5월 27일생

1980년 : 한양대학교 전자공학과(공학사)

1983년 : Pennsylvania State Univ. 전자공학(공학석사)

1989년 : Purdue University 전자공학(공학박사)

1985년 9월 ~ 1989년 5월 : Purdue University 연구조교

1989년 9월 ~ 1994년 2월 : University of Texas 조교수

1994년 3월 ~ 현재 : 한양대학교 전자공학과 부교수

주관심분야 : 영상처리, 컴퓨터 비전, 패턴인식, 워터마크