

회전에 견고한 DCT 기반 영상 검색

DCT-Based Images Retrieval for Rotated Images

김남이*, 송주환**, 유강수**
 전북대학교 영상공학과*, 전주대학교 교양학부**

Namyee Kim(ynk1115@jbnu.ac.kr)*, Juwhan Song(jwsong@jj.ac.kr)**,
 Kangsoo You(gsyoun@jj.ac.kr)**

요약

일반적인 영상 검색의 경우 질의 영상과 같은 영상 또는 최대한 유사한 영상을 결과로써 보여지게 되는데 회전 영상의 경우에는 일반적인 영상 검색 방법들의 성능이 현저하게 떨어지는 경향이 있다. 이에 따라 본 논문에서는 회전 영상에 대해서도 검색 성능이 우수하도록 하기 위해 질의 영상과 DB 영상에 대해 Harris Corner Detector 알고리즘을 통해 특징점을 구하고, 그 특징점을 토대로 Principal Components Analysis 알고리즘을 이용해 물체의 주축을 구하여 회전 불변 상태의 영상으로 전환한다. 제안한 기법은 Wang의 원본 1000장의 영상과 30°, 45°, 90°, 135°, 180° 로 회전된 총 6000 장의 지름이 256 크기인 자연 영상을 가지고 실험한 결과, 기존의 기법과 비교하였을 때 회전 영상에 대해서도 우수한 검색 성능을 보였다.

■ 중심어 : | Harris 검출기 | PCA | 회전 영상 | 영상 검색 |

Abstract

The image retrieval generally shows the same or similar images to a query image as a result. In the case of rotated image, however, its performance tends to be debased significantly. We propose a method to ensure a reliable image retrieval of rotated images as follows; First, to obtain feature points of query/DB images by Harris Corner Detector; and then, utilizing the feature points, to find the object's axis and query/DB images into rotation invariant images with Principal Components Analysis algorithm. We have experimented with 6,000 natural images which are 256 pixels in diameter. They are 1,000 Wang's images and their rotated images by 30°, 45°, 90°, 135° and 180°. The simulation results show that the proposed method retrieves rotated images more effectively than the conventional method.

■ keyword : | Harris Corner Detector | Principal Components Analysis | Rotated Image | Image Retrieval |

I. 서 론

현대의 컴퓨터와 통신기술의 발달로 영상정보의 생성과 유통이 활발하게 이루어지고 있다. 이에 따라 영상 및 비디오 등을 중심으로 대량의 영상 데이터가 발

생하였고 이를 효율적으로 검색하고자 하는 높은 수준의 멀티미디어 검색 기술에 대한 연구가 진행 되어 왔다. 멀티미디어를 검색 하는 방법에는 다음과 같이 두 가지로 나눌 수 있다[1-3].

첫 번째, 텍스트 기반 검색 방법은 대상이 되는 각각

의 영상에 사람이 문자 주제를 일일이 부과하고 이를 이용하여 영상을 검색하는 방법이다. 이 방법은 제한된 범위 내에서는 정확도가 높다는 장점을 가지고 있지만 방대한 영상 정보에 대해 사람이 일일이 주제를 부여하기가 쉽지 않고, 같은 영상에 대해 다른 주제가 주어졌다면 검색이 불가능하다는 단점이 있다.

두 번째, 내용 기반 검색방법은 영상 내의 특징정보인 컬러(Color), 질감(Texture), 형태(Shape) 등을 자동으로 추출하고, 이를 색인으로 이용하여 검색을 수행함으로써 텍스트 기반 검색 기법의 단점을 극복할 수 있다. 컬러정보는 관측 위치나 크기 변화, 조명 변화 등에 어느 정도 무관하게 적용 가능하다는 장점을 가지고 있다. 질감(Texture)정보는 물체 표면의 거칠 정도와 밝기 변화와 같이 시각적으로 고유한 패턴을 분석하여 나온 물체의 표면 특성이라고 할 수 있다. 형태정보는 표현하는 방법의 어려움으로 인해 색상이나 질감보다는 이용이 제한되어 한정된 시스템에서 사용된다.

영상 검색의 경우, 질의 영상과 같은 영상 또는 최대한 유사한 영상을 결과로써 보여주게 되는데, 다양한 각도로 회전된 영상이 질의 영상으로 입력이 되거나 DB의 후보영상이 회전 되어 있는 경우의 검색에는 상당한 어려움이 따른다.

본 논문에서 제안한 기법은 Harris Corner Detector 알고리즘[4]으로 영상의 특징점을 구하고, 그 특징점들을 PCA (Principal Components Analysis) 알고리즘[5]을 통해, 축을 맞추어 회전 불변 영상으로 변환한다. 그 후, DCT Filter-Bank를 이용해 특징벡터를 구하여 유사 영상을 검색하였다.

본 논문의 구성은 II장에서는 기존의 내용기반 영상 검색 방법에 대하여 기술하고, III장에서는 제안한 방법에 대하여 기술한다. 그리고 IV장에서는 실험방법 및 실험결과를, 마지막으로 V장에서 결론을 맺는다.

II. 관련 연구

일반적으로 내용기반 영상검색에서 주로 사용하는 것은 컬러 히스토그램 방법이다. 이 방법은 영상의 색

상 정보에 대한 표현을 효과적으로 나타낼 수 있으며, 각 색상 값에 속하는 화소의 수에 대한 통계적인 정보를 이용하는 것이 주된 내용이다. 이는 처리 및 구현이 간단하여 많이 사용되고 있다.

Swain과 Ballard[6]가 1991년도에 제안한 히스토그램 인터섹션 기법은 계산과 정보 표현에 있어서 간결하기 때문에 많은 상황에서 유용하게 사용되고 있지만, 조명의 변화와 영상내의 물체 크기에 민감하다는 단점이 있다. 또한 부족한 위치정보로 인하여 영상 내에서의 객체가 가지는 2차원적 공간 정보가 상실될 수 있다는 단점이 있다.

Pass[7] 등은 컬러 히스토그램 방법이 공간정보가 부족한 것을 해결하기 위해 영상 내 화소간의 공간적인 접성을 바탕으로 히스토그램 빈을 나누는 CCV (Color Coherence Vector) 를 제안하였다. 그러나 CCV는 영상상의 화소 위치에 크게 의존하기 때문에 큰 시점 변화에 민감하고, 공간정보를 얻기 위해 주변 화소의 관계를 고려하기 때문에 계산량이 많은 단점이 있다.

최근에는 색상, 질감, 모양 특징이 각각의 장단점을 보완하여 줄 수 있기 때문에 2-3가지의 특징을 혼용하여 검색 성능을 높이는 방법들[8]이 많이 제안되고 있다. 2009년, 컬러공간에서의 DCT Filter-Bank을 이용한 영상검색에 관한 연구[9]는 컬러정보와 형태정보를 반영 할 수 있는 검색 방법을 제안하였다. 이 방법은 다음과 같은 단계를 통해 검색을 수행하였다.

Step 1. [그림 1]과 같이 정규화된 256×256 크기의 RGB 영상에 대한 채널 분리



(a) 원영상 (b) R채널 (c) G채널 (d) B채널

그림 1. 원영상 Indian과 RGB 채널

Step 2 앞 단계의 분리된 각 채널에 대한 DCT 적용 - 8×8 DCT를 수행하여 1개의 DC성분과 63개의 AC 성분을 획득한다.

Step 3 DCT Filter-Bank 생성

- 각 채널에 대한 DC성분에 대해서는 양자화를 통하여 bin의 수를 줄이고 같은 방향의 AC성분 ($AC_{01}:90^\circ$, $AC_{11}:135^\circ$, $AC_{10}:180^\circ$)의 특징 값을 이용하여 히스토그램 feature bin을 획득한다. 추출된 주요한 AC성분에 대해서는 Otsu 이진화를 적용하여, 영상 내의 객체에 대해서 어느 정도의 형태 정보를 얻고, Otsu 이진화가 적용된 AC성분에 대해서 수평/수직방향으로 투영 히스토그램을 이용하여 특징 값으로 산출 하고, 먼저 산출된 DC성분들의 특징 값과 함께 검색에 이용한다. 다음 [그림 2]는 검색을 위한 RGB 채널의 feature bin에 대한 원영상과 히스토그램을 나타낸 것이다.

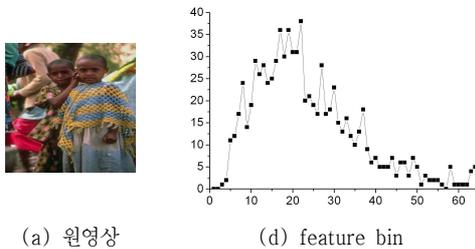


그림 2. 원영상과 DC, AC성분의 feature bin

이 연구[9]는 DCT Filter-Bank의 특성상 AC성분의 방향성을 이용하기 때문에, 질의 영상과의 회전각이 다른 후보 영상을 검색하고자 할 때에는 검색을 하지 못하는 경향이 있다. 따라서 본 논문에서는 컬러공간에서의 DCT Filter-Bank를 이용한 영상검색에 관한 연구[9]의 회전 영상 검색의 비효율성을 극복하기 위해 회전 불변 상태의 영상으로 변환하여 영상 검색을 행한다.

III. 제안한 방법

질의 영상이 임의 각도로 회전되어 있는 경우이거나 데이터베이스에 존재하는 회전 영상에 대한 검색에는 II 장에서 살펴본 것처럼 [9]의 검색 성능이 떨어지는 경향이 있다.

따라서 본 논문에서는 질의 영상과 DB 영상에 대하여

정규화된 256×256 크기를 갖는 회전 불변 영상으로 전환 시킴으로써 영상 검색의 성능을 향상시키도록 한다.

제안한 방법은 첫째, 아래 [그림 3]과 같이 256×256 크기의 원영상 1000장을 각각 30° , 45° , 90° , 135° , 180° 로 회전하고 지름이 256인 원형으로 잘라내어 DB에 기록한다. 둘째, [그림 3]의 회전 변형 영상에 대하여 Harris Corner Detector 알고리즘[4]을 통해 특징점을 추출한다. 이 특징점을 PCA 알고리즘[5]을 이용해 축을 맞추어 회전 불변 영상으로 전환한다. [그림 4]는 Harris Corner Detector 알고리즘을 통해 나타난 특징점과 PCA 알고리즘으로 주축을 찾은 결과를 나타낸 것이다.

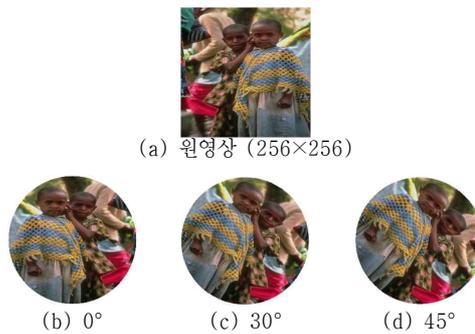


그림 3. 회전 영상과 지름이 256인 원형 영상



그림 4. Harris Corner Detector에 의한 특징점

PCA 알고리즘은 다음과 같은 과정을 거쳐 주축을 찾는다.

1. 물체의 중심 (모든 정점들의 x, y 좌표를 평균)을 구하여 이를 물체좌표계의 원점으로 한다.

$$\begin{pmatrix} m_x \\ m_y \end{pmatrix} = \sum_{i=1}^N \begin{pmatrix} x_i \\ y_i \end{pmatrix} / N \quad (1)$$

여기에서, (x_i, y_i) 는 물체 정점의 좌표, (m_x, m_y)

는 물체 중심점의 좌표, N 은 물체를 구성하는 정점의 수를 의미한다.

2. 식 (1)을 이용하여 공분산 배열 (Covariance Matrix)을 구한다.

$$C = \frac{1}{N-1} \sum_{i=1}^N \begin{pmatrix} cov(x,x) & cov(x,y) \\ cov(y,x) & cov(y,y) \end{pmatrix} \quad (2)$$

여기에서, $cov(x,y)$ 는 다음과 같이 표현된다.

$$cov(x,y) = \frac{1}{N-1} \sum_{i=1}^N (x_i - m_x)(y_i - m_y) \quad (3)$$

단, $cov(x,y) = cov(y,x)$ 이 성립한다.

4. 공분산 배열 C 의 고유 값 (Eigenvalue) 을 구한다.
5. 4 단계에서 구한 고유 값에 의해 고유벡터를 구한다.
6. 고유 벡터를 이용하여 물체의 주축을 구한다.

PCA 알고리즘을 이용하여 고유 벡터에 의해 물체의 주축을 구하였다. 이렇게 주축이 구해지면 주축을 기준으로 회전 불변 영상을 구할 수 있다.

셋째, 영상의 특징 벡터를 구한 후, AC성분에 대해서는 Otsu 이진화를 적용하여, 영상 내의 객체에 대해서 어느 정도의 형태 정보를 얻고, Otsu 이진화가 적용된 AC성분에 대해서 수평/수직방향으로 투영 히스토그램을 이용하여 특징 값으로 산출 하고, 먼저 산출된 DC성분들의 특징 값과 함께 검색에 이용한다.

본 논문에서 제안한 전체적인 검색 시스템의 전체적인 구성도는 [그림 5]와 같다.

IV. 실험결과 및 성능평가

본 논문의 실험은 Microsoft Visual C#.Net 2008로 하였다. 질의 영상과 데이터베이스를 구성하는 후보 영상은 “SIMPLcity: Semantics-Sensitive Integrated matching for Picture Libraries[10]”의 100장씩 10개의 서로 다른 주제를 갖는 총 1000장으로 구성 되어 있는

James Z. Wang[11]의 데이터베이스를 이용하여 1000장에 대해 각각 30°, 45°, 90°, 135°, 180° 로 회전되어 지름이 256의 크기를 갖는 자연 영상으로 구성되어 있다.

또한 회전 질의 영상이나 회전 후보 영상을 위해 테스트 영상을 30°, 45°, 90°, 135°, 180° 회전하여 본 논문에서 제안한 방법과 기존 방법의 성능을 평가하였다. 10종류의 6000장의 자연 영상에 대해 일정한 크기로 정규화하고 특징 정보를 미리 추출하여 데이터베이스에 저장하였다.

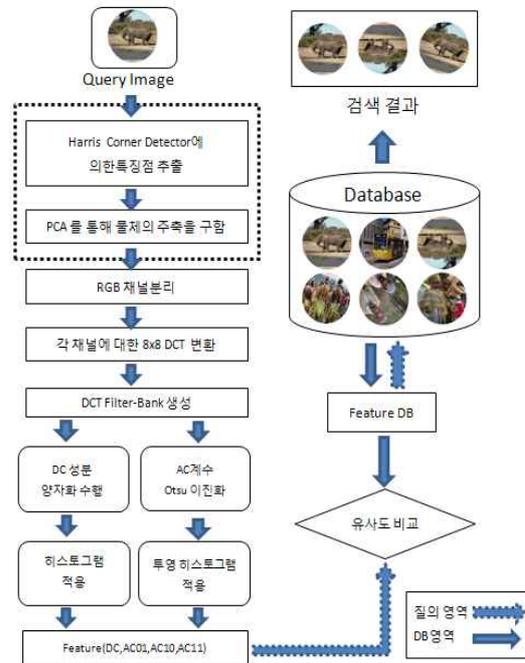


그림 5. 구현한 검색 시스템의 전체 구성도

실험 결과를 통해, 정확율 (Precision Rate) 과 재현율 (Recall Rate) 로 검색 효율성을 평가하여 비교 분석하였다. 다음 식 (4)은 정확율과 재현율을 계산하기 위한 식이다.

$$Recall = \frac{R_r}{T_r}, \quad Precision = \frac{R_r}{T_r} \quad (4)$$

여기에서 T_r 는 검색대상 중 관련된 영상의 총 수, T_r 은 검색된 영상의 총 수를 그리고 R_r 은 검색된 영상 중

관련된 영상의 수를 나타낸다.

[그림 6]은 구현한 검색시스템에서의 Elephant 영상에 대한 검색 결과를 보여주고, [표 1]과 [표 2]는 식 (4)의 정확율과 재현율에 대하여 기존 방법[9]와 제안 방법의 측정치를 나타낸 것이다. [그림 7]은 Indian, Building, Elephant 영상에 대한 성능을 비교한 그래프이다.



그림 6. 구현한 검색 시스템에서의 Elephant 검색

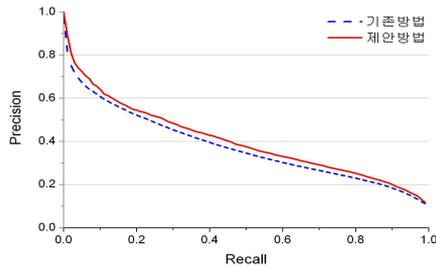
표 1. 정확율 측정 비교

이미지 그룹	기존 방법	제안 방법
indian	0.42	0.45
beach	0.17	0.18
building	0.18	0.20
bus	0.31	0.32
dinosaur	0.98	1.00
elephant	0.31	0.33
flower	0.42	0.41
horse	0.38	0.40
mountain	0.16	0.18
food	0.33	0.36

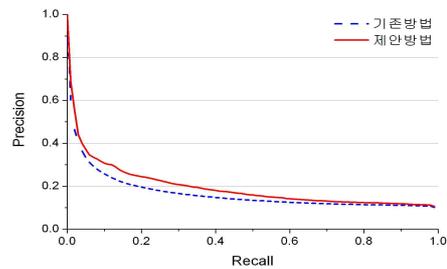
성능평가를 통해, 임의 각도로 회전된 영상에 대한 질의가 있을 경우나 DB에 회전 후보 영상이 있을 경우에도 본 논문에서 제안한 방법을 이용하는 것이 기존 방법보다 효율적인 검색 성능을 얻을 수 있음을 알 수 있었다.

표 2. 재현율 측정 비교

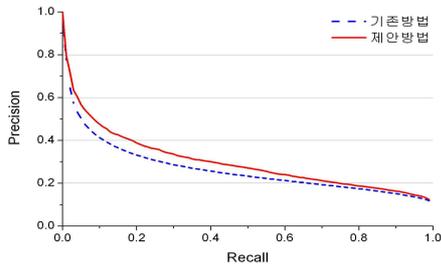
이미지 그룹	기존 방법	제안 방법
indian	0.35	0.37
beach	0.14	0.15
building	0.15	0.16
bus	0.26	0.26
dinosaur	0.82	0.83
elephant	0.26	0.27
flower	0.35	0.34
horse	0.32	0.33
mountain	0.13	0.15
food	0.27	0.29



(a) Indian



(b) Building



(c) Elephant
 그림 7. 구현한 검색 시스템의 성능평가

V. 결론

본 논문에서는 아래와 같은 2 가지 측면에서 영상 검색이 가능하도록 제안하였다.

- 임의 각도로 회전된 질의 영상에 대한 검색
- DB에 회전 후보 영상이 존재할 때의 영상 검색

이러한 검색이 가능하도록 하기 위해서는, 256×256 크기의 영상을 지름이 256 크기가 되도록 정규화한 후, Harris Corner Detector 알고리즘을 통해 특징점을 추출한다. 이러한 특징점을 이용해 PCA 알고리즘으로 물체의 주축을 구해 회전 불변 영상으로 전환하여 실험하였다. 정확율과 재현율을 이용하여 평가한 결과, 기존 방법보다 성능이 우수함을 알 수 있었다.

향후, 배경의 영향이나 조명의 변화, 확대나 축소가 이루어진 영상에 대해서도 검색 성능을 높일 수 있는 연구가 계속되어야 할 것이다.

참 고 문 헌

[1] Y. Rui and T. S. Huang, "Image Retrieval : Current Techniques, Promising, Directions and Open Issues," J. of Computer Visual Communication and Image Representation, Vol.10, pp.39-62, 1999.
 [2] 노형기, 황본우, 문종섭, 이성환, "내용기반 영상 검색 기술의 현황", 대한전자공학학회논문지, 제25

권, 제8호, pp.798-806, 1998.
 [3] 노진구, "내용기반 영상정보 검색기술에 관한 이론적 고찰", 한국도서관정보학회지, 제31권, 제1호, pp.229-258, 2000.
 [4] C. Harris and M. Stephens, "A combined Corner and Edge Detector," Proc. of 4th Alvey Vision Conf., Vol.15, pp.147-151, 1988.
 [5] L. Smith, "A Tutorial on Principal Components Analysis," http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf, 2002.
 [6] M. J. Swain and D. H. Ballard, "Color Indexing," International Journal of Computer Vision, Vol.7, No.1, pp.11-32, 1991.
 [7] G. Pass, R. Zabih, and J. Miler, "Comparing Images Using Color Coherence Vectors," Proceedings of the fourth ACM International Conference on Multimedia, pp.65-73, 1996.
 [8] S. K. Saha, A. K. Das, and B. Chanda, "CBIR using Perception Based Texture and Colour Measures," 17th International Conference on Pattern Recognition(ICPR'04), Vol.2, 2004.
 [9] 김용호, 이석환, 윤용인, 이상근, 최중수, "컬러성분과 DCT Filter-Bank을 이용한 영상검색에 관한 연구", 제22회 신호처리합동학술대회, 2009.
 [10] James Z. Wang, Jia Li, and Gio Wiederhold, "SIMPLcity: Semantics-Sensitive Integrated Matching for Picture Libraries," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.23, No.9, pp.947-963, 2001.
 [11] <http://wang.ist.psu.edu/docs/related.shtml>

저 자 소 개

김 남 이(Namyee Kim)

정회원



- 2006년 8월 : 전북대학교 산업기술대학원 정보기술학과(공학석사)
- 2007년 ~ 현재 : 전북대학교 대학원 영상공학과 박사수료

<관심분야> : 영상처리, 멀티미디어, 컴퓨터비전

송 주 환(Juwhan Song)

정회원



- 1995년 2월 : 전주대학교 전자계산학과 졸업(이학사)
- 1997년 2월 : 전북대학교 전산통계학과(이학석사)
- 2003년 8월 : 전북대학교 전산통계학과(이학박사)

▪ 2006년 9월 ~ 현재 : 전주대학교 교양학부 교수

<관심분야> : 컴퓨터그래픽스, 멀티미디어, 영상처리

유 강 수(Kangsoo You)

정회원



- 1991년 : 전북대학교 컴퓨터공학과(공학사)
- 1994년 : 전북대학교 컴퓨터공학과(공학석사)
- 2005년 : 전북대학교 영상공학과(공학박사)

▪ 2006년 ~ 현재 : 전주대학교 교양학부 교수

<관심분야> : 영상처리, 컴퓨터 비전, 멀티미디어 콘텐츠 관련분야