

MPEG-7 칼라 스트럭처 기술자와 도미넌트 칼라 기술자 추출의 최적화 방법

유정수^o 낭종호

서강대학교 컴퓨터공학과

yjs@mlneptune.sogang.ac.kr jhnang@sogang.ac.kr

An Optimization Method for Extraction of MPEG-7 Color Structure Descriptor and Dominant Color Descriptor

Jeongsoo Yu^o Jongho Nang

Department of Computer Science and Engineering, Sogang University

MPEG-7[1]은 멀티미디어 메타데이터 기술 표준으로서 파트 3인 Visual 파트에서는 이미지/비디오 검색에 이용할 수 있는 시각 정보 기술자가 정의되어 있다. 일반적으로 색상은 매우 효과적으로 사용이 가능한 시각 특성 중 하나이며, 현재 MPEG-7 color structure descriptor(CSD)와 dominant color descriptor(DCD)를 포함하여 여러 색상 기술자를 정의해 놓았다.

시각 정보의 추출은 이미지 검색 및 분류 등 여러 응용에 널리 쓰이고 있다. 그러나 대부분 추출 비용이 매우 높다는 문제점이 있다. 특히 최근 컴퓨팅 성능이 제한된 휴대 기기 환경에서도 이와 같은 저 수준 정보를 사용하는 응용이 많이 개발되는 추세여서 이와 같은 저 수준 정보 추출의 비용을 감소시키는 방법이 필요하다고 볼 수 있다.

MPEG-7 색상 기술자들 중 CSD는 가장 우수한 검색 성능을 나타내며[2], DCD는 매우 컴팩트한 특징 데이터 포맷을 가지므로 인덱싱에 유리하다는 장점을 갖는다. CSD는 전역적인 칼라 히스토그램과 지역적인 구조를 동시에 표현한다. 추출은 먼저 HSV 공간의 픽셀을 HMMD 색상 양자화를 거친 후 정사각형 윈도우(8 × 8 내의 각 칼라의 출현 여부를 카운팅하는 방식을 통해서 공간 적인 칼라 구조를 결정하게 된다. DCD는 이미지 전체의 칼라를 몇 개의 dominant color로 양자화 하여 각각의 비중과 공간적 응집도를 통해서 기술된다. 칼라 양자화는 Generalized Lloyd Algorithm (GLA)[3]를 통해서 이루어진다. 한편, MPEG-7은 기술 형식에 대하여 정의함과 동시에 XM(eXperimental Model)[4]를 통해서 추출 및 비교 방법에 대하여 구체적으로 제시하고 있다. [표 1]은 640 × 480 이미지에 대하여 400MHz ARM9 PDA 환경에서 XM에서 제공하는 실험용 프로그램을 사용하여 CSD와 DCD의 추출 성능을 측정한 것으로 실제 추출 비용이 매우 높은 것을 알 수 있다.

본 논문에서는 CSD와 DCD의 추출 성능 향상 방법에 대하여 제안하였다. CSD의 경우는 8 × 8 윈도우 이동 시 이전 단계와 중복되는 픽셀 읽기를 제거하는 방법을 통해 ARM9 환경에서 기존의 약 20% 수준으로 추출 시간을 감소시켰다. DCD의 경우는 GLA 대신 LBA[5]방법을 사용하였고, [5]의 LBA를 통한 클러스터링 과정 중 그룹핑 단계에서 일반적인 이미지의 특성을 이용한 ARP (Approximate-Rectangle-Partition) 방법을 사용하여 추출비용을 ARM9 환경에서 기존 LBA 적용 시의 약 15% 정도로 현저하게 감소시켰다. [표 2]와 [표 3]은 각각 CSD와 DCD에 대하여 본 논문에서 제안한 최적화 방법의 성능 향상 결과이다. 제안한 방법을 적용하여 추출 성능을 향상시킴으로써 멀티미디어 데이터의 인덱싱을 구축하거나 휴대 기기 환경에서 저 수준 특징을 이용한 응용의 개발 시 전체 시스템의 성능 향상에 도움이 될 것이라고 기대된다.

표 1 : MPEG-7 색상 기술자의 추출 시간 (640×480, 400MHz ARM9)

	Color Structure	Scalable Color	Color Layout	Dominant Color
추출 시간(ms)	4,500ms	250ms	150ms	23,000ms

표 2 : CSD 최적화의 성능 측정 (640×480, 400MHz ARM9 PDA)

		최적화 전	최적화 후
픽셀 읽기		$(W-8)(H-8)(8 \times 8)$	$(W-8)(H-8)8 + (H-8)(8 \times 8)$
Shift (register operation)		—	$(W-8-1)(H-8-1) \times M$
Zero-setting		$(W-8)(H-8) \times M$	$(H-8) \times M$
추출 시간	PC	190ms	30ms
	PDA	4,500ms	900ms

표 3 : DCD 최적화의 성능 측정 (640×480, 400MHz ARM9 PDA)

	GLA	LBA	LBA with ARP/RARP
추출 시간	23,000ms	1,200ms	170ms

참고 문헌

[1] ISO/IEC JTC1/SC29/WG11, *Information Technology Multimedia Content Description Interface-Part3 : Visual*.2001

[2] T. Ojala, M. Aittola & E. Matinmikko, " Empirical evaluation of MPEG-7 XM color descriptors in content-based retrieval of semantic image categories," in *Proceedings of International Conference on Pattern Recognition*, pp. 1021- 1024, 2002.

[3] A.Gersho and R.M.Gray, " Vector Quantization and Signal Compression," *Kluwer Academic Publishers*,1993.

[4] ISO/IEC JTC1/SC29/WG11, *MPEG-7 Visual part of eXperience Model* Version 11.0.2001.

[5] N.C.Yang, W.H.Chang, C.M.Kuo, T.H.Li, " A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval," *Journal of Visual Communication and Image Representation*, 2008