

지역 칼라와 질감을 활용한 블록 기반 영상 검색 기술자 설계

박성현 · 이용환* · 김영섭*

†단국대학교 응용컴퓨터공학과

*단국대학교 전자공학과

Design of Block-based Image Descriptor using Local Color and Texture

Sung-Hyun Park, Yong-Hwan Lee[†] and Youngseop Kim*

[†]Department of Applied Computer Engineering, Dankook University

*Department of Electronic Engineering, Dankook University

Abstract

Image retrieval is one of the most exciting and fastest growing research fields in the area of multimedia technology. As the amount of digital contents continues to grow users are experiencing increasing difficulty in finding specific images in their image libraries. This paper proposes an efficient image descriptor which uses a local color and texture in the non-overlapped block images. To evaluate the performance of the proposed method, we assessed the retrieval efficiency in terms of ANMRR with common image dataset. The experimental trials revealed that the proposed algorithm exhibited a significant improvement in ANMRR, compared to Dominant Color Descriptor and Edge Histogram Descriptor.

Key Words : Content-based Image Retrieval, Block-base Image Descriptor, Image Search, Local Color Texture, Image Feature Extraction

1. 서 론

최근 스마트폰, 디지털 카메라, 캠코더 등 멀티미디어 장비의 발전에 따라, 저렴한 가격으로 멀티미디어 디바이스가 보급되면서 대용량, 다량의 데이터들이 축적되고 있다. 멀티미디어 기기를 통해 수집되는 영상, 음성, 동영상 등의 멀티미디어 콘텐츠들 중 가장 많은 사용자층을 가지는 영상 콘텐츠는 시간이 지남에 따라, 기하급수적으로 데이터의 양이 증가하고 있다. 이에 따라 영상 검색은 최근 멀티미디어 관련 기술에서 부각되고 있는 연구 분야중의 하나가 되었다[1].

영상 검색(Image Retrieval)은 디지털 영상 데이터베이스에서 사용자가 원하는 적절한 영상을 열람(Browse)하고 검색(Search)하는 시스템을 의미하며[14], 검색 유형에 따라 2가지 형태로 나뉜다[2]. 첫 번째 형태는 키워드(Keyword) 중심으로 검색을 지원하는 텍스트

기반 영상 검색(Text-based Image Retrieval)이다. 여기에는 기존의 웹 기반 영상 검색 기능을 제공하는 구글, 야후와 네이버 검색 등이 포함된다. 이러한 텍스트 기반의 웹 영상 검색은 키워드와 주석을 기반으로 검색을 수행하며, 키워드 기록자의 주관적인 해석이 반영되어 키워드가 생성되기 때문에, 검색 만족도가 떨어지는 문제점이 있다. 또한 대용량, 다량의 영상 데이터베이스에서 키워드 생성에 따른 작업 시간과 작업 부하가 높다는 단점이 있다. 이러한 단점을 보완하기 위한 방법으로, 내용-기반 영상 검색(CBIR, Content-based Image Retrieval)이 연구되고 있다. 이러한 CBIR은 검색을 위한 영상 특성을 자동으로 추출하고, 이를 저장하고 있다가 사용자 질의 영상에 따라 유사도(Similarity)를 평가하여, 적절한 영상을 반환하는 컴퓨터 비전의 응용분야이다[14]. 영상의 특성을 추출하기 위해, 주로 칼라, 질감, 모양 등과 같은 저급의 시각적(Low-level Visual) 특징을 활용한다[3]. CBIR 시스템의 핵심 알고리즘인 영상 검색 기술자(Image Descriptor)는 칼

[†]E-mail : hwany1458@empal.com

라, 질감, 모양, 공간 정보 등과 같은 영상의 내용(Content)과 관련된 특징을 추출한 다음, 영상 데이터베이스로부터 미리 추출하여 저장하고 있는 각각의 대상 영상들의 특징과 유사도를 평가하여, 사용자 질의 영상과 유사한 영상을 검색하여 반환하는 기능을 담당한다[4].

본 논문에서는 HSV 칼라 공간에서 특징을 추출함으로써, 밝기와 채도에 덜 민감한 데이터를 효율적으로 검출할 수 있는 영상 검색 기술자를 제안한다. 제안하는 영상 기술자는 입력된 영상의 전체 히스토그램 및 중심 영역과 외곽 영역에서 칼라 분리를 위한 각 영역별 히스토그램을 생성한 후, 칼라 특징 정보를 추출한다. 또한 영상을 일정 비율의 중첩되지 않는 부분 영상으로 나누고 부분 영상별 에지 히스토그램을 추출하여 질감 특징 정보를 추출한다.

본 논문의 구성은 다음과 같다. 2장에서 내용 기반 영상 검색에 대한 기존의 연구 수행 방법을 살펴보고, 3장에서 제안하는 검색 기술자 알고리즘에 대해 설명한다. 4장에서 실험 데이터를 통해 제안 알고리즘의 성능을 분석하고, 5장에서 결론과 향후 연구 방향으로 마무리 한다.

2. 관련 연구

영상 검색과 관련하여, 다수의 범용적인 영상 검색 엔진이 연구 목적으로 또는 상용화된 제품으로 개발되고 있다. 대표적인 내용-기반 영상 검색 방법으로 최초의 IBM의 QBIC을 시작으로, 콜롬비아 대학교의 VisualSeek와 WebSeek, 버클리대학교의 Blobworld, MIT의 PhotoBook 등이 연구 개발되었으며[5], 표준화와 관련하여 MPEG-7 XM이 있다[6]. 이러한 검색 툴들은 내용 기반으로 칼라, 질감, 모양과 공간 정보를 이용하여 영상의 특징을 추출하였다. 이외에 최근 영상 표준인 JPEG에서 JPSearch 분야를 구성하여 검색 방법과 검색 기술자, 메타데이터 포맷 등에 대한 표준화 연구가 활발히 진행되고 있다[1].

본 논문의 주제인 내용 기반 검색에서의 시각적 특징 추출과 관련하여, Jacobs는 Query-by-Sketch 기반의 다중-해상도 영상 검색 알고리즘을 제시하였다[7]. 해당 논문에서는 입력 영상을 사용자의 스케치로 받아 스케닝한 디지털 영상 데이터베이스를 대상으로 검색을 수행하였다. 그러나 그림을 대상으로 데이터베이스를 구축하였으며, 사진 영상과 같이 해상도가 높지 않은 영상을 대상으로만 실험을 수행하였다. Wu는 다중-해상도 영상 검색을 지원하는 검색 기술자를 연구하였다[8]. [8] 연구에서는 텍스트-기반 키워드 검색 방법과

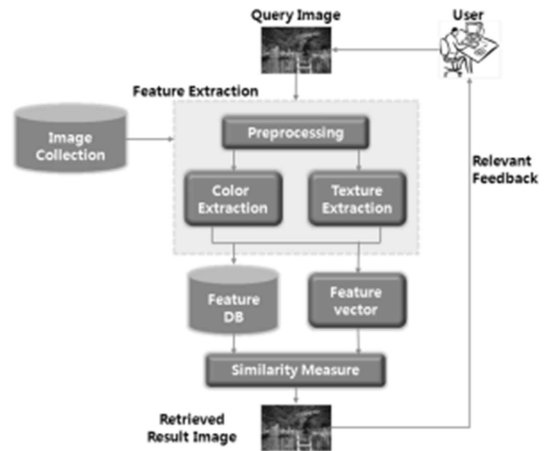


Fig. 1. 제안 영상 검색 시스템 구조도.

칼라 히스토그램을 결합한 형태로 다중-해상도 영상 데이터베이스를 다루었다. 그러나 키워드 생성 과정에서 주관적 해석의 문제점을 가지고 있으며, 다국어 지원이 되지 않아 활용도가 떨어지는 단점을 지닌다. Laila는 내용 기반 검색에서의 시맨틱 갭(Semantic Gap) 문제를 해결하기 위해, 입력 영상으로부터 MPEG-7의 대표적인 기술자인 칼라 레이아웃 기술자(CLD)와 에지 히스토그램 기술자(EHD)를 통해 시각적 특징을 추출하여 검색하고, 사용자의 선택 기능을 제공하여 사용자 피드백을 유도하는 상호동작적인 검색 시스템을 제시하였다[9].

3. 제안 검색 기술자 알고리즘

본 논문에서 제안하는 전체 시스템은 영상의 칼라 정보와 질감 정보를 활용한 검색 기술자로 구성된다. 칼라 기술자에서는 입력된 영상을 HSV 칼라 공간으로 변환한 후, 영상의 전체 히스토그램과 영상의 중심 영역과 외곽 영역에서의 칼라 분리를 위한 각 영역별 히스토그램을 추출하여 특징 벡터를 추출한다. 질감 기술자에서는 입력 영상을 일정 비율의 부분 영상으로 나누어 영역별 부분 영상의 에지 히스토그램을 추출하여 영상에서의 객체 이동에 보다 강건하도록 설계한다. Fig. 1은 제안 기술자를 이용한 전체 검색 시스템 구조를 보여준다.

3.1. 영역 기반 지역 칼라 특징 추출

전역 칼라 히스토그램은 계산이 쉽고 영상 콘텐츠의 회전 및 평행 이동에 강인한 특징을 갖는다. 그러나 영

상의 중심 영역과 외곽 영역에 대한 분리된 특징 정보를 가지지 않기 때문에, 영상의 전체 칼라 분포에 대한 유사도만을 측정할 수 있다. 이는 영상의 조명 강도 변화 및 양자화 오류와 같은 잡음 간섭(Noisy Interference)에 민감하다는 단점이 있다. 또한 영상 특징 표현에 대한 특징 벡터의 차원이 증가하는 문제점을 가진다[10]. 이러한 문제를 해결하기 위해, 제안 알고리즘에서는 중심 영역과 외곽 영역을 분리하여, 각 영역별 히스토그램을 통해 칼라 정보를 추출한다. 영역 분리를 위한 모델 수립을 위해, [15]에서 정의한 영역 분리에 대한 정의를 적용한다.

- 1) 중심 영역의 칼라는 영상의 중앙 부근에 위치한다.
- 2) 영상의 가장 많은 부분을 차지하는 칼라는 외곽 영역으로 정의한다.
- 3) 영상의 외곽 영역 칼라는 영상의 배경 영역에 많은 양이 분포한다.
- 4) 배경 영역에 존재하는 외곽 영역 칼라를 제외한 칼라는 중심 영역 칼라로 정의한다.
- 5) 중심 영역 칼라는 외곽 영역과는 차별화된 칼라 및 질감으로 구성된다.
- 6) 중심 영역은 경계 위치에 비교적 강한 경계 성분을 포함한다.

분리된 영역에 따라, 수식 (1)과 수식 (2)을 이용하여 중심 영역 히스토그램과 외곽 영역 히스토그램을 각각 계산한다.

$$\begin{aligned}
 BG_H &= \sum P_{ij} \\
 (0 \leq i < 1 \times \text{width} / 8 \\
 \text{OR } 7 \times \text{width} / 8 \leq i < \text{width} \\
 \text{OR } 0 \leq j < 1 \times \text{height} / 8 \\
 \text{OR } 7 \times \text{height} / 8 \leq i < \text{height})
 \end{aligned} \quad (1)$$

$$\begin{aligned}
 FG_H &= \sum P_{ij} \\
 (3 \times \text{width} / 8 < i < 3 \times \text{width} / 4) \\
 \text{OR } 3 \times \text{height} / 8 < j < 3 \times \text{height} / 4)
 \end{aligned} \quad (2)$$

수식 (1)에서 BG_H 는 외곽 영역 히스토그램을, 수식 (2)에서 FG_H 는 중심 영역 히스토그램을 의미하며, 각각은 수식 (3)을 이용하여 계산된다. P_{ij} 는 영상의 픽셀을, width 와 height 는 영상의 너비와 높이를 나타낸다.

$$H_{c_i} = \text{prob}_{p \in I} [p \in I_{c_i}] \quad (3)$$

수식 (3)에서 I 는 칼라 영상을, c_i 는 칼라 값을 의미하며, H_{c_i} 는 이미지 영역에서의 칼라 c_i 를 갖는 픽셀의 발생 빈도수를 나타낸다.

3.2. 블록 기반 지역 에지 특징 추출

영상의 질감 특징 추출과정에서는 MPEG-7의 에지 히스토그램 기술자(Edge Histogram Descriptor)를 이용하여 블록 단위의 8 가지 에지 형태를 추출한다. 입력된 영상을 16개(4×4)의 겹치지 않는 블록으로 구분하고, 지역(local) 영역에서 8 개 방향성 에지(Edge)로 표현되는 국부 에지 히스토그램을 계산하여 공간적 분포를 표현한다.

먼저, 에지의 방향성 벡터를 추출하기 위해 영상 블록을 9개의 부분 블록으로 나눈 후, 각 부분 블록에 0부터 8까지 번호를 할당함으로써 (i, j)번째 영상 블록의 9개의 부분 블록에 대한 평균 그레이 레벨을 $A_k(i, j)$ 와 수식 (4)로 표현한다.

$$A_k(i, j) = \frac{1}{s_height \times s_width} \sum_{m=0}^{s_height} \sum_{n=0}^{s_width} I_{ij}(m, n) \quad (4)$$

수식 (4)에서 s_height 와 s_width 는 각각 블록의 넓이와 높이를 의미한다.

나뉘어진 각 부분 블록은 8 방향성(Direction) 에지에 대한 필터 계수를 이용하여 블록의 평균 밝기 값과 컨볼루션(Convolution) 연산을 수행한 크기를 해당 에지의 강도로 나타낸다. 45°방향의 각 8개 에지 방향 검출 필터 계수를 이라 하면, 번째 영상 블록에 대한 8 개 에지 강도 히스토그램은 수식 (5)를 이용하여 계산한다.

$$m_n = \left| \sum_{k=0}^8 A_k(i, j) \times f_{45^\circ}(i, j) \right| \quad (5)$$

수식 (5)에서 A_k 는 수식 (4)로부터 도출된 값이며, f_{45° 는 45°의 에지 검출 필터 계수이다.

계산된 한 개의 영상 블록 내 8 개 에지 강도들 중에서 가장 큰 값을 수식 (6)을 이용하여 구하고, 에지 강도를 가지는 해당 영상 블록은 최대 강도를 갖는 에지의 성분으로 표현된다. m_n 은 수식 (5)로부터 계산된 8 개의 에지 강도 히스토그램이다.

$$\max \left\{ \begin{array}{l} m_0 \cdot (i, j), m_{45^\circ} \cdot (i, j), m_{90^\circ} \cdot (i, j), \\ m_{135^\circ} \cdot (i, j), m_{180^\circ} \cdot (i, j), m_{225^\circ} \cdot (i, j), \\ m_{270^\circ} \cdot (i, j), m_{315^\circ} \cdot (i, j) \end{array} \right\} \quad (6)$$

4. 실험

제안 기술자 알고리즘의 검색 효율성을 평가하기 위해, MPEG-7에서 검색 효율을 평가하는 평가 기준인 ANMRR (Average Normalized Modified Retrieval Rank)를 수식 (7)을 이용하여 계산한다.

$$ANMRR = \frac{1}{Q} \sum_{q=1}^Q NMRR(q) \quad (7)$$

ANMRR은 대표 질의 영상에 대한 질의로부터 검색된 영상들 중에서 질의한 영상과 같은 카테고리에 속하는 영상의 개수와 해당 영상들의 순위를 이용하여 검색 효율성을 평가한다. ANMRR은 0에 가까울수록 좋은 검색 결과를 나타내며[12], 수식 (8)을 이용하여 계산된다.

$$K = \min(4 * NG(q), 2 * GTM)$$

$$AVR(q) = \sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)} \quad (8)$$

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

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

수식 (8)에서, $NG(q)$ 는 질의 영상 q 에 대한 적절한 이미지(Ground truth image)로 정의된 수이며, $K(q)$ 는 모든 질의 q 에 대한 $NG(q)$ 의 최대값인 GTM 에 대해 $\min(4 * NG(q), 2 * GTM)$ 으로 계산된 값이다. $R(k)$ 는 질의 결과에서의 영상 k 에 대한 순위이며, AVR 은 질의 q 에 대한 평균 순위를 의미한다.

실험 데이터로는 MPEG-7에서 제공하는 영상 데이터베이스인 Common Color Dataset(CCD)와 다양한 주제의 자연 영상 데이터셋을 추가하여 사용하였다. CCD에는 총 5,400 여 개의 영상으로 구성되어 있으며, Common Color Query(CCQ)로 50개의 영상과 GTS (Ground Truth Sets)을 제공한다[12]. 이를 바탕으로, 자연 영상 600 여 개를 추가하여 총 6,000 개의 영상 데이터베이스를 이용하여 검색 효율성을 실험 평가하였다. 실험 환경으로, Pentium Dual-Core 3.1GHz, 4GB 메모리, Windows 7 Enterprise 운영체제 환경에서 프로그래밍 툴로 Microsoft Visual Studio 2010을 이용하여 프로토타입 시스템을 구현하였다.

Fig. 2는 제안 검색 기술자의 칼라 특징을 통해 검색한 결과 예시이다. Table 1과 같이, 제안 검색 기술자의 ANMRR의 값이 기존의 대표 칼라 기술자(Domi-



(a) “흰색 꽃” 질의 영상



(b) “흰색 꽃” 질의에 대한 검색 결과

Fig. 2. 칼라 특징 검색 결과 예시.

Table 1. DCD와 제안 방법의 ANMRR 비교

	Dominant Color Descriptor [11]	제안 방법
ANMRR	0.46221	0.22651

nant Color Descriptor) [11] 보다 0.23570만큼 낮은 0.22651를 보임으로서 향상된 검색 효율을 보였다. 전역 칼라 히스토그램(Global Color Histogram) 또는 대표 칼라 기술자(DCD)를 이용한 검색 방법은 전체 영역의 칼라 정보를 활용하는 반면, 제안 알고리즘과 같이 영상의 중앙 영역과 외곽 영역을 분리하여 질의를 수행할 경우, 보다 좋은 검색 효율을 보였음을 알 수 있다.

Fig. 3은 제안 검색 기술자의 질감 특징을 통해 검색한 결과 예시이다. Table 2는 제안 검색 기술자의 질감 특징을 기반한 검색 결과와 MPEG-7 EHD의 검색 결과를 비교하였다. 검색 결과의 ANMRR에서, EHD보다 제안 검색 기술자가 0.18372만큼 낮은 수치인 0.32651 값을 보임으로서 보다 높은 검색 효율성을 보였다.



(a) “새” 질의 영상



(b) “새” 질의에 대한 검색 결과

Fig. 3. 질감 특징 검색 결과 예시

Table 2. EHD와 제안 방법의 ANMRR 비교

	Edge Histogram Descriptor [13]	제안 방법
ANMRR	0.51023	0.32651

5. 결론 및 향후 연구 과제

본 논문에서는 영상을 블록 기반으로 분리하고 칼라와 질감 정보를 활용하여 내용 기반 영상 검색 기술자를 제안하였다. 칼라 특징은 블록 기반으로 중심 영역과 외곽 영역을 나누고 해당 영역에서의 칼라 분포 정보를 추출하였다. 질감 특징은 분리된 블록에서 8개 에지 방향 검출 필터를 통해 에지 강도 성분을 추출하였다. MPEG-7 CCD와 자연 영상을 대상으로, 제안하는 검색 기술자 알고리즘의 성능을 평가하였으며, ANMRR 값에서 기존의 대표 칼라 기술자(DCD)와 에지 히스토그램 기술자(EHD) 보다 각각 0.23570과 0.18372 만큼의 낮은 값을 보임으로써, 보다 좋은 검색 효율성을 가졌다.

본 연구를 통하여 제안한 칼라와 질감 특징 추출 기법은 다양한 다른 기법들과 함께 조합하여 이용될 수 있다. 향후 과제로, 영상의 컨텍스트 정보를 결합하여, 내용 기반 영상 검색의 문제점인 시맨틱 갭(Semantic Gap)을 해결하는 영상 검색으로 확장할 예정이다.

감사의 글

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 콘텐츠산업기술지원사업으로 수행되었음(과제번호: R2012050022).

참고문헌

1. Mun-Kew Leong, Wo Chang, “JPSearch-24800, Part-1: Framework and System Components”, ISO/IEC JTC1/SC29/WG1N3684. 2005.
2. R. Datta, D. Joshi, J. Li, J.Z. Wang, “Image Retrieval: ideas, influences, and trends of the new age”, ACM Computing Surveys, 40(2), pp.1-60, 2008.
3. N. Singhai, S.K. Shandilya, “A Survey on Content Based Image Retrieval Systems”, International Journal of Computer Application, vol.4, 2010.
4. Torres, R.S., Falcao, A.Z., “Content-based image retrieval: Theory and applications”, In Brazilian Symposium on Computer Graphics and Image Processing, pp. 161-185, SIBGRAPO, 2006.
5. Gareth Loy, Jan-Olof Eklundh, “A Review of Benchmarking Content-based Image Retrieval”, MUSCLE/ImageCLEF Workshop on Image and Video Retrieval Evaluation, 2005.
6. Akio Yamada, Robert O’challaghan, S.K. Kim, “MPEG-7 Visual Part of Experimentation Model ver 27.0”, ISO/IEC JTC1/SC29/WG11N7808, 2006.
7. Charles E. Jacobs, Adam Finkelstein, David H. Salesin, “Fast Multi-resolution Image Querying”, ACM Proceedings of Conference on Computer Graphics and Interactive Techniques, pp.277-286, 1995.
8. Li Wu, “Multi-resolution of Image Retrieval in Image Databases System”, International Conference on Artificial Intelligence and Computational Intelligence, pp.243-246, 2010.
9. Laila Shoukry, Sebastian Klenk, Gunther Heidemann, “MPEG-7 Feature Visualization for CBIR Systems”, International Conference on Computer Theory and Applications (ICCTA), pp.86-90, 2010.
10. Ju Han, Kai-Kuang Ma, “Fuzzy Color Histogram and its Use in Color Image Retrieval”, IEEE Transaction on Image Processing, vol.11, no.8, 2002.

11. Wong K. Man, "Content based Image Retrieval using MPEG-7 Dominant Color Descriptor", Master Thesis, City University of Hong Kong, 2004.
12. K. Wong, K. Cheung and L. Po, "MIRROR: An Interactive Content Based Image Retrieval System," IEEE International Symposium on Circuits and Systems, vol. 2, pp. 1541-1544, 2005.
13. Dong Kwon Park Chee Sun Won, Soo-Jun Park, "Efficient Use of MPEG-7 Edge Histogram Descriptor", ETRI Journal, 24(2), 2002.
14. Website, <http://en.wikipedia.org>
15. Sung-Hyun Park, "A Study on Efficient Image Retrieval using Color and Texture Feature", Ph.D. Thesis, Dankook University, 2010.

접수일: 2013년 11월 15일, 심사일: 2013년 12월 2일,
게재확정일: 2013년 12월 16일