

PIM 기반 국부적 Co-occurrence 행렬 및 normalised correlation 를

이용한 효율적 비디오 검색 방법

김규현, 정세윤, 전병태, 이재연, 배영래
한국전자통신연구원

Video image retrieval on the basis of subregional co-occurrence matrix texture features and normalised correlation

Kyuheon Kim, Seyoon Jeong, Byung Tae Chun, Jae Yeon Lee, Younglae Bae
Image Processing Department, Computer • Software Technology Lab.
Electronics and Telecommunications Research Institute
161 Kajong Dong, Yusong Gu, Daejon, S. Korea
kkim@etri.re.kr

ABSTRACT

This paper proposes the simple and efficient image retrieval algorithm using subregional texture features. In order to retrieve images in terms of its contents, it is required to obtain a precise segmentation. However, it is very difficult and takes a long computing time. Therefore, this paper proposes a simple segmentation method, which is to divide an image into high and low entropy regions by using Picture Information Measure (PIM). Also, in order to describe texture characteristics of each region, this paper suggest six different texture features produced on the basis of co-occurrence matrix. For an image retrieval system, a normalised correlation is adopted as a similarity function, which is not dependent on the range of each texture feature values. Finally, this proposed algorithm is applied to a various images and produces competitive results.

1. 서론

컴퓨터의 발전에 따른 멀티미디어 데이터가 급속하게 널리 사용됨에 따라, 이 멀티미디어 데이터를 보다 효과적으로 검색할 방법에 대한 연구가 시급히 요구되어지고 있다. 이러한 이유로, 영상에 포함되어 있는 내용을 영상 특징을 통해 묘사하여 영상 정보를 검색하는 내용기반 영상 검색 기술에 대한 기술이 제안되어지고 또한 많은 관심을 갖게 되었다[1][2].

질감정보는 전통적으로 영상을 나타내는 중요한 정보로 인식되어져 왔다. 질감 정보 자체는 해당

물체가 차지하는 영역의 크기에 따라 같은 재질에 대한 질감 정보도 다른 특징을 갖고 있기에 정확하게 정의하기가 어려우나, 일반적으로는 해당 영역의 특징을 픽셀 단위의 값에 의한 것보다는 여러 개의 픽셀로 구성된 최소 영역을 차지하는 텍셀에 의해 표현되고 있다. 이러한 질감정보의 다양성에 입각하여 영상을 인식할 경우, 한 개의 영상은 몇가지 나아가서는 여러가지의 다른 질감 정보를 갖고 있는 영역들이 모여서 구성된 하나의 모자이크형 조합으로 인식이 가능하다. 이때, 만약, 이러한 다른 질감 정보를 갖는 영역에 대한 질감 정보가 정확히 추출되어진다면, 영상은 그 내용의 질감 정보에 따라 검색이 가능해질 것이다. 하지만, 현실적으로 영상을 질감 정보에 따라 각각의 영역으로 구분하는 정확한 세그멘테이션 작업은 많은 계산 시간을 요하며 또한 그 정확도에서도 많은 문제점을 갖고 있는 현실이다. 그러나, 현실적으로 인간은 한 영상이 갖고 있는 모든 다른 영역에 대한 정확한 인지보다는 그 영상이 포함하고 있는 중요한 내용물에 대한 인지로 그 영상을 타 영상과 구분하며 또한 유사 여부를 판단하고 있다. 즉, 한 영상의 내부에 대한 자세하고 정확한 인식이 그 영상을 유사여부를 판단하는 가장 필요한 요소이나, 중요한 몇가

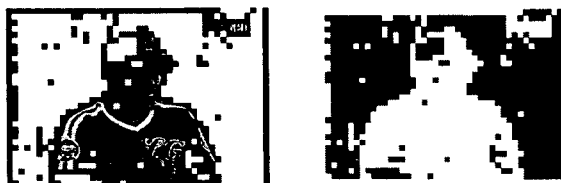
지 내용물 정보로도 충분히 만족할 만한 유사 검색 검색물을 얻을 수 있다는 것이다. 본 논문은 한 영상의 내부에 대한 자세한 정보를 얻기 위한 복잡한 세그멘테이션 작업대신, Picture Information Measurement (PIM)을 이용하여 해당 영상을 엔트로피가 높고 낮은 영역으로 구분하는 두개의 특징영역으로 구분하였다 [3]. 이렇게 나누어진 영역에 대한 각각의 질감 정보를 얻기 위하여 co-occurrence 행렬을 이용한 6가지의 질감 정보 특징량의 사용을 제안하였다 [4]. 이렇게 얻어진 특징량들은 영상의 유사도에 기반한 검색 작업을 위하여 사용되어진다. 이때, 각각의 특징 값들은 각각이 나타내는 특징들이 다르기에 그 값의 분포에서도 서로 다른 영역을 보여준다. 이에, 보다 정확한 검색을 얻기 위하여 본 논문은 normalised correlation 을 사용하였다 [5].

2. Picture Information Measurement

본 논문에서는 영상의 영역을 분리하는데 엔트로피를 이용하고 있다. Chang 은 Shannon 엔트로피에 이론적 기반을 두고 영상의 공간 정보까지 반영하는 PIM 이라는 영상에 적합한 엔트로피를 제안하였다. PIM 의 정의는 식 (1)과 같다 [3].

$$PIM = \sum_{l=0}^{L-1} (h(i) - Max_l h(i)) \quad (1)$$

여기서 $h(i)$ 는 각 해당 블록(또는 영상)의 히스토그램을 의미한다. PIM 값은 히스토그램 중 가장 큰 도수의 히스토그램 값과 해당 블록의 전체 픽셀 수와의 차이로 계산된다.



(a) High Entropy 영역

(b) Low Entropy 영역

그림1. PIM 엔트로피를 이용한 영역 분리

그림 1은 이러한 PIM의 엔트로피를 이용하여 영상을 엔트로피가 높고 낮은 영역으로 구분한 예를 보여 주고 있다. 그림 1-(a)는 엔트로피가 높은 영역 영상이고 그림 1-(b)는 엔트로피가 낮은 영역 영상이다. 그림에서 알 수 있듯이, 엔트로피가 높은 부분에 영상 내부의 오브젝트에 대한 정보가, 또한 엔트로피가 낮은 영역에는 배경에 대한 정보가 포함되어 있음을 알 수 있다. 이처럼, 영상 내부의 오브젝트에 대한 정보를 얻기 위하여 사용되는 복잡하고 많은 계산을 요구하는 세그멘테이션 대신, PIM을 이용하여 보다 손쉽게 영상 내부의 오브젝트에 대한 영역을 분리함으로써, 영상의 내용의 공간 정보를 추출할 수 있으며, 또한 영상 내용을 무시한 전체적인 영상 특징 정보 보다 공간정보에 입각한 보다 충실한 내용기반 영상 검색을 이루어 낼 수 있다.

3. 질감 정보

본 논문에서는 PIM에 의해 분리된 영역을 하나의 질감을 갖고 있는 영역으로 구성되어 있다고 가정하고 있다. 이러한 가정하에서, 단일한 질감의 특징을 갖고 있는 영역에 대한 질감 정보를 추출하기 위하여 Co-occurrence 행렬을 이용한 질감 특징 정보를 사용하였다.

Co-occurrence 행렬은 그레이 값 i 와 j 를 갖는 2개의 픽셀이 거리 r 에서 어떤 방향 θ 로 얼마나 자주 나타나는 것 인가를 나타내는 행렬로 본 논문에서는 다음과 같이 표기 하였다.

$$P(i, j, r, 0^\circ) = \#\{(k, l), (m, n) \in \Omega, |k - m| = 0, |l - n| = r, Y(k, l) = i, Y(m, n) = j\} \quad (2)$$

$$P(i, j, r, 45^\circ) = \#\{(k, l), (m, n) \in \Omega, ((k - m) = r, (l - n) = -r) \text{ or } ((k - m) = -r, (l - n) = r), Y(k, l) = i, Y(m, n) = j\} \quad (3)$$

$$P(i, j, r, 90^\circ) = \#\{(k, l), (m, n) \in \Omega, |k - m| = r, |l - n| = 0, Y(k, l) = i, Y(m, n) = j\} \quad (4)$$

$$P(i, j, r, 135^\circ) = \#\{(k, l), (m, n) \in \Omega, ((k - m) = r, (l - n) = r) \text{ or } ((k - m) = -r, (l - n) = -r), Y(k, l) = i, Y(m, n) = j\} \quad (5)$$

윗 식에서 #은 조건을 만족하는 픽셀들의 개수를 나타낸다. 상기에서 설명한 바와 같이 4개의 서로 다른 방향 성분을 고려하기에, 비교되어지는 픽셀의 총수는 달라진다. 이러한 개수의 차이에서 오는 오류를 시정하기 위하여, 다음과 같은 정규화 상수를 각각의 방향 성분 co-occurrence 행렬에 사용하였다.

$$\kappa_H = \angle IV_y (IV_x - 1) \quad (6)$$

$$\kappa_V = \angle IV_x (IV_y - 1) \quad (7)$$

$$\kappa_D = \angle (IV_x - 1)(IV_y - 1) \quad (8)$$

상기에서 정의된 각 방향의 Co-occurrence 행렬을 이용하여 여러가지의 많은 질감 정보 특징들이 개발 되었다. 많은 특징값들의 사용은 보다 많은 계산 시간을 요구하기에 본 논문에는 계산 시간과 정확도의 딜레마를 해결하기 위해 뚜렷한 각각의 특징들을 나타내는 다음과 같은 6개의 특징값을 사용하였다. [4][6][7]

a) Maximum probability:

$$f_{MP} = \underset{i,j}{MAX} \left(\frac{P(i,j,r,\theta)}{R} \right) \quad (9)$$

b) Angular Second Moment: uniformity

$$f_{CON} = \sum_{n=0}^{N_x-1} n^2 \left\{ \sum_i \sum_j \left(\frac{P(i,j,r,\theta)}{R} \right) \right\}_{where |i-j|=n} \quad (10)$$

c) Contrast:

(11)

d) Correlation

$$f_{COR} = \frac{\sum_i \sum_j (ij) \frac{P(i,j,r,\theta)}{R} - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (12)$$

e) Sum of squares: variance

$$f_{IDM} = \sum_i \sum_j \left(\frac{1}{1+(i-j)^2} \left(\frac{P(i,j,r,\theta)}{R} \right) \right) \quad (14)$$

$$f_{VAR} = \sum_i \sum_j (i-\mu)^2 \left(\frac{P(i,j,r,\theta)}{R} \right) \quad (13)$$

f) Inverse difference moment:

4. Normalised correlation 및 검색 알고리즘

질감 특징량을 구하기 위한 알고리즘은 다음과 같다.

1. 입력영상을 8x8의 윈도우로 나눔
2. 영상 내의 해당 윈도우의 PIM 및 평균, 분산 값을 구함
3. 해당 윈도우의 분산값이 해당 PIM의 분산값보다 클 경우, 해당 윈도우 영역을 높은 엔트로피 영역, 그리고 낮을 경우 낮은 엔트로피 영역으로 나눔.
4. 분리된 2개의 영역에 대한 Co-occurrence 행렬 기반 6개의 질감 특징 정보 구함.

상기에서 설명된 특징량을 서로 비교하여 해당 영상의 유사도를 측정한다. 일반적으로 유사도 측정에 사용되는 유클리디언 거리는 특징량의 값의 분포가 상이하게 다를 때 정확한 유사도 측정이 어려우기에, 본 논문은 아래에 정의된 normalised correlation을 사용하였다. 하단에 정의된 normalised correlation은 특징량 각각의 값에 관계없이 독립된 유사도 값을 나타내 준다 [5].

$$NC(Q,I) = \sum_{j=1}^n \frac{\langle F_{Q_j}, F_{I_j} \rangle}{\sqrt{\langle F_{Q_j}, F_{Q_j} \rangle \langle F_{I_j}, F_{I_j} \rangle}} \quad (17)$$

$$\langle F_{Q_j}, F_{I_j} \rangle = F_{Q_j}^T F_{I_j} \quad (18)$$

5. 실험 결과

본 논문에서는 MPEG7에서 사용되고 있는 컬러 영상 데이터 중 500장을 선택하여 실험 데이터 베이스로 사용하였다. 그림 2-4는 실험 검색 결과이다. 그림 2에서는 부드러운 질감 정보 질의 영상이고, 그림 3은 커진 질감 정보가 대부분을 차지하는 질의의 검색이다. 제안한 알고리즘이 영상 내부에 2개의 커다란 오브젝트가 있다는 것을 가정하였어도, 그림 2-3의 질의 영상 처럼 영상 내부에 커다란 오브젝트에 해당하는 부분이 없는 영상에 대해서도 성공적인 검색 결과를 보여주고 있다.

또한 그림 4는 2개의 커다란 오브젝트로 구성된 질의 영상의 검색 결과이다. 본 실험 결과를 통해서, 제안한 기술이 영상을 2개의 커다란 영역으로 나누지만, 질의 영상 내부의 오브젝트 수에 상관없이 좋은 결과를 나타냄을 알 수 있다.

6. 결론

본 논문은 영상 내부의 질감 정보 특징을 이용하여 영상을 보다 손쉽게 검색할 수 있는 알고리즘을 제안하였다. 제안된 알고리즘은 PIM을 사용하여 영상을 2개의 엔트로피 영역으로 나누고, 각각의 영역에서 Co-occurrence 행렬을 이용하여 질감 특징 정보를 추출하였다. 또한, 유클리디언 거리가 각 특징량 값들에 좌우되는 한계를 극복하기 위하여, 본 논문에서는 normalised correlation을 사용하였다.

그러나, 영상을 2개의 엔트로피 영역으로 나눌경우, 원 영상의 자세한 질감 정보를 잃을 수 있는 위험성도 있다. 이러한 문제점은 영상 검색이 갖고 있는 시간의 효율성과 검색의 정확도라는 딜레마의 문제이기도 하다. 이러한 딜레마를 해결하기 위하여, 제안된 검색 기술이 2개 이상의 영역으로 확장되어 상세한 질감 정보를 구할 수 있도록 하는 연구가 요구되어진다.

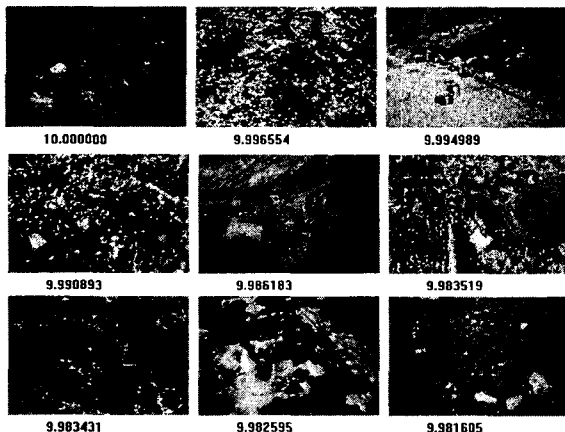


그림 2. 검색 결과 (부드러운 질감의 질의 영상)

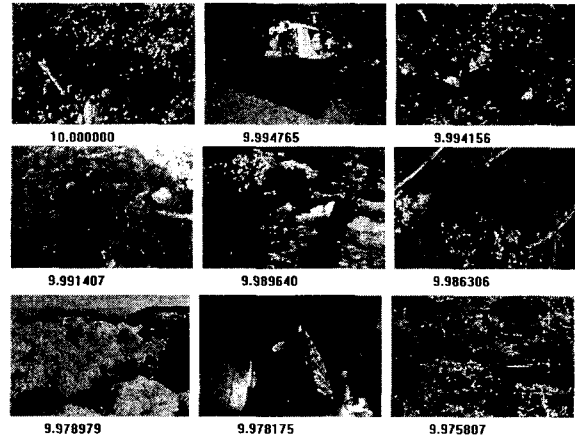


그림 3. 검색 결과 (키친 질감의 질의 영상)

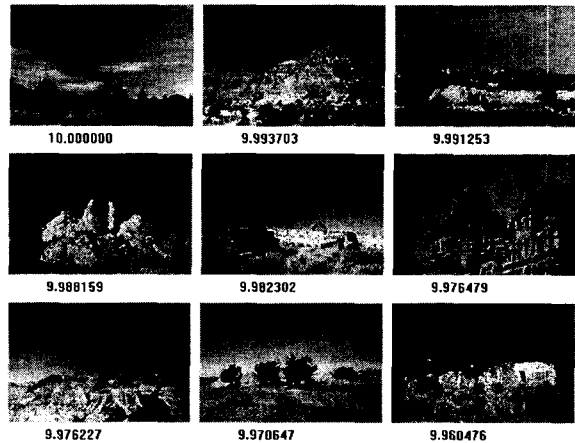


그림 4. 검색 결과

참고문헌

- [1] M. Stricker and A. Dimai, "Color Indexing with Weak Spatial Constraints," *Proc. of SPIE: Storage and Retrieval for Image and Video Databases IV*, Vol. 2666, pp. 29-40, San Jose, Feb. 1996
- [2] B.S. Manjunath and W.Y. Ma, "Texture features for browsing and retrieval of image data", *IEEE Trans. On PAMI*, pp. 837-842, Aug. 1996
- [3] J. R. Smith, Shin-Fu Chang, "Tools and techniques for Image Retrieval", *Proc. of SPIE: Storage and Retrieval for Image and Video Databases IV*, Vol. 2666, pp. 426-437, San Jose, Feb. 1996.
- [4] R.M. Haralick, K. Shanmugsm and I. Dinstein, "Texture features for image classification," *IEEE. Trans. On Systems, Man and Cybernetics*, vol. 3, pp. 610-621, 1973
- [5] A.K. Jain, *Fundamentals of digital image processing*, London, UK, Prentice-Hall, Inc., 1989
- [6] L.M. Linnett and A.J. Richardson, "Texture segmentation using directional operators," *Proc. ICASSP, IEEE Int. Conf. Acous, Speech, Signal Processing*, vol. 4, pp. 2309-2312, 1990
- [7] R.M. Haralick, "Statistical and structural approaches to texture," *Proc. of the IEEE*, vol. 67, pp. 786-804, 1979