

레터논문 (Letter Paper)

방송공학회논문지 제20권 제5호, 2015년 9월 (JBE Vol. 20, No. 5, September 2015)

<http://dx.doi.org/10.5909/JBE.2015.20.5.778>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

비정형객체 검색을 위한 향상된 분할영역 색 기술자

정 현 일^{a)}, 김 해 광^{a)†}

Enhanced Boundary Partition Color Descriptor for Deformable Object Retrieval

Hyun-il Jung^{a)} and Hae-kwang Kim^{a)†}

요 약

비정형객체 검색을 위한 향상된 경계 영역 분할에 기반을 둔 색 기술자를 제안한다. 이 논문에서는 하나의 객체를 내부영역과 경계 영역으로 나눠 각각으로부터 기술자를 추출하는 기존의 기술자를 향상하는 기술을 제안한다. 먼저 주어진 영상으로부터 비정형객체영역과 배경영역을 분할한다. 분할된 객체의 중심 좌표를 구하고 이 중심 좌표를 기준으로 해서 전체 객체를 $N \times N$ 블록으로 나눈다. 각 블록은 배경픽셀의 유무를 확인하여 경계블록과 내부블록으로 구분한다. 하나의 객체 영상은 내부블록들로 구성된 내부영역과 경계 블록들로 구성된 경계영역으로 구분하여 각 영역에 대해 MPEG-7 주조색 기술자 (dominant color descriptor)를 추출하여 최종 기술자를 추출한다. 비정형객체 검색 성능 실험을 위하여 핸드백 영상들을 다양한 시점으로 변화하여 구성된 1,973장의 실험 데이터베이스를 구축하였으며, 검색 성능을 ARR(Average Retrieval Ratio)를 사용하여 기존의 MPEG-7 주조색 기술자와 비교하였다.

Abstract

The paper presents a new way of visual descriptor for deformable object retrieval on the basis of partition based description. The proposed descriptor technology partitions a given object into boundary area and interior area and extracts a descriptor from each area. The final descriptor combines these descriptors. From a given image, deformable object is segmented. The center position of the deformable object is calculated. The object is partitioned into $N \times N$ blocks on the basis of the given center position. Blocks are classified as boundary area and interior area depending on the pixels in the block. The proposed descriptor consists of extracted MPEG-7 dominant descriptors from both the boundary and interior area. The performance of proposed method is tested on a database of 1,973 handbag images constructed with view point changes. ARR (Average Retrieval Rate) is used for the retrieval accuracy of the proposed algorithm, compared with MPEG-7 dominant color descriptor.

Keyword : boundary descriptor, deformable object, object retrieval, color descriptor

a) 세종대학교 컴퓨터공학부(Dept. of Computer Eng, Sejong Univ.)

† Corresponding Author : 김해광(Haekwang Kim)

E-mail: hkkim@sejong.ac.kr

Tel: +82-2-3408-3757

ORCID: <http://orcid.org/0000-0002-0302-0627>

· Manuscript received August 3, 2015; Revised September 24, 2015; Accepted September 24, 2015.

1. 서론

영상 검색은 전역 기술자 기술로부터 시작하였으나, 지역적인 정보의 손실로 영상검색 성능에 있어서 문제가 많았다. 계산적으로 복잡한 지역 기술자 기술들이 연구되어 왔으며, 특히 SIFT (Scale Invariant Feature Transform) 기술은 영상 검색뿐만이 아니라 객체 인식, 객체 추적 등 다양한 응용분야에서 많이 사용되고 있으며 속도 등의 성능 향상을 위한 많은 변형이 연구되어 왔다^[1]. MPEG-7에서는 CDVS (Compact Descriptor for Visual Search) 라는 지역 기술자에 대한 표준이 완료되었다^[2]. ACC (Agglomerative Correspondence Clustering) 기술은 하나의 영상과 영상 내의 객체가 비정형적으로 변화된 다른 영상에서 객체들을 정합시키는 기술이다^[3]. 비정형객체의 검색에 있어서는 이러한 기존의 지역 기술자들이 성능을 보이지 못하는데, 이것은 비정형변화에 대한 모델을 정립하는 것이 매우 어렵기 때문이다. 비정형객체 검색 응용을 위하여 하나의 영상 객체를 경계영역과 내부영역으로 구분하여 각각의 영역으로부터 기술자를 추출하여 종합함으로써 기존의 전역기술자의 한계를 극복하는 분할영역 기술이 제안되었다^[4]. 이 논문에서는 분할영역기술의 성능을 향상시키는 기술이 제안된다. 실험을 위한 전역 기술자로서는 MPEG-7 전역 기술자 중에서 구조색 기술을 사용한다^[5]. 2장에서는 논문에서 제안하는 분할영역 기술자 추출 과정 알고리즘을 설명하고 3장에서는 성능 실험 결과에 설명하고 분석하며 4장에서 결론을 맺는다.

II. 제안하는 분할영역 색 기술자 추출 과정

그림 1은 제안하는 기술자의 추출과정을 보여준다. 비정형객체가 포함된 영상이 입력되면 영상은 객체영역과 배경영역으로 분할된다. 객체 영역으로부터 중심위치가 계산되며, 중심위치를 기반으로 $N \times N$ 블록으로 분할된다. 각 블록들은 배경 픽셀의 포함 유무에 따라 내부블록과 경계블록으로 구분된다. 내부블록으로 구성된 내부영역과 경계블록으로 구성된 경계영역에 대해 각각 MPEG-7 구조색 기술자 추출과정을 적용한다. 추출된 2 개의 MPEG-7 구조색 기술자를 결합하여 분할영역 기술자를 만든다. 제안하는

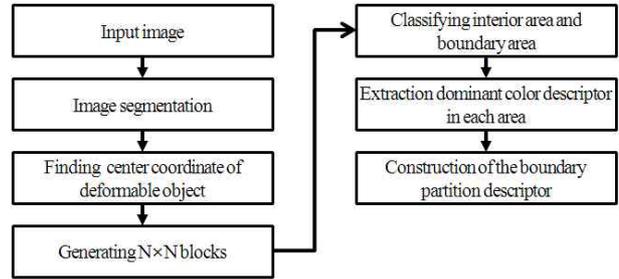


그림 1. 분할영역 색 기술자 과정
 Fig. 1. The process of boundary partition color descriptor

기술이 기존의 분할영역 기반 기술과 다른 점은 기존의 방법은 각 블록으로부터 기술자를 먼저 추출하고, 경계영역 기술자는 경계영역에 해당하는 블록들의 기술자들의 평균으로 구하고 내부영역 기술자는 내부영역에 해당하는 블록들의 기술자들의 평균으로 구하는데 반해, 제안하는 기술은 블록들마다 기술자를 먼저 구하지 않고 경계영역을 구성하는 블록들로 구성된 영역으로부터 하나의 기술자를 추출하고 내부영역을 구성하는 블록들로 구성된 영역으로부터 하나의 기술자를 추출하는데 있다. 기존의 기술은 블록단위로 기술자의 통계로 기술자를 구성하기 때문에 경계영역과 내부영역의 전역적인 구조색 특징을 표현하지 못한다. 제안하는 기술은 경계영역, 내부영역의 각 영역 전체에 대하여 하나의 기술자를 추출하여 영역 전체의 구조색 특징을 구할 수 있다. 자세한 세부 추출과정을 아래 기술한다.

1. 객체의 위치와 크기에 따른 블록 영상 구현

그림 2는 주어진 영상으로부터 $N \times N$ 의 블록 객체 영상을 만드는 과정을 보여주고 있다. 그림 2 a)는 주어진 원 영상이고 그림 2 b)는 객체분할에 의해서 얻어진 객체 영상을 보여준다. 객체의 중심 좌표는 MPEG-7 질감 기술자 (Homogeneous texture descriptor) 에서 사용되는 방법을 이용하여 계산하였다. 객체 영역 내부만을 포함하는 가장 큰 사각형을 찾아 이 사각형의 중심 좌표를 객체 영역의 중심 좌표로 사용한다. 그림 2 c)는 객체내부의 가장 큰 사각형과 사각형의 중심인 객체의 중심좌표를 보여준다. 그



그림 2. 영상의 N×N 블록 분할 과정
 Fig. 2. The N×N block splitting processing of image

림 2 d)는 객체의 중심좌표를 기반으로 구해진 9 × 9의 블록객체 영상을 보여준다. 객체 영역을 N × N 개의 블록으로 분할하는데 이 중심좌표가 한 블록의 중심위치가 되도록 분할한다. N은 고정되며 제안기술의 실험에서는 5 × 5와 9 × 9의 블록개수가 사용되었다. 블록개수가 고정됨으로써, 블록들의 크기는 객체의 크기로부터 정해진다. 핸드백 객체의 중심 좌표에서부터 N × N 개의 블록을 만들어 나가야 하기 때문에 전체 블록 객체 영상의 크기는 핸드백 객체 영역의 크기와 같거나 그보다 크다.

2. 내부 및 경계 영역 구분 및 기술자 추출

먼저 각 블록은 블록 내부에 하나의 픽셀이라도 배경 값일 경우 경계블록으로, 그렇지 않은 경우 내부블록으로 한다. 그림 3은 각각 경계영역과 내부영역으로 나뉘어 만들어진 영상을 보여준다.

내부영역과 경계영역 각각에 대해 MPEG-7 주조색 기술자를 추출한다. MPEG-7 기술자는 주조색에 대한 색상, 색 분포, 분산 그리고 공간 상관성으로 구성된다. 제안한 기술자는 경계영역에서 추출한 MPEG-7 주조색 기술자와 내부영역에서 추출한 MPEG-7 주조색 기술자로 구성한다.

III. 실험 결과 및 분석

제안된 기술의 성능 실험을 위하여 핸드백 영상 데이터베이스를 구성하였다. 실험 775개의 다양한 핸드백 영상으로부터 시점변화를 적용하여 총 1,973장의 핸드백 영상으로 구성하였다. 그림 4는 데이터베이스에 있는 핸드백 영상과 시점 변화된 핸드백 영상들의 예를 보여준다. 각 줄의 맨 왼쪽 영상이 원 영상이며, 오른쪽에 있는 영상들이 원 영상의 핸드백이 시점 변화된 영상들이다. 성능 실험은 1,973개의 데이터베이스에 있는 모든 영상을 질의로 하여 해당하는 시점 변환된 영상들을 얼마나 잘 검색하는가를 ARR(Average Retrieval Rate)로 측정하였다. 식 (1)은 ARR을 구하는

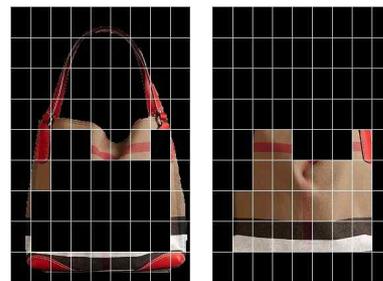


그림 3. 영역 구분 영상: 경계블록 영상(왼쪽), 내부영역 영상(오른쪽)
 Fig. 3. The boundary and interior block images



그림 4. 핸드백 영상 데이터베이스
 Fig. 4. The database of handbag image

식을 보여준다. ARR은 RR (Retrieval Rate)의 평균으로 구한다. NQ 는 총 질의 영상의 수이다 (실험에서는 1,973). 하나의 질의 영상 q 에 대해 RR_q 을 구하는 식에서 NG_q 는 질의 영상에 해당하는 시점 변환된 영상의 수 (정답 수), NR_q 는 가까운 거리로 정렬된 영상 중에서 상위 NG_q 의 2배 안에 포함되는 시점 변환된 영상의 수이다.

$$ARR = \frac{1}{NQ} \sum_i RR_i, \quad RR_q = \frac{NR_q}{NG_q} \quad (1)$$

표 1은 제안한 알고리즘을 5×5 및 9×9 블록개수로 했을 때의 성능과 추출시간을 기존의 MPEG-7 주조색 기술자 (DCD), 기존의 분할 기술자 (BIDCR) 와 비교한 결과이다. ARR로 측정된 검색성능에 있어서 DCD를 사용할 때의 결과가 37.30%에 비해 기존의 분할기술자의 성능은 37.75%로서 거의 성능 향상이 미약하다. 하지만 논문에서 제안한 기술의 5×5 의 경우가 39.39%로서 상당한 의미 있는 성능향상을 보이며 9×9 의 경우 43.5%로서 블록의 개수를 증가함으로써 더욱 많은 성능 향상을 보인다. 객체 영역을 $N \times N$ 개의 블록으로 나눔에 있어서 N 이 증가할수록 블록의 크기가 감소한다. 따라서 경계영역과 내부영역을 보다 정교하게 구분하여 표현할 수 있다. 가장 좋은 결과인 9

표 1. 제안한 알고리즘과 기존 기술자의 실험 결과 비교
 Table 1. The comparison of ARR extraction time

	DCD	BIDCR	Proposed (5×5)	Proposed (9×9)
ARR(%)	37.30	37.75	39.39	43.50
Time(/s)	0.32	1.48	1.49	1.51

$\times 9$ 의 제안된 기술이 달성한 43.5%의 결과도 유용성의 측면에서 매우 부족한 수치인데 이것은 비정형객체의 검색이 매우 힘들기 때문이다. 제안한 기술자를 추출하는데 걸리는 시간은 MPEG-7 주조색 기술자와 달리 두 개의 영역으로 나눠 기술자를 추출하기 때문에 시간이 더 소모된다. 기술자의 바이트 크기는 제안하는 기술의 내부영역과 경계영역에서 각각 기술자를 추출하기 때문에 DCD에 비해 2배가 크다는 단점이 있다.

IV. 결 론

이 논문은 비정형 객체의 검색을 위한 기술자로서 내부 영역과 경계영역에 대해 각각 기술자를 추출하여 검색 성능을 향상시키는 방법을 제안하고 실험하였다. ARR로 측정된 검색 성능에 있어서 제안한 방법이 기존의 기술자에 비해 상당한 향상을 보여준다. 또한 블록개수의 증가에 있어서 보다 나은 성능을 보여준다. 제안한 기술자는 추출속도와 기술자의 바이트 크기에 있어서 기존의 방법에 비해 좋지 않은 단점이 있다. 앞으로 알고리즘 검색 성능의 향상, 최적 블록개수의 설정, 추출시간 단축, 바이트 크기의 축소 및 압축에 대한 연구를 진행할 예정이다.

참 고 문 헌 (References)

- [1] D. G. LOWE, "Distinctive Image Features from Scale-Invariant Keyoints", International Journal of Computer Vision, vol. 60, pp.91-110, November, 2004
- [2] ISO/MPEG N4320, "MPEG-7 Requirements Document,v 15", F. Pereira, ed., MPEG Requirements Group, Sydney, July 2001
- [3] M. Cho J. Lee and K. Lee, "Feature Correspondence and Deformable Object Matching via Agglomerative Correspondence Clustering", IEEE 12th International Conference on Computer Vision, pp.1280-1287, October, 2009
- [4] H. I. Jung and H. K. Kim, "Boundary and Interior Dominant Color Descriptor for Handbag Deformable Object Retrieval", Workshop on Image Processing and Image Understanding, February, 2015
- [5] B. S. Manjunath, Jens-Rainer Ohm, Vinod V. Vasudevan, and Akio Yamada, "Color and Texture Descriptors", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 11, NO. 6, pp.703-715, June, 2001