

## 영상 검색의 속도 향상을 위한 차원 축소를 최적화

이 세 한, 이 주 호, 조 정 원, \*최 병 옥  
한양대학교 전자통신전파공학과, \*한양대학교 정보통신학부  
전화 : 02-2290-0363 / 핸드폰 : 019-390-2312

### Optimization of Condensation Ratio for Fast Image Retrieval

Se-Han Lee, Chu-Ho Yi, Jung-Won Cho, Byung-Uk Choi  
Department of Electrical and Computer Engineering, Hanyang University  
\*Division of Information and Communications, Hanyang University  
E-mail : sehani@ihanyang.ac.kr

#### Abstract

This paper suggests the condensed two-stage retrieval method for fast image retrieval in the content-based image retrieval system, and proves the validity of the performance. The condensed two-stage retrieval method reduces the overall response time remarkably while it maintains relevance with the conventional exhaustive search method. It is explained by properties of the Cauchy-Schwartz inequality. In experimental result, it turns out that there is an optimal value of condensation ratio which minimizes the overall response time. We analyze the optimal condensation ratio by modeling a similarity computation time mathematically.

#### I. 서론

내용기반 영상 검색(content-based image retrieval)은 검색의 대상이 되는 영상 정보들을 색상, 형태, 질감 등의 시각에 의한 특징들로 기술하고, 이 특징 벡터들 간의 유사도를 비교함으로써 원하는 영상을 찾아내는 것으로 정의된다. 일반적으로 검색시스템의 성능은 적합성과 신속성으로 평가할 수 있는데, 내용기반

검색에서는 특징 벡터들 간의 유사도를 기반으로 검색이 이루어지므로 특징 벡터의 종류와 차원 수가 시스템의 속도와 밀접한 관계를 가지고 있다. 적합성을 높이기 위한 연구들은 질의에 부적합한 결과는 가능한 검색하지 않아야 한다는 전제를 바탕으로 특징 기술자들을 혼합하거나 차원수를 높이는 데에 집중되어 왔다. 그러나 특징 벡터의 차원 수를 높이면, 유사도 비교시간이 길어져 검색시간이 증가하게 된다. 이러한 문제점을 해결하기 위하여 특징 벡터의 차원 수를 줄이는 부분에 관한 많은 연구가 진행되어 왔다. 대표적인 방법으로는 부집합 선택 (variable-subset selection) [1], 다중차원 스케일링 (MDS, multidimensional scaling) [2][3], 기하 해싱 (geometric hashing) [4][5]을 들 수가 있다. 부집합 선택은 원 특징 벡터의 기저 벡터들을 선형 변환에 의하여 사상시켜서 변화량 (variance)이 적은 기저를 대상으로 차원을 줄인다. KLT(Karhunen-Loève transform), SVD(singular value decomposition), PCA(principle component analysis) 등의 방법이 이에 속하며, 데이터에 종속적이고 계산량이 많다는 단점을 가지고 있다. 다중차원 스케일링은 원 특징 벡터를  $k$ -feature extraction function에 의해 비선형적인 방법으로  $n$ 차원에서  $k$ 차원으로 줄인다. SOM(self-organized map)을 이용해 클러스터링을 한 후에 각 클러스터에 대해 선형 PCA를 적용시키는 비선형 PCA나 FastMap이 여기에 해당한다. 마지막으로 기하 해싱은 영상의 정보를 기하학적인 관계에 의한 특징 점들로 표현을 하고, 이것을 동일한 크기의 해시

테이블에 저장을 한다. 차원 축소는 해싱 함수에 의해 이루어지는데 함수 설계가 복잡하고, 원 특징 벡터가 저차원일 경우에만 사용할 수 있다. 특징 벡터의 차원을 축소시키는 이러한 방법들은 검색 순위에 영향을 미치거나 탐색구조를 변화시키는 문제점이 있고, 전체 차원의 원 특징 벡터를 대상으로 검색했을 경우의 적합성을 그대로 보장할 수가 없다는 단점이 있다.

본 논문에서는 시스템의 적합성을 그대로 유지하고 결과 순위에 영향을 미치지 않으면서 검색 속도를 향상시킬 수 있는 축소 2단계 검색 방법 (condensed two-stage search method)에 대해 소개한다. 그리고 실험에 의해 확인된 특징 벡터의 최적 차원 축소율 (condensation ratio)을 통계적 추정에 의한 방법으로 모델링하여 전체 응답시간을 최소화시킬 수 있는 최적 축소율을 분석한다.

## II. Cauchy-Schwartz 부등식

원 특징 벡터의 차원 수를  $m$ , 전체 데이터베이스 내 영상 수를  $n$  이라고 하면, 데이터베이스 내의 영상에 대한 특징 벡터와 질의 영상에 대한 특징 벡터를 (1) 과 같이 나타내고, 임의의  $m'$  차원으로 축소된 특징 벡터를 (2)와 같이 표시할 수 있다.

$$D^{(m)} = \{d_1, d_2, \dots, d_m\}, Q = \{q_1, q_2, \dots, q_m\} \quad (1)$$

$$D^{(m')} = \{d'_1, d'_2, \dots, d'_{m'}\}, Q' = \{q'_1, q'_2, \dots, q'_{m'}\} \quad (2)$$

원 특징 벡터를 2배씩 축소시킬 경우, 축소된 특징 벡터 성분은 식 (3)에 의해 그 이전 단계의 특징 벡터보다 항상 크거나 같음을 알 수 있다.

$$d'_i = \sqrt{(d_{2i-1})^2 + (d_{2i})^2} \geq \max(d_{2i-1}, d_{2i}), (1 \leq i \leq m') \quad (3)$$

본 논문에서는 특징 벡터간의 유사도를 내적에 의한  $\cos\theta$  값으로 정의한다. 이러한 방법은 L2 norm을 사용하면서도, 유사도 값을 0에서 1로 정규화 한다는 장점이 있다. 차원 축소된 특징 벡터에 의한 유사도 역시 원 특징 벡터에 의한 유사도 보다 항상 크거나 같은 성질을 나타낸다. 이것은 (4)의 Cauchy-Schwartz 부등식에 의해 설명된다.

$$\sqrt{\{\sum d_i^2\} \cdot \{\sum q_i^2\}} \geq \sum (d_i \cdot q_i) \quad (4)$$

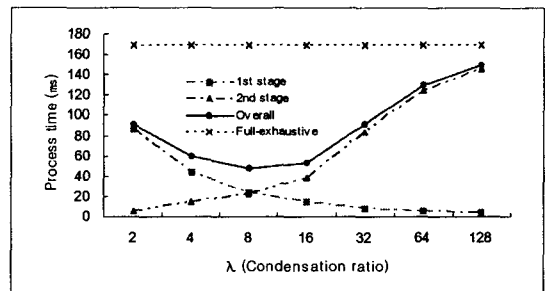
여기서 (4)의 우변은 원 특징 벡터로 계산한 유사도를 나타내고 좌변은 차원 축소된 특징 벡터로 계산한 유사도를 나타내게 되며, 차원 축소된 특징 벡터로 계산한 유사도 역시 차원 축소율이 보다 낮은 단계의 유사도 보다 항상 크거나 같다는 것을 알 수 있다.

## III. 축소 2단계 검색 방법

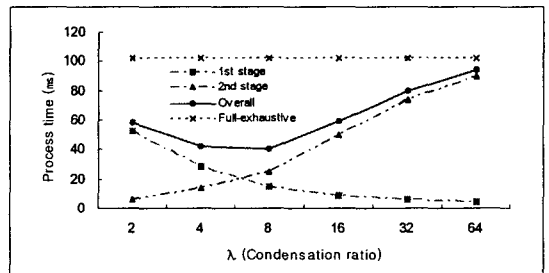
본 장에서는 위에서 살펴본 Cauchy-Schwartz 부등식의 성질을 이용한 축소 2단계 검색 방법에 대해 설명한다. 축소 2단계 검색 방법에서는 Cauchy-Schwartz 부등식의 특성에 의해 특징 벡터의 차원을 축소하기 때문에 전체 검색과 비교해 볼 때, 적합성을 그대로 유지하고, 결과 순위에 영향을 미치지 않으면서도 응답 시간을 현저히 줄일 수 있다. 특징 벡터의 차원 축소율을 (5)와 같이 정의할 때,  $\lambda$ 가 커질수록 1차 검색 시간이 줄어드는 반면, 1차 검색 결과인 후보 영상의 개수가 많아져 2차 검색 시간은 증가하게 된다.

$$\lambda = \frac{D_o \text{ (원 특징 벡터의 차원수)}}{D_c \text{ (축소 특징 벡터의 차원수)}} \quad (5)$$

반대로  $\lambda$ 값이 작아질수록 1차 검색 시간이 증가하여 후보 영상의 개수는 적어지는 반면, 2차 검색 시간은 줄어든다. 그림 1은 유사도 기준 0.8에서 원 특징 벡터가 256, 128 차원일 때, 각각  $\lambda$ 값의 변화에 따른 1차 검색 시간과 2차 검색 시간, 그리고 전체 응답시간의 변화를 보여준다. 그림 1의 결과로부터, 축소 2단계 검색 방법에서 전체 응답 시간을 최소화 시키는 최적의  $\lambda$ 값이 존재함을 확인할 수 있다.



(a) 256 차원



(b) 128 차원

그림 1.  $\lambda$ 값의 변화에 따른 응답시간

### IV. 유사도 계산 복잡도 모델링

실험 과정에서 최적의 차원 축소율이 존재함을 확인하였다. 본 장에서는 축소율과 1차 검색 후보 영상의 개수와의 관계를 분석하여, 전체 응답 시간의 대부분을 차지하는 유사도 계산 복잡도를 모델링한다. 그림 2는 512차원의 특징 벡터를 2배씩 단계적으로 축소할 경우, 질의 영상에 대한 전체 데이터베이스 내의 영상에 대한 유사도 분포를 나타낸다.

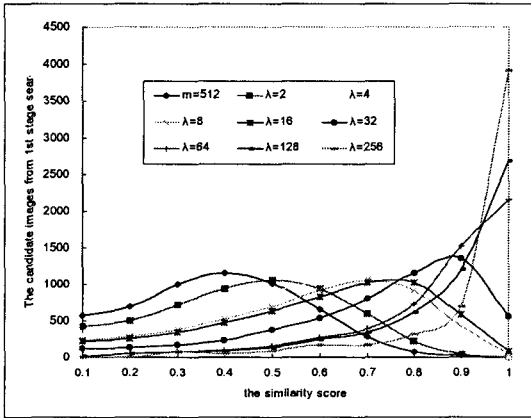


그림 2. λ변화에 따른 전체 영상에 대한 유사도 분포

그림 2의 결과로부터, 특징 벡터의 차원 축소에 따른 유사도 분포는 베타(beta) 분포를 따르는 것을 알 수 있고, 전체 차원 검색과 축소 2단계 검색의 유사도 계산 복잡도는 각각 (6), (7)과 같이 나타낼 수 있다.

$$C_F = m \times n \tag{6}$$

$$C_c = \left(\frac{m}{\lambda}\right) \times n + m \times (\xi(x) \times n) \tag{7}$$

(7)의 첫 번째 항은 1차 검색의 복잡도, 두 번째 항은 2차 검색의 복잡도를 각각 나타내며, λ는 특징벡터의 축소율, ξ(x)는 (8)과 같이 임의의 유사도 기준 k에 의한 1차 검색 후보 영상과 전체 데이터베이스 내의 영상 수의 비율을 나타내는 함수이다.

$$\xi(x) = \int_k^1 f(x) dx \tag{8}$$

f(x)는 베타 분포의 확률밀도함수이며 (9)와 같이 표현된다. 이 때, a=0, b=1인 경우에는 (10)과 같이 정규 베타 분포형태로 정의되며, (11)의 B(α, β)는 베타 함수를 나타낸다. 베타 분포는 α,β의 두 모수(parameter)에 의해 형태가 결정되며, 본 논문에서는 표본 데이터

로부터 α,β 두 모수를 추정하여 1차 검색 후보 영상 개수의 비율을 나타내는 ξ(x)를 근사화한다.

$$f(x) = \frac{1}{B(\alpha, \beta)} \frac{(x-a)^{\alpha-1} (b-x)^{\beta-1}}{(b-a)^{\alpha+\beta-1}} \tag{9}$$

$(a \leq x \leq b, \alpha > 0, \beta > 0)$

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \tag{10}$$

$(0 \leq x \leq 1, \alpha > 0, \beta > 0)$

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt \tag{11}$$

확률분포의 모수를 추정하기 위한 방법으로는 최우법(maximum likelihood estimation)과 적률법(method of moments) 등의 방법이 있다. 최우법은 추출된 표본의 확률이 최대가 되도록 모수를 추정하는 방법이다. 일반적으로 최우법은 표본의 양이 충분히 많을 때, 가장 효과적인 추정치를 얻을 수 있으며 다른 모수 추정 방법과의 신뢰도 기준으로 주로 사용된다. 그러나 표본의 크기가 작을 경우에는 추정 결과가 잘 일치하지 않는다는 단점이 있다. 그에 반해 적률법은 비교적 간단하여 가장 많이 사용하는 방법 중의 하나로 모집단의 적률과 표본 데이터의 적률을 같게 하여 모수를 추정하는 방법이다. 하지만 적률법은 분포 함수가 한 쪽으로 많이 치우치거나, 고차 적률로 갈수록 불안정한 추정치를 얻게 되는 단점이 있다. 여기서는 β와 λ의 표본 자료가 많지 않고, 2차 적률까지만 사용하면 되므로 적률법에 의해 모수 α,β를 아래와 같이 추정한다.

$$\bar{\alpha} = \mu \left( \frac{\mu(1-\mu)}{\sigma^2} - 1 \right) \tag{12}$$

$$\bar{\beta} = (1-\mu) \left( \frac{\mu(1-\mu)}{\sigma^2} - 1 \right) \tag{13}$$

표 1에서 특징 벡터의 차원 축소율에 따른 유사도 분포는 α가 고정된 상태에서 β가 축소율이 커짐에 따라 지수적으로 감소하는 형태임을 알 수 있다. 그림 3에서와 같이 모수 β와 축소율 λ와의 관계는 (14)와 같이 지수함수 형태로 표현할 수 있고, 표본 데이터로부터 최소제곱추정법(least square method)을 이용하여 근사화 할 수 있다.

$$\beta = p e^{-q \log_2 \lambda} \tag{14}$$

그림 4는 각 유사도 기준 k에서 차원 축소율에 따른 (7)의 복잡도 값을 나타내고 있다. 이 결과로부터, 유사도 기준 k가 높아지면 복잡도는 전체적으로 낮아지

표 1.  $\lambda$ 에 따른  $a, \beta$ 의 변화

$\lambda$	$a$	$\beta$
m=512	2.592620	4.026676
2	2.607978	3.162902
4	2.576512	2.310515
8	2.445409	1.728886
16	2.350184	1.510862
32	2.307367	0.868431
64	2.487203	0.403617
128	2.221846	0.306001
256	1.997740	0.145231

표 2. 실험에 의한  $\lambda$ 와 추정에 의한  $\lambda$

m	$\lambda_{opt}(실험)$	$\lambda_{opt}(추정)$
m=512	16	18.822892
m=256	8	10.297873

본 논문의 실험은 MPEG-7의 CCD(common color dataset) 5466장과 CCQ(common color queries) 50장에 대하여 실시하였고, MPEG-7의 색상기술자인 CSD(color structure descriptor)를 사용하였다.

### V. 결론

본 논문에서는 기존 전체 검색의 적합성을 그대로 유지하면서 검색 속도를 월등히 향상시킬 수 있는 축소 2단계 검색 방법에 대해 설명하였다. 여기서 전체 응답시간을 최소화시키는 축소율  $\lambda$ 값은 원 특징 벡터를 사용하여 계산한 유사도 분포곡선을 베타 분포로 추정하여, 차원  $m$ 의 특징 벡터에 대해 고정적인  $a$ 값과,  $\lambda$ 가 커질수록 지수감쇄를 하는  $\beta$ 값으로 추정할 수 있음을 알 수가 있었다. 즉, 임의의 차원의 특징벡터의 분포를 알면, 전체 응답시간을 최소화시키는 특징 벡터의 차원 축소율  $\lambda$ 를 측정할 수가 있다. 결론적으로 유사도 분포를 연속확률분포로 추정하여 유사도 계산 복잡도를 모델링 할 수 있었으며, 제안한 축소 2단계 검색 방법의 성능을 검증할 수가 있었다.

### References

- [1] B.V. Bonnlander and A.S. Weigend, "Selecting input variables using mutual information and nonparametric density estimation," *In Proc. of Int. Symp. on Artificial Neural Networks*, Taiwan, pp.312-321, Dec. 1994.
- [2] M. Beatty and B.S. Manjunath, "Dimensionality reduction using multidimensional scaling for content-based retrieval," *In Proc. of IEEE Int. Conf. on Image Processing*, USA, pp.835-838, Oct. 1997.
- [3] C. Faloutsos and K.-I. Lin, "FastMap: a fast algorithm for indexing, data-mining, and visualization of traditional and multimedia data sets," *In Proc. of 1995 ACM SIGMOD Int. Conf. on Management of Data*, USA, pp.163-174, May 1995.
- [4] A. Califano and R. Mohan, "Multidimensional Indexing for Recognizing Visual Shapes," *IEEE trans. on Pattern Analysis & Machine Intelligence*, Vol. 16, No. 4, pp. 373-392, Apr. 1994.
- [5] B.S. Manjunath, Philippe Salembier, Thomas Sikora, *Introduction to MPEG-7, Multimedia Content Description Interface*, John Wiley & Sons Ltd., 2002.
- [6] Vittorio Castelli, Lawrence D. Bergman, *Image Databases, Search and Retrieval of Digital Imagery*, John Wiley & Sons Inc., 2002.

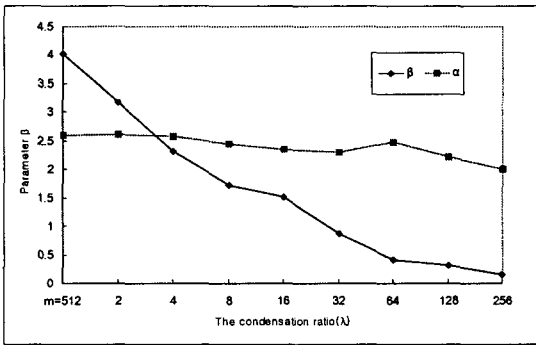


그림 3.  $\lambda$ 에 따른  $a, \beta$ 의 변화도

고,  $k$ 가 낮아지면 복잡도는 높아진다. 또한 유사도 기준  $k$ 는 전체적인 복잡도에는 영향을 미치지 않는다. 그러므로 추정된  $\lambda$ 값은 표 2와 같으며, 실질적으로 사용하는  $\lambda$ 값이 2의 멱승형태이기 때문에 임의의  $m$  차원의 원 특징벡터에 대한 최적의  $\lambda$ 값을 가까운 2의 멱승형태로 추정할 수 있다.

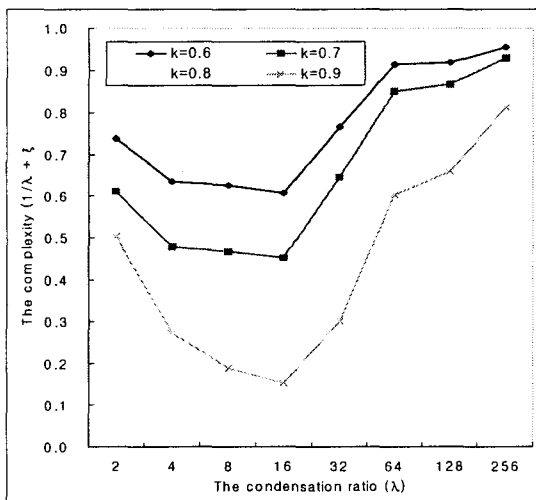


그림 4.  $k$ 와  $\lambda$ 에 따른 복잡도 변화