

MPEG-7 칼라 및 에지 서술자를 이용한 영상 검색

강 희 범, 박 동 권, 원 치 선, 박 수 준*

동국대학교 전자공학과 / 한국전자통신연구원*

Image Retrieval Using MPEG-7 Color and Edge Descriptors

Hee Beom Kang, Dong Kwon Park, Chee Sun Won, and Soo Jun Park*

Dept. of Electronic Eng. Dongguk University, Seoul, 100-715, Korea

*ETRI, Taejon, 305-600, Korea

E-mail: cswon@dongguk.edu

요 약

본 논문에서는 MPEG-7의 에지 히스토그램 서술자와 컬러 레이아웃 서술자를 조합하여 검색 성능을 향상시키는 방법을 제안한다. 에지 특징 정보는 영상의 콘텐츠를 표현하기 위한 중요한 요소로 고려되어진다. 이것은 특별히 컬러의 단점을 보완하기 위해서 유용하다. 반면 컬러 특징 정보는 구조적인 단순함과 빠른 동작 속도에 의해 영상 검색에 넓게 사용되어진다. 본 논문에서는 앞에서 언급한 두 특징 정보를 잘 표현하는 에지 히스토그램 서술자와 컬러 레이아웃 서술자를 사용하였다. 실험 결과로 제안한 방법이 자연 영상에 대하여 두드러지게 검색 성능을 향상시켜주는 것을 확인할 수 있다

1. 서 론

디지털 방송 TV의 도입에 따른 채널 수의 증가와 인터넷의 보급에 따라 많은 양의 디지털 비디오 및 영상 정보가 저장되고 있다. 따라서, 사용자는 관심을 갖는 콘텐츠 또는 프로그램에 대한 빠르고 효율적인 접근을 위한 도구가 필요하다. 이와 같은 문제를 해결하기 위해 현재 MPEG-7에서 많은 서술자들이 개발되어[1] 우수한 검색 성능이 입증되었다. 본 논문에서는 검색 성능을 더욱 향상시키기 위하여 MPEG-7의 에지 히스토그램 서술자[1][2]와 컬러 레이아웃 서술자[1][5][6]를 조합하여 검색 성능이 우수한 서술자를 선택하는 방법에 의한 검색을 제안한다. 에지 특징 정보는 영상의 콘텐츠를 표현하기 위한 중요한 요소로 고려되어진다. 이것은 특별히 컬러의 단점을 보완하기 위해서 유용하다. 반면 컬러 특징 정보는 구조적인 단순함과 빠른 동작 속도에 의해 영상 검색에 넓게 사용되어진다. 그러므로, 위에서 언급한 두 특징 정보를 잘 표현하는 에지 히스토그램 서술자와 컬러 레이아웃 서술자는 서로 보완적인 관계로 사용됨으로써 향상된 영상 검색 결과를 만들 수 있다. 제안한 방법은 두 단계로 나누어진다. 첫번째 단계는 MPEG-7 서술자들을 조합하여 1차 검색 결과를 보여준다. 두번째 단계는 1차 검색 결과를 토대로 적합한 서술자를 선택하고 렐러번스 피드백[3][4]에 의하여 사

용자가 만족할 때 까지 반복된다. 본 논문의 구성은 다음과 같다. 2장에서는 실험에 사용한 MPEG-7 서술자에 대하여 기술하고, 3장에서는 MPEG-7 서술자들을 조합하는 방법에 대하여 설명하며, 4장에서 MPEG-7 서술자를 선택하는 방법에 대하여 설명하고, 5장에서 실험 결과 및 성능을 보이고 마지막 장에서 제안한 방법의 결론을 맺는다.

2. MPEG-7의 칼라 및 에지 서술자 소개

에지 히스토그램 서술자는 다섯 가지 에지 형태를 국부(local) 영상 영역에서 공간적 분포로 표현한다. 네개의 방향성 에지와 한개의 비방향성 에지가 있는 각 국부 영역은 부영상(sub-image)이라 한다[1]. 보다 명확하게 말하면 16개의 겹치지 않는 각각의 부영상에 대해 5개의 빈(bins)으로 구성된 국부 에지 히스토그램이 생성된다. 각각의 부영상에 대하여 5가지 형태의 에지가 존재하므로 총 $16 \times 5 = 80$ 의 히스토그램 빈(histogram bins)을 갖는다. 게다가, 부영상은 이미지 블록(image block)으로 나뉘므로 에지 형태 정보는 이미지 블록으로부터 추출되어진다[2]. 컬러 레이아웃 서술자는 컬러의 공간적 분포로 표현된다. 영상 또는 비디오 매칭에 사용되는 것 외에도 스케치 대 영상의 매칭과 같이 컬러에 대하여 레이아웃에 기초한 검색(layout based retrieval)을 목표로 하고 있다. 이것은 전체 영상 또는 일정부분의 영상에 사용되어질 수 있다. 추출은 다음과 같이 진행된다. 영상을 8×8 개의 블록으로 나눈다. 각 블록 내의 픽셀에 대하여 하나의 우세한 컬러(dominant color)가 선택된다. 8×8 영상 결과는 DCT 변환을 사용하여 계수의 열로 변환된다. 이것은 할당된 비트의 수에 적합하게 양자화된다. 기본적으로 계수는 Y에 대하여 6, Cb와 Cr에 대하여 각각 3이다[1].

3. MPEG-7 서술자들을 조합하는 방법

조합 방법은 그림 1의 순서도와 같으며 각각의 서술자를 독립적으로 사용하여 영상을 검색한다. 검색 결과로부터 에지 특징을 대상으로 유사도가 가장 떨어지는 $E_{D_{MAX}}$ 와 칼라 특징에 대해 유사도가 가장 떨어지는

C_D_{MAX} 를 구한다.

$$E_D_{MAX} = \text{MAX}\{E_dist_1, \dots, E_dist_b, \dots, E_dist_N\} \quad (1)$$

여기서 E_dist_i 는 i 번째 영상과 쿼리(Query) 영상과의 에지 히스토그램 서술자에 의한 유사도 측정값을 나타낸다. 이때 N 은 총 영상의 개수를 나타낸다.

$$C_D_{MAX} = \text{MAX}\{C_dist_1, \dots, C_dist_b, \dots, C_dist_N\} \quad (2)$$

여기서 C_dist_i 는 i 번째 영상과 쿼리 영상과의 컬러 레이아웃 서술자에 의한 유사도 측정값을 나타낸다. 각 영상의 E_dist , C_dist 값을 구해진 E_D_{MAX} 와 C_D_{MAX} 값으로 식 (3), (4)와 같이 각각 나누어 0~1 사이로 정규화된 유사도 측정치 E_D_i 와 C_D_i 를 구한다.

$$E_D_i = \frac{E_dist_i}{E_D_{MAX}} \quad (i = 1, \dots, N) \quad (3)$$

$$(E_D_i = [0, 1])$$

$$C_D_i = \frac{C_dist_i}{C_D_{MAX}} \quad (i = 1, \dots, N) \quad (4)$$

$$(C_D_i = [0, 1])$$

최종적으로 각 서술자 별로 정규화 된 유사도 E_D_i , C_D_i 의 합을 통하여 각 영상에 대해 새로운 유사도 T_D_i 를 식(5)에서와 같이 구할 수 있다.

$$T_D_i = E_D_i + C_D_i \quad (i = 1, \dots, N) \quad (5)$$

각 영상 별로 구해진 T_D_i 가 낮은 순으로 21개의 결과 영상을 보여준다. 본 실험을 위해 사용한 영상의 개수 N 은 11639이다. 그림 1은 제안한 조합 방법의 순서도이다.

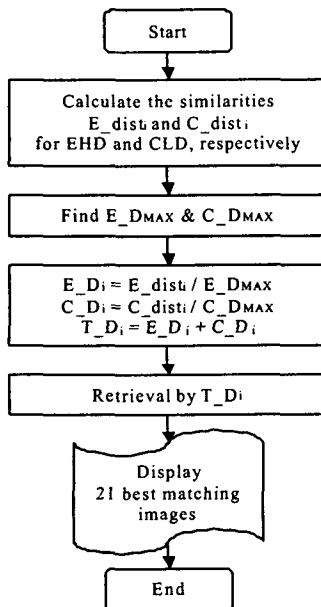


그림 1. MPEG-7 서술자 조합 방법의 순서도

분명히 두개의 서술자를 이용한 영상 검색의 결과가 단지 한개의 영상 서술자를 이용한 검색의 결과보다 더 좋은 결과를 낳는다. 예를 들어, 첫번째 쿼리 영상(코끼리)에 대하여 에지 히스토그램 서술자에 의한 1차 검색 결과 6개의 관련 영상을 찾았으며 컬러 레이아웃 서술자에 의해서는 3개의 관련 영상을 찾았다. 반면에 조합 방법에 의한 1차 검색 결과 7개의 관련 영상을 찾았음을 볼 수 있으며 관련 영상들이 상위 레벨에 위치함을 확인할 수 있다. 두번째 쿼리 영상(농구경기)에 대하여 에지 히스토그램 서술자에 의한 1차 검색 결과 2개의 관련 영상을 찾았으며 컬러 레이아웃 서술자에 의해서는 5개의 관련 영상을 찾았다. 반면에 조합 방법에 의한 1차 검색 결과 8개의 관련 영상을 찾았음을 볼 수 있으며 역시 관련 영상들이 상위 레벨에 위치함을 확인할 수 있다. 그림 2, 그림 3은 두개의 쿼리 영상에 대한 관련 영상의 수로 조합 방법의 성능을 에지 히스토그램 서술자와 컬러 레이아웃 서술자를 독립적으로 사용하는 방법과 비교하여 보여주고 있다.

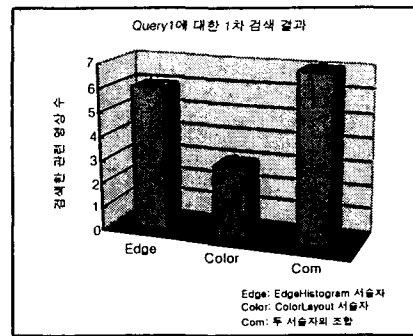


그림2. Query1에 대한 1차 검색 결과

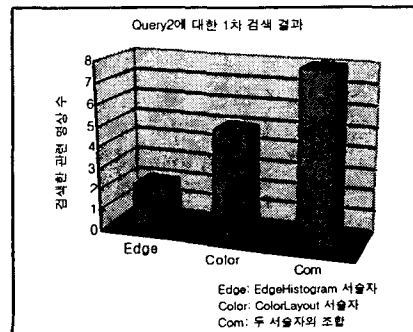


그림3. Query2에 대한 1차 검색 결과

4. MPEG-7 서술자를 선택하는 방법

3장에서와 같이 단순히 두 서술자의 유사도를 통합하므로써 어느 정도 검색성능을 향상시킬 수 있지만 영상에 따라서는 두 서술자 중에 특히 어떤 한 쪽이 다른 한 쪽보다 더 영향을 끼치는 경우가 있다. 이와 같은 상황을 적용적으로 적용하므로써 검색의 성능을 더욱 향상시킬 수 있다. 따라서, MPEG-7 서술자인 에지 히스토그램 서술자와 컬러 레이아웃 서술자를 조합한 방법에 의한 1차 검색 결과인 21개의 영상에 대하여 먼저 사용자

가 관련 영상을 선택한다. 선택된 관련 영상 별로 에지 히스토그램 서술자에 의한 검색 결과(E_NMRR)와 컬러 레이아웃 서술자에 의한 검색 결과(C_NMRR)를 식(6), (7), (8), (9), (10), (11)을 통하여 구한다.

$$E_AVR = \frac{\sum_{k=1}^{E_K} E_Rank(k)}{E_K} \quad (6)$$

$$C_AVR = \frac{\sum_{k=1}^{C_K} C_Rank(k)}{C_K} \quad (7)$$

$$E_MRR = E_AVR - 0.5 - \frac{E_K}{2} \quad (8)$$

$$C_MRR = C_AVR - 0.5 - \frac{C_K}{2} \quad (9)$$

$$E_NMRR = \frac{E_MRR}{2 * E_K + 0.5 - 0.5 * E_K} \quad (10)$$

$$C_NMRR = \frac{C_MRR}{2 * C_K + 0.5 - 0.5 * C_K} \quad (11)$$

여기서 E_K , C_K 는 에지 히스토그램 서술자와 컬러 레이아웃 서술자 별로 쿼리 영상에 대한 유사 영상 (ground truth image)의 개수를 대신하여 사용되고 E_Rank , C_Rank 는 에지와 칼라에 대해 각각 유사영상이 검색된 순위를 나타낸다. E_NMRR , C_NMRR 은 쿼리 영상에 대하여 유사영상 모두 상위순위로 검색하였을 경우 0의 값을 가지게 되고 모두 검색에 실패하였을 경우 1의 값을 가지게 된다. 구하여진 E_NMRR 과 C_NMRR 을 비교하여 그 중 우수한 서술자를 선택한다. 2차 검색 결과는 선택된 서술자에 의한 검색 결과 영상으로 나타내어진다. 2차 검색 이후의 과정은 선택된 서술자에 의한 검색 결과에 대하여 사용자의 만족 여하에 따라 렐러번스 피드백[3][4]이 반복된다. 이때 반복적으로 서술자가 결정될 수 있으며 선택된 서술자에 의하여 새로운 검색 결과를 보여주게 된다.

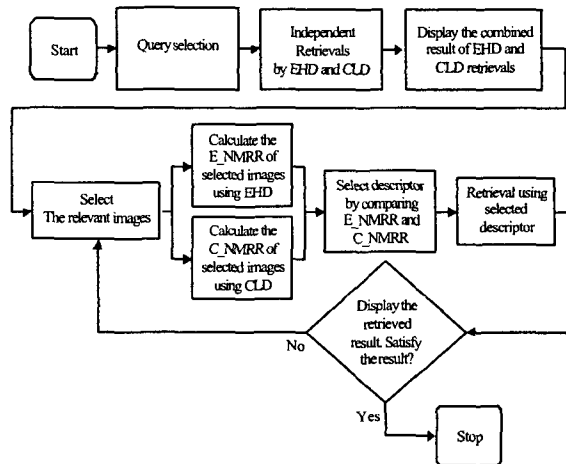


그림 4. MPEG-7 서술자를 선택하는 블록도

그림 4는 에지 히스토그램 서술자와 컬러 레이아웃 서술자 중에 우수한 서술자를 선택하는 방법에 대해 제안

하는 방법의 전체 블록도이다. 제안하는 방법은 반복적으로 서술자를 선택할 수 있기 때문에 서술자의 특성을 잃지 않고 검색을 할 수 있으며 사용자가 선택한 관련 영상들에 의하여 자동적으로 성능이 우수한 서술자를 선택한다. 그러므로, 위의 방법은 에지 히스토그램 서술자와 컬러 레이아웃 서술자 각각의 단점을 보완하여 향상된 검색 결과를 보여준다.

5. 실험 결과

MPEG-7 서술자들을 이용한 관련 영상 검색을 위하여 제안한 방법에 11639개의 자연 영상을 사용하였다. 쿼리 영상은 그림 5, 그림 6, 그림 7의 중앙 아래쪽에 위치하며 좌상에 위치한 영상 일수록 유사도가 높은 순으로 검색된 영상들이다. 그림 5는 에지 히스토그램 서술자만을 사용한 검색 결과이고 그림 6은 컬러 레이아웃 서술자만을 사용한 검색 결과이다. 그림 7은 에지 히스토그램 서술자와 컬러 레이아웃 서술자를 사용하는 제안한 방법에 의한 검색 결과이다. 검색 결과들이 보여주는 바와 같이 에지 히스토그램 서술자만을 사용한 검색 결과는 첫번째 쿼리 영상(코끼리)에 대하여 컬러 레이아웃 서술자 보다 좋은 결과를 보여 준다. 그러나, 두번째 쿼리 영상(농구경기)에 대해서는 그렇지가 못한 것을 볼 수 있다. 대조적으로 컬러 레이아웃 서술자를 사용한 검색은 에지 히스토그램 서술자를 사용한 검색과 반대의 결과를 보인다. 그러나, 제안한 방법에 의한 검색은 그림 7과 같이 두 쿼리 영상에 대하여 우수한 결과를 보여준다. 그림 5, 그림 6, 그림 7의 검은 실선으로 표시된 그림들은 쿼리 영상에 대하여 유사한 영상으로 인정된 것들을 표시한다.

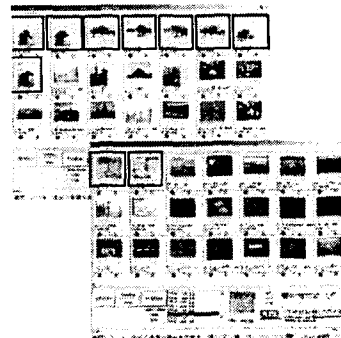


그림 5. EHD에 의한 검색 결과

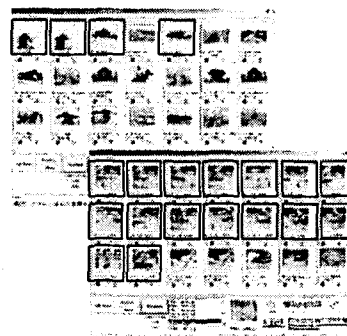


그림 6. CLD에 의한 검색 결과

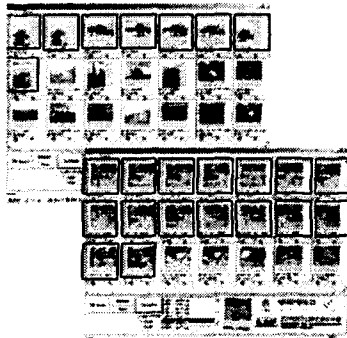


그림 7. 제안한 방법에 의한 검색 결과

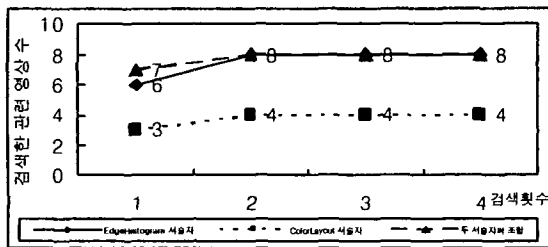


그림 8. Query1에 대한 실험 결과 비교

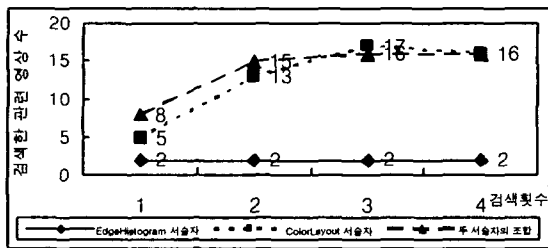


그림 9. Query2에 대한 실험 결과 비교

그림 8, 그림 9는 검색 횟수에 따라 검색된 관련 영상의 수를 나타내어 그림 5, 그림 6, 그림 7에 나타난 결과들을 비교하고 있다. 그림 10은 두개의 쿼리 영상에 대하여 검색 횟수에 따라 검색된 관련 영상의 수를 식 (12), (13), (14)를 통하여 구하여진 정규화 된 값으로 나타내고 있다.

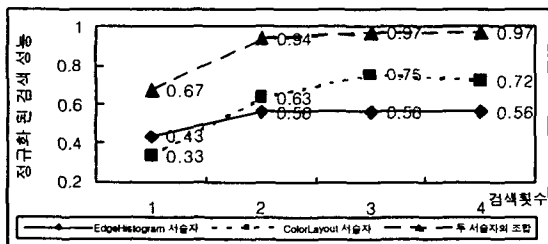


그림 10. 정규화 된 검색 성능 비교

$$E_k = \left(\frac{E_k - Q_1}{Max_Q_1} + \frac{E_k - Q_2}{Max_Q_2} \right) / 2 \quad (k=1, \dots, 4) \quad (12)$$

$$C_k = \left(\frac{C_k - Q_1}{Max_Q_1} + \frac{C_k - Q_2}{Max_Q_2} \right) / 2 \quad (k=1, \dots, 4) \quad (13)$$

$$P_k = \left(\frac{P_k - Q_1}{Max_Q_1} + \frac{P_k - Q_2}{Max_Q_2} \right) / 2 \quad (k=1, \dots, 4) \quad (14)$$

여기서 $E_k - Q_1$, $C_k - Q_1$, $P_k - Q_1$ 은 에지 히스토그램 서술자, 컬러 레이아웃 서술자 그리고 제안하는 방법 별로 쿼리 1에 대하여 검색 횟수에 따라 검색된 관련 영상의 수를 나타내고 $E_k - Q_2$, $C_k - Q_2$, $P_k - Q_2$ 는 쿼리 2에 대하여 나타낸 것이다. k 는 검색 횟수이며 Max_Q_1 , Max_Q_2 는 쿼리 1과 쿼리 2에 대하여 검색된 관련 영상의 최대 수이다. 따라서 E_k , C_k , P_k 는 두개의 쿼리 영상에 대하여 정규화 된 검색 성능 값을 나타낸다.

6. 결론

본 논문에서는 검색 성능을 향상시키기 위하여 MPEG-7의 에지 히스토그램 서술자와 컬러 레이아웃 서술자를 조합하여 검색 성능이 우수한 서술자를 선택하는 방법을 제안했다. 첫번째 단계로 MPEG-7 서술자들을 조합하는 방법으로 각각의 서술자 특징에 대하여 유사도가 가장 떨어지는 값을 구하여 각각의 영상에 대한 거리값을 나누어 줌으로써 정규화를 시켜주고 거리값을 합산하여 관련 영상을 보여주었다. 두번째 단계로는 MPEG-7 서술자를 선택하는 방법으로 서술자 조합에 의한 1차적 검색 결과인 21개의 영상에 대하여 먼저 사용자가 관련 영상을 선택한다. 선택된 관련 영상 별로 에지 히스토그램 서술자에 의한 검색 결과와 컬러 레이아웃 서술자에 의한 검색 결과를 비교하여 그 중 우수한 서술자를 선택하였다. 두번째 단계 이후의 과정은 선택된 서술자에 의한 검색 결과에 대하여 사용자의 만족 여하에 따라 랠러번스 피드백이 반복된다. 결론적으로 제안한 방법은 반복적으로 서술자를 선택할 수 있기 때문에 서술자의 특성을 잃지 않고 검색을 할 수 있으며 사용자가 선택한 관련 영상들에 의하여 자동적으로 성능이 우수한 서술자를 선택한다. 제안한 방법은 향상된 관련 영상 검색을 위하여 서로 보완적이며 검색 성능을 상당히 향상시킴을 알 수 있다.

7. 참고 문헌

- [1] ISO/IEC JTC1/SC29/WG11/W4062, "FCD 15938-3 Multimedia Content Description Interface - Part 3 Visual," Singapore, Mar. 2001
- [2] D.K. Park, Y.S. Jeon, C.S. Won, S.J. Park, "Efficient use of local edge histogram descriptor," Int. Workshop on Standards, Interoperability and Practices, ACM, pp. 52-54, Marina del Rey, CA, Nov. 4, 2000.
- [3] Y. Rui, T. S. Huang, and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," in Proc. IEEE int. Conf. on Image Proc., 1997.
- [4] Zhong Su, Hongjiang Zhang, Shaoping Ma, "Using Bayesian Classifier in Relevant Feedback of Image Retrieval", IEEE, 2000
- [5] Greg Pass, Ramin Zabih, and Justin Miller. Comparing images using color coherence vectors. In Proc. ACM Conf. On Multimedia, 1996
- [6] J. Huang, S.Kumar, W.-J. Zhu, and R. Zabih. Image indexing using color correlogram. In Proc. Of IEEE Conf. On Computer Vision and Pattern Recognition, 1997