

레터논문 (Letter Paper)

방송공학회논문지 제18권 제2호, 2013년 3월 (JBE Vol. 18, No. 2, March 2013)

<http://dx.doi.org/10.5909/JBE.2013.18.2.319>

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

## 영상 검색을 위한 점진적 블록 크기 기반의 효율적인 손실 좌표 압축 기술

최 경 민<sup>a)</sup>, 정 현 일<sup>a)</sup>, 김 해 광<sup>a)†</sup>

### Gradual Block-based Efficient Lossy Location Coding for Image Retrieval

Gyeongmin Choi<sup>a)</sup>, Hyunil Jung<sup>a)</sup>, and Haekwang Kim<sup>a)†</sup>

#### 요 약

MPEG-7 CDVS (Compact Descriptor for Visual Search)분야에서 표준화하고 있는 현대의 모바일 디바이스 및 서버에서 사용되는 영상검색과 매칭 알고리즘들은 SIFT(scale invariant feature transform)와 SURF(speeded up robust features) 같은 강인한 디스크립터를 기반으로 하는 특징 점에 의한 알고리즘으로 이루어진다. 이러한 특징 점들은 크게 좌표와 디스크립터로 나누어져 있다. 빠르고 정확한 검색을 위해서 특징 점들은 디바이스에서 서버, 또는 서버에서 디바이스로 자유롭게 전송이 되어야 하므로 과거에 여러 압축 알고리즘들이 제안되었다. 이 논문에서는 특징 점들의 분포 및 연관성 등을 관찰하고 연구하여 좌표의 정보를 효율적으로 압축하면서 정확도를 보존할 수 있는 점진적 블록 크기 기반의 손실 좌표 압축 알고리즘을 제안한다. 실험 결과로부터 현재 가장 효율이 좋은 알고리즘 보다 특징 점당 비트가 평균적으로 0.3~0.4bit(5%~6%) 감소하고 정확도(TP,FP,TN)가 데이터 종류에 따라 유지되거나 미약하게 상승하는 결과를 얻었다.

#### Abstract

Image retrieval research activity has moved its focus from global descriptors to local descriptors of feature point such as SIFT. MPEG is Currently working on standardization of effective coding of location and local descriptors of feature point in the context mobile based image search driven application in the name of MPEG-7 CDVS (Compact Descriptor for Visual Search). The extracted feature points consist of two parts, location information and Descriptor. For efficient image retrieval, we proposed a novel method that is gradual block-based efficient lossy location coding to compress location information according to distribution in images. From experimental result, the number of average bits per feature point reduce 5~6% and the accuracy rate keep compared to state of the art TM 3.0.

Keyword : MPEG-7 CDVS, image retrieval, Lossy location coding, Gradual block-based

a) 세종대학교 컴퓨터공학과 (Department of Computer Engineering, Sejong University)

† Corresponding Author : 김해광(Haekwang Kim)

E-mail: [hkkim@sejong.ac.kr](mailto:hkkim@sejong.ac.kr)

Tel: +82-31-745-7881

※ 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 콘텐츠 산업 기술 지원 사업으로 수행되었음.

· Manuscript received December 12, 2012 Revised February 27, 2013 Accepted February 27, 2013



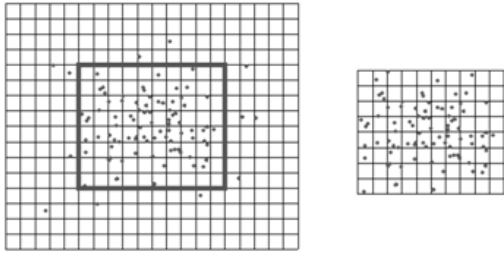


그림 3. Peking 대학교 Core-area 방법, 중앙의 특징 점이 밀집된 곳만 잘라서 코딩하는 방법  
 Fig 3. Core-area method from Peking, the method by using core-area for compression

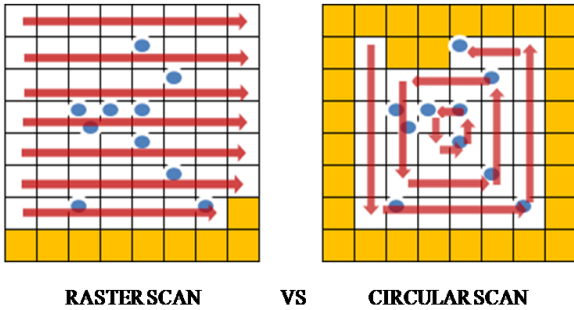


그림 4. Huawei의 Circular scan orderd 방법, 압축의 스캔 순서를 이미지의 중앙으로부터 외각으로 하는 방법  
 Fig 4. Circular scan order method from Huawei, the method that compression begins from center of image to outcorner by circular way

하지만 특징 점의 분포를 보고 더욱 효율적인 위치 정보 (x,y)의 압축을 위해 점진적인 블록을 사용하는 알고리즘을 고안하게 되었다.

## II. 제안 알고리즘

앞선 기술 소개의 그림 3에서와 같이 SIFT나 SURF로 추출된 특징 점들은 대부분 영상의 중간에 모여 있는 것을 알 수 있다. 따라서 이 전의 모든 알고리즘들이 선택했던 균일 블록 기반의 압축에서 정확도 유지와 압축률 향상을 위해 유동적인 블록 크기 기반의 알고리즘을 제안한다.

SIFT에 의해 특징 점이 추출된 이미지의 위치 정보(x,y)를 압축하기 위하여 크게 3단계의 압축단계로 나눌 수 있

고, 각각의 단계는 영역 1~ 영역 3까지 구분되어 순차적으로 압축 된다.

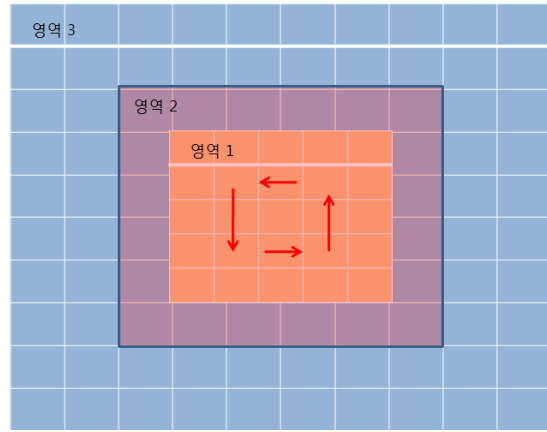


그림 5. 제안 하는 알고리즘 영역 1~영역 3이 나타내는 이미지의 공간 예시  
 Fig 5. Proposed algorithm example Area1~ Area3 in image

표 1. 6x6블록 크기 영상 일 때 각 영역 정보  
 Table 1.Each area information as 6x6 block size image

	블록크기	Context 방법	스캔 오더	특징점블록 분포량
영역 1	5x5	sum[0]	circular	많음
영역 2	7x7	sum[0]	circular	중간
영역 3	7x7	sum[n]	circular	적음

먼저 영역 1에서는 표 1에서와 같이 특징 점의 분포량이 가장 많은 구간이기 때문에 sum-based cotext를 사용하지 않고 이미 트레이닝 된 cotext 정보를 이용하여 6x6보다 사이즈 1작은 5x5블록으로 인코딩 한다. 이때 많은 블록으로 인해 정확도는 향상되지만 압축률은 떨어진다.

두 번째 영역 2의 코딩은 영역 1 지역보다 듬성듬성한 특징 점들의 분포를 가지기 때문에 5x5 사이즈가 아닌 기존의 6x6보다 사이즈 1 큰 7x7 블록 사이즈로 코딩하게 되고 마찬가지로 중심에서 가까운 쪽부터 circular scan하여 코딩한다. 영역 2는 영역 3을 코딩하기 위해 sum-based context 정보 수집을 하기 위한 기반을 마련하는 역할을 하게 되고 마찬가지로 미리 트레이닝 한 영역 2의 context를 사용하여 압축한다.

끝으로 영역 3 코딩은 미리 인코딩된 영역 2의 블록 정보

를 이용하여 코딩하게 되는데 기존의 6x6 보다 1 사이즈 큰 7x7 블록사이즈를 이용하여 인코딩하게 된다. 한 가지 중요한 점은 매우 듬성듬성 분포한 특징 점 블록들의 상호 관계가 중심 지역만큼 많지 않기 때문에 주변에 몇 개의 특징 점을 포함한 블록이 있느냐 에 따라 context가 달라지는 sum-based context 코딩을 하게 된다. 또한 기존의 6x6 으로 코딩 했을 때 보다 블록의 개수는 적어 비트효율은 좋아지지만 그만큼 정확도는 떨어지게 되는데 이는 영역 1~ 영역 3까지 각각 정확도와 비트효율 상호 효과로 전체적으로는 더 좋은 결과를 얻게 된다.

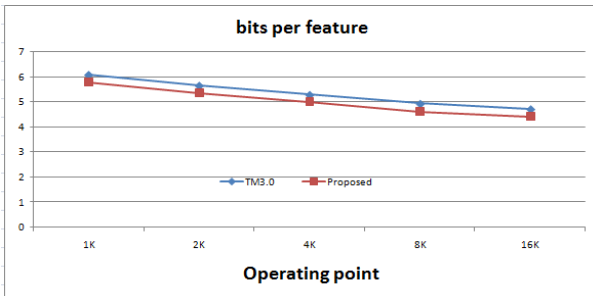


그림 6. operating point 별 bits per feature  
Fig 6. bits per feature on each operating point

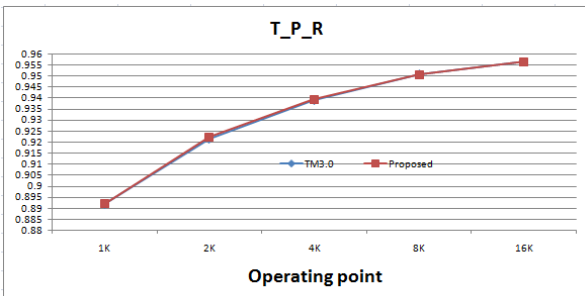


그림 7. operating point 별 True Positive 값  
Fig 7. True positive rates on each operation point

### III. 실험 및 결과

모든 실험은 MPEG-7 CDVS standard dataset을 사용하고, TM3.0의 알고리즘과 제안된 알고리즘을 적용한 각각의 카테고리별 영상들의 Operation point 별 평균 값들을 이용하여 그래프로 나타냈다. 비교요소는 압축률을 확인할 수 있는 특징 점 하나당 비트수와 정확도를 확인할 수

있는 TP값을 기준으로 정리하였다. 그림 6 에서와 같이 기존 TM3.0<sup>[9]</sup>보다 특징 점당 비트수가 약5%~6%정도감소한 것을 볼 수 있고 그림 7에서와 같이 TP값은 TM3.0과 거의 비슷하거나 데이터 카테고리 종류에 따라 미세하게 높거나 낮은 것을 확인 할 수 있다.

### IV. 결론 및 향후 연구방향

이 논문에서는 빠르고 정확한 영상 검색을 위한 점진적 블록 크기 기반의 효율적인 손실 좌표 압축 기술을 제안 했다. 영상의 특징 점들의 분포가 2차원 가우시안 형태를 띠는 것을 관찰하여 점진적 블록 크기를 이용하여 TM3.0과 비교하여 정확도를 유지하고 압축률을 5%~6% 높였고 이후 특징 점들의 분포를 더욱 자세히 분석하여 보다 높은 효율을 가진 알고리즘 개발을 계획하고 있다.

### 참 고 문 헌

- [1] Yuri Reznik. Compact Descriptors for Visual Search: Applications and Use Scenarios Requirements Subgroup.MPEG N11529. 2010.
- [2] ISO/MPEG N4320, MPEG-7 Requirements Document, v 15, F. Pereira, ed., MPEG Requirements Group, Sydney, July 2001.
- [3] D. Lowe, Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110,November 2004.
- [4] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF:speeded up robust features," in European Conference on Computer Vision, Graz, Austria, May 2006, pp.404-417.
- [5] S. Tsai, D. Chen, G. Takacs, V. Chandrasekhar, J. P. Singh, and B. Girod, "Location Coding for Mobile Image Retrieval," in Proc. International Mobile Multimedia Communications Conference, Sept. 2009.
- [6] ISO/IEC JTC1/SC29/WG11/M24758, "Improvements to location coding : Reponse to CE3 - Feature Point Location Coding", Geneva, CH, April 2012
- [7] ISO/IEC JTC1/SC29/WG11/M24782, "Peking University's Response to the CDVS Core Experiment 3", Geneva, CH, April 2012
- [8] ISO/IEC JTC1/SC29/WG11/M24802, "Huawei's response to CDVS Core Experiment 3", Geneva, CH, April 2012
- [9] ISO/IEC JTC1/SC29/WG11/W12929, "Test Model 3: Compact Descriptors for Visual Search," Stockholm, SE, Jul. 2012.