

PCA를 이용한 데이터베이스 검색에 있어서의 최적 MPEG-7 디스크립터 조합에 관한 연구

김현민, 최윤식
연세대학교 전기전자공학과
e-mail : beautifulife@hanmail.com

A study of optimal MPEG-7 descriptor composite in database searching
using PCA

Hyun-Min Kim, Yoon-sik Choe
Department of Electrical and Electronic Engineering
Yonsei University

Abstract

When we search database with a query image, the retrieval efficiency will vary from each kind of descriptor. Even the best representative descriptor, it results a few useless images that don't match with query image. This type of error can be reduced by adopting another descriptor which extracts features in different way.

At present, the choice of descriptors is base on intuitive and experimental method. By theoretic accessing to the problem of descriptor choice, we can solve the given problem in the objective and rational way.

In this study, we intend to make a composite of descriptors that can reduce retrieval error by adopting principal component analysis.

I. 서론

이미지 기반 데이터 검색을 할 때 검색 정확도는 어떠한 알고리즘을 사용하느냐에 따라 달라지게 된다. MPEG-7내에 정의 되어있는 비주얼 디스크립터들은 이러한 이미지 기반의 검색을 위해 응용되어지며 각 디스크립터 들은 서로 상이한 알고리즘을 사용하여 데이터 검색을 한다. 이 MPEG-7 비주얼 디스크립터를 사용하여 이미지 기반 데이터 검색을 하는 경우 한 가지 디스크립터만을 사용하는 경우보다는 두 가지 이상의 디스크립터들을 동시에 사용하여 검색을 하는 것은 검색 효과를 더 높일 수 있다. 본 연구에서는 특정 부류의 쿼리영상(query image)에 있어서 어떠한 디스크립터들을 사용하여 디스크립터 조합을

생성 하는 것이 최적의 성능을 가질 수 있는 지에 대해 연구 하였고, 그 방법으로서 주성분 분석(principal component analysis)방법을 사용하였다.

II. 이론적 배경

2.1 MPEG-7 비주얼 디스크립터

MPEG-7 비주얼 스탠다드(visual standard)[1]의 주된 목표는 전송되거나 저장된 이미지의 표준화된 디스크립터를 제공하는 것이다. 디스크립터는 기술하는 방법에 따라 크게 두 가지 부류로 나눌 수 있는데, 하나는 저수준(low-level) 디스크립터이고 다른 하나는 고 수준(high-level) 디스크립터

터이다. 이때, 저 수준 디스크립터는 영상의 색이나 질감, 모양 등에 대한 정보로서 완전히 자동으로 추출될 수 있다. 반면, 고 수준 디스크립터는 추상 단계기술에 관한 내용으로, 의미론적 정보를 담고 있다.

MPEG-7 스탠다드의 비주얼 파트는 여러 디스크립터들을 정의하고 있다. 이들 모두가 비주얼 미디어(visual media)에서 특징 추출을 한다는 의미의 디스크립터는 아니고 몇몇 디스크립터는 디스크립터 집단의 구조를 나타내기도 한다. 가장 기본적인 디스크립터들은 컬러레이아웃(color layout), 컬러 스트럭처(color structure), 도미넌트 컬러(dominant color), 스케일러블 컬러(scalable color), 에지 히스토그램(edge histogram), 호모지니어스 텍스처(homogeneous texture), 텍스처 브라우징(texture browsing), 리즌 셰이프(region shape) 등이 있다.

2.1. 주성분 분석(principal component analysis)

PCA(Principal Component Analysis)[2]의 목적은 오리지널 변수(original valuables)들의 선형조합으로 이루어진 새로운 변수들을 생성함으로써, 데이터 집단내의 변수들의 수를 줄이는데 있다. PCA 는 기하학 적으로 오리지널 좌표축을 새로운 직교좌표축(orthogonal axis)으로 축 변환(axis transformation)을 시키는 것이고, 한 개의 전체 데이터를 대표할 수 있는 최적 라인을 찾는 것이 알고리즘 내용의 중심 된 포인트이다 . 전체 데이터를 대표하는 최적라인은 샘플 포인트로부터 라인까지의 수직 거리의 제곱이 최소가 되게 하는 제한 조건 하에서 구할 수 있으며, 이렇게 구해진 변환된 좌표축에서의 가장 큰 몇 개의 성분은 오리지널 데이터의 variation의 대부분을 설명할 수 있다[8][9].

III. 디스크립터 선정에 있어서의 PCA의 적용

본 논문에서는 특정 패턴의 쿼리영상을 검색 하는데 있어서 어떠한 디스크립터의 조합이 가장

유용한가에 초점을 맞추어 연구하였다. 이러한 맥락에서 주성분 분석을 이용하였다.

3.1 PCA 적용 방법

본 논문에서는 MPEG-7 저수준 디스크립터중 비주얼 디스크립터인 컬러레이아웃 디스크립터(colorlayout descriptor,CLD), 호모지니어스 텍스처 디스크립터(homogeneous texture descriptor), 에지 히스토그램 디스크립터(edge histogram descriptor)를 사용하여 특정 쿼리영상에 대한 최적 디스크립터를 찾고자 한다

제안한 방법은 크게 네 단계로 나뉜다. 먼저 디스크립터 단위로 유사한 영상들의 공간적 분포를 고려해 주 방향 성분을 각각 구한다. 다음 단계로 구한 주 방향 성분을 두 번째 주 방향 성분 분석에 알맞은 형태로 수정한 후 두 번째 주방향 분석을 수행한다. 마지막으로 구한

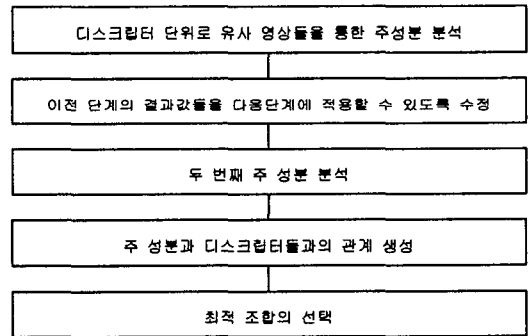


그림 1. 제안한 알고리즘의 순서도

주 성분 방향을 디스크립터들과 매칭(matching) 시킴으로써 최적의 디스크립터 조합을 구한다. 그림 1에는 이상과 같은 과정을 순서도로 나타내었다.

본 논문의 알고리즘의 첫 번째 단계인 디스크립터 단위로 유사 영상들을 통한 주성분 분석을 위해 유사한 영상 n 장을 사용하여 디스크립터 하나를 사용했을 때 추출된 계수 값들을 이용해 입력 행렬(X)을 구성한다. 유사한 영상 한 장에 있어

서 추출된 계수 값들을 세로 방향으로 나열 하였다. 계수 값들의 종류가 P 개 였으면 적용된 영상의 장 수가 n 장이므로 행렬의 크기는 $P \times n$ 이 된다. 이 행렬의 원소들을 standardize 시킨후 스캐터 행렬 S 를 식(1)과 같이 구한다.

$$S = XX^T \quad (1)$$

이 행렬 S 를 대각화 함으로서 고유값과 고유벡터를 구할 수 있다.

이러한 대각화 과정을 네 개의 디스크립터 D_A, D_B, D_C, D_D 마다 동일하게 적용하여 각각의 고유 값들 중 최대 고유 값과 이에 해당하는 고유 벡터들을 구함으로써 다음 단계의 과정에 이용한다. 이렇게 각 디스크립터별로 따로 분석을 하는 이유는 모든 디스크립터들을 한꺼번에 분석함으로써 생기는 블록 매트릭스간의 계산 복잡성에 있다.

이번에는 앞서 구한 네 쌍의 고유 값과 고유 벡터를 강제적으로 대각화 위치에 할당한다. 이들 각각의 고유 값에 해당하는 고유벡터들은 사실 서로 직교하지 않기 때문에 본 논문에서는 이러한 대각화 과정을 유사 대각화라고 이름 지었다. 유사 대각화 과정은 식 (5-15)과 같이 표현된다.

$$S \cong TYT^{-1} \quad (2)$$

이 유사 대각화 된 식을 실제적으로 서로 직교하는 고유벡터들로 이루어진 행렬과 대각 행렬의 곱으로 표현 함으로써 최종 적으로 대각화가 가능하다. 즉, 식 (3)과 같은 식이 성립된다. 이 식에서 P 는 고유 벡터들로 이루어진 행렬이고 D 는 최종 적으로 대각화 된 행렬이다.

$$S \cong TYT^{-1} = PDP^{-1} \quad (3)$$

마지막으로 주성분 방향과 디스크립터간의 관계를 알아보기위해 D 의 최대 고유 값에 해당하는 고유 벡터와 각 디스크립터간의 유사도를 조사하

여 가장 유사도가 큰 두 개의 디스크립터를 해당 쿼리영상에 적합한 디스크립터 조합의 성분으로 선정하였다.

IV. 실험 결과 및 분석

본 연구에서는 텍스처영상과 이진영상의 경우로 나누어 특정쿼리 영상에 있어서의 최적 디스크립터 조합을 선정하는 실험을 수행하였다.

실험에 사용한 텍스처 쿼리영상과 이진 쿼리영상은 그림 2,3과 같다.

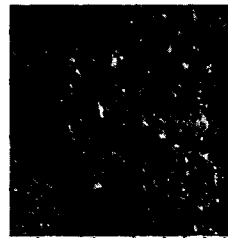


그림 2. 텍스처 쿼리영상



그림 3. 이진 쿼리영상

두 경우 모두, 데이터 마이닝(data mining)을 위해 쿼리영상과 유사한 영상을 10장 씩을 사용하였다.

유사 대각화를 이용한 최종 대각화를 통해구한 고유 값들은 각각 표 1,2와 같다.

고유 값1	고유 값 2	고유 값 3	고유 값 4
6.7709	0.7354	3.3459	0.0004

표 1 텍스처영상의 경우 계산된 고유 값

고유 값1	고유 값 2	고유 값 3	고유 값 4
8.0356	0.005	0.8387	2.8909

표 2 이진영상의 경우 계산된 고유 값

표1,표2에서 각각 최대 고유 값 6.7709, 8.0356을 선택하고 각 고유 값에 해당하는 고유 벡터와 디스크립터들과의 상관도(correlation)을 구한다. 그 결과를 표3,표 4에 나타내었다.

순위	디스크립터 조합	상관도의 합
1	CLD+HTD	1.0122
2	CLD+EHD	0.9070
3	CLD+DCD	0.8503
4	HTD+EHD	0.5244

5	HTD+DCD	0.4677
6	EHD+DCD	0.3625

표 3. 텍스트영상의 경우
디스크립터조합 효용순위

순위	디스크립터 조합	상관도의 합
1	HTD+EHD	1.5446
2	HTD+CLD	1.2341
3	HTD+DCD	1.1808
4	EHD+CLD	1.0185
5	EHD+DCD	0.9652
6	CLD+DCD	0.6547

표 3. 이진영상의 경우
디스크립터조합 효용순위

표4,표5에는 해당 쿼리에대하여 모든조합의 경우에 있어서 실제 검색결과를 토대로 조합 효과도를 따져서 순위별로 기록하였다.

순위	디스크립터
1	CLD+HTD
1	CLD+DCD
1	HTD+EHD
4	EHD+DCD
5	HTD+DCD
6	CLD+EHD

표 4

순위	디스크립터
1	CLD+HTD
1	CLD+EHD
1	HTD+EHD
4	HTD+DCD
5	EHD+DCD
6	CLD+DCD

표 5

표4 텍스트영상의 경우 실험적 방법을통한 디스크립터조합 효과도 순위

표5 이진영상의 경우 실험적 방법을통한 디스크립터조합 효과도 순위

표2,3과 표4,5를 비교해 보면 알 수 있듯이 제안한 방법을 통한 최적 디스크립터의 선정은 실험적 방법을 통한 조합의 선정과 유사한 결과를 보이고있다.

V. 결론

본 논문에서는 PCA를 이용한 MPEG-7 최적 디스크립터의 조합 선정에 관한 연구를 수행하였다. 실질적 방법을 통한 디스크립터 조합의 선정 결과와 제안한 방법을 통한 조합 선정 결과를 비교해보면, 두 가지 방법의 결과가 정확히 일치하지는 않고, 전체적인 추이가 유사한 모습

을 보인다. 이에 대한 원인으로서는 실험에 있어서의 오차가 존재한다는것과, 제안한방법에서는 간략화(approximation)을 수행했기 때문에 정확한 값과는 차이가 있는 것으로 파악된다. 하지만 실험적 방법과 제안한 방법으로 얻어진 디스크립터들의 효용도 순위를 비교해 본다면 하나 하나의 순위는 맞지 않지만 전체적인 순위 흐름은 유사한 것으로 파악되었다. 즉 제안한 방법에 있어서 상위 1,2인가, 3,4위인가는 실험적 방법과 차이점을 보였지만 특정 조합이 상위 순위에 있는가 하위 순위에 있는가에 있어서는 유사한 결과가 나왔다.

이러한 결과를 토대로 영상 검색에 있어서 어떠한 조합을 사용하는 것은 효율 적이지 못하며 어떠한 조합을 사용하는 것은 효율 적이다 라는 판단을 내릴 수 있으므로 본 연구에서 제안 한 방법을 통한 디스크립터 조합의 선정은 의미를 갖는다.

참고문헌

- [1] Information Technology Multimedia Content Description Interface-Part 8: Extraction and use of MPEG7 descriptions, July, 2002
- [2] Chao Huang; Tao Chen; Chang, E.; Speaker selection training for large vocabulary continuous speech recognition, Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASSP '02). IEEE International Conference on , Volume: 1 , 13-17 May 2002 Page(s): I-609 -I-612 vol.1