

고차국소 자기상관함수를 이용한 에지 특징벡터의 생성과 유사이미지에의 적용

윤미진⁰ 오근석^{**} 김판구^{*}
^{*}조선대학교 전자계산학과, ^{**}광주보건대학
 (mjyoon⁰, pkkim)[@]mina.chosun.ac.kr^{*}, okseok[@]www.kjhc.ac.kr^{**}

Edge Feature Vector Extraction using Higher-Order Local Autocorrelation and Its Application in Image Retrieval

Mi-Jin Yoon⁰ Kun-Sèok Oh^{**} Pan-Koo Kim^{*}
^{*}Dept. of Computer Science, Chosun University
^{**}Department of Digital Media, Gwangju Health College
 요약

본 논문에서는 자기상관함수의 국소적 특징을 사용하여 에지 특징을 추출한 후, 이를 이용해 유사이미지를 검색하는 방법을 제시한다. 자기상관함수의 국소적 특징을 이용하여 이미지를 검색할 경우 크기, 밝기, 색상등과 같은 이미지 요소가 서로 다를 경우에도 영향을 받지 않고 에지 특징정보를 추출해 낼 수 있다. 이는 얻어진 에지 특징을 이미지 크기와 고차 국소 자기상관함수의 변위에 의해 변하지 않도록 정규화를 하고, 동일 이미지에 대해 밝기가 조금 달라지면 검색효율이 떨어지는 점을 해결하기 위해 거리척도로서 방향여현거리(direction cosine distance)를 이용함으로써 가능하다. 이렇게 추출된 특징벡터를 자기조직화 맵에 의하여 클러스터링하고, 유사이미지 검색의 효율성을 비교해본 결과, 본 논문에서 제시한 방법을 사용하여 검색한 경우 재현율이 기존의 방법에 비해서 비교적 높은 수치를 나타냈다.

1. 서론

다량의 데이터베이스로부터 이미지를 검색하는 것은 최근 중요한 연구테마이다. 대부분의 사용자가 인터넷을 이용하여 멀티미디어 데이터베이스로부터 이미지를 탐색하거나, 표시하거나 한다. 여기에서 이미지를 어떻게 검색하는가가 큰 문제이다. 같은 이미지에 대해 개인이 붙이는 주석이 다르다는 점을 고려하여, 탐색을 할 때의 키워드와 주석을 기술하는데 걸리는 시간에 구애받지 않는, 내용기반 이미지검색(Content-based Image Retrieval: CBIR)이 필요하게 되었다. 이것은 서로 의미 내용이 많은 이미지데이터로부터 서로 가까운 특징량이 얻어진다는 논리에 근거를 두고 있다.

본 논문에서는, Haar 웨이블릿에 의하여 이미지를 주파수 분해하여, 그 고주파성분을 이용하여 에지 이미지를 생성한다. 그 에지 이미지로부터 고차국소 자기상관(Higher Order Autocorrelation)함수라는 기법을 이용하여 에지에 대한 특징 정보를 추출한다. 대규모의 이미지데이터베이스에서는 이미지는 서로 다른 사이즈, 밝기, 색상 등으로 저장되어 있으므로, 이러한 요소에 불변한 검색을 행하기 위하여 기존의 기법[2]에서는 이루어지지 않았던 고차국소 자기상관 함수의 정규화 및 방향여현거리(direction cosine distance)를 도입했다.

이미지의 클러스터링에는 SOM[1]을 이용했다. SOM의 학습은 다수의 특징벡터로 구성된 데이터집합을 2차원 상에 맵핑하여, 특징벡터와의 유사성을 발견하여 서로 유사한 특징벡터끼리 맵의 인접영역에 배치된다. 또한 질의 이미지에 의한 검색을 지원하는 시스템을 구현했다.

2. 관련연구

내용기반 이미지검색 시스템에는 IBM의 QBIC시스템[2]이 있다. QBIC은 이미지로부터 색상의 히스토그램, 텍스처, 형상, 오브젝트의 공간적 관련성을 특징량으로서 검색을 수행한다. 질의방법으로서는 예시이미지를 지정하거나, 형상을 스케치함으로써 검색이 가능하다. 그러나 이 시스템은 특징추출, 검색에 시간이 걸린다. Jacobs는 웨이블릿을 이용한 유사이미지 검색시스템을 구현했다[3]. VisualSEEK[4]는 이미지영역의 공간관계 조합과 압축영역으로부터의 특징추출을 이용하여 검색을 수행한다. 특징량은 최빈출 색상의 집합과 웨이블릿을 이용한 텍스처이다. SIMPLicity[5]은 Stanford 대학에서 개발한 세맨틱스 타입의 이미지 데이터베이스이다. 인덱스작성에는 4x4화소의 블록으로 분할하여 분류한다. 특징량으로서는 색상, 텍스처, 위치정보가 사용되었다.

대규모의 이미지데이터베이스에서 검색스피드를 향상시키기

위하여 특징벡터를 조직화하여 클러스터링 하는 것은 매우 중요하다. Oja는 PicSOM[6] 시스템에서 Tree 구조 자기조직화 맵(TS-SOM)을 제안하고 있다. 이는 SOM의 맵층을 계층화함으로써 특징벡터를 축약한 것이다. 다만, 이 시스템은 동적인 변화에 대응하지 못하고, 새롭게 이미지가 도입된 경우 많은 시간을 들여 SOM의 맵을 재학습하지 않으면 안된다.

3. 인덱스와 클러스터링

3.1 Haar 웨이블릿 변환

먼저 이미지의 색상공간인 RGB공간을 YIQ공간으로 변환한다. 에지이미지 작성에는 밝기정보인 Y값만을 이용한다.

Y값만을 이용함으로써

- (i) 색상에 의존하지 않는 검색: 자동차 등 모양은 달랐지만, 색상이 다른 것에 대한 검색
- (ii) 특징벡터의 차원수의 축약: RGB공간에서는 차원수가 3이지만, Y값만을 이용함으로써 차원수가 1이 된다.

가 가능하다.

다음으로, 이미지에 대하여 웨이블릿 변환을 수행한다. 웨이블릿 변환으로 얻어진 웨이블릿 계수공간에서 색상, 텍스처, 형상정보 등의 다양한 이미지내용에 대응한 특징량을 시스템 측에서 효율적으로 추출해 낼 수 있다. 1회 분해처리를 수행한 후에는 3개의 고주파 서브이미지와 1개의 저주파 서브이미지로 분해된다. 분해 후에는 저주파 성분 LL, 가로방향 차분이미지 LH, 세로방향 차분이미지 LH, 대각선방향 차분이미지 HH 등 4개의 서브 이미지로 이루어진다.(그림1)

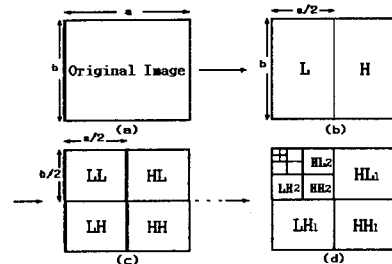


그림 1 웨이블릿 분해: (a) 원시이미지 (b) 세로 방향 분해후의 이미지 (c) 가로방향 분해후의 이미지. 저주파 성분 LL과 고주파 성분 HL, LH, HH로 분해된다. (d) 분해를 저주파 성분에 대하여 재귀적으로 반복한 후의 이미지

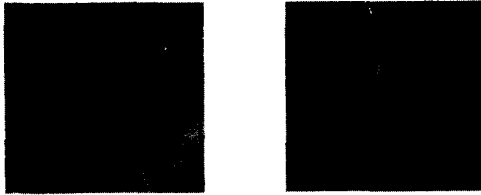


그림 2 (왼쪽):원시이미지 오른쪽:1회 분해후의 이미지

본 논문에서는 이미지의 에지 특징정보를 추출하기 위하여 Haar 웨이블릿 변환을 이용한다. Haar 웨이블릿 변환을 이미지에 적용한 경우 높은 압축률로 인한 블록 Distortion이 생긴다는 결점이 있지만, 본 논문에서의 기법에는 그다지 영향을 주지 않는다.

Haar 웨이블릿 변환을 이용한 장점은 다음과 같다.

- (i) Haar의 기저함수가 단순하므로 분해가 고속
Haar 기저 c 와 상세계수 d 는

$$\begin{cases} c_k^{j+1} = \frac{1}{\sqrt{2}}(c_{2k-1}^j + c_{2k}^j) \\ d_{k+\frac{1}{2}}^{j+1} = \frac{1}{\sqrt{2}}(c_{2k-1}^j - c_{2k}^j) \end{cases} \quad (1)$$

where $1 \leq k \leq L/2$

로 계산한다. 단, c_m^j, d_m^j 은 분해레벨 l , 왼쪽 위의 이미지 기준으로 한 m 번째의 화소값, L 은 저주파 성분的高度 (height), 또는 너비(width)를 나타낸다. 이 식이 나타내는 바와 같이 계산은 매우 단순하며 고속으로 수행된다.

- (ii) 이미지의 축소로 특징추출시의 계산량이 적어진다.
웨이블릿 변환은 2:1의 다운샘플링을 수행한다. 그림1에 나타난 바와 같이 웨이블릿 변환은 1회 분해할 때마다 1/2사이즈가 된다. 이미지는 2차원이므로 화소수는 1/4이다. 그러므로 특징추출의 경우 계산이 고속으로 수행된다.
- (iii) 에지특징만이 아니라 다른 특징추출에도 이용될 수 있다.

본 논문에서는 에지특징에 대해서만 언급하지만, 색상, 텍스처 등 그 이외의 특징정보를 추출하는 경우에도 이용되며, 그러한 특징정보는 유사검색을 지원하는 요소들이다.

3.2 에지이미지 작성

이미지를 웨이블릿 변환하여 계층구조로 표현할 수가 있다. 웨이블릿 변환 후 서브이미지 HL과 LH는 각각 가로방향, 세로방향의 에지를 표시한 것이다. 이 성질을 이용하여 에지 이미지를 작성한다. LH 서브이미지의 좌표(m,n)에 대응하는 요소를 $v_{m,n}$, 마찬가지로 HL 서브이미지의 요소를 $h_{m,n}$ 로 하고, 대응하는 에지이미지의 요소를 $e_{m,n}$ 로 하면 에지이미지의 요소는 다음과 같이 하여 계산할 수 있다.

$$e_{m,n} = \sqrt{(v_{m,n})^2 + (h_{m,n})^2} \quad (2)$$

where $1 \leq m \leq w, 1 \leq n \leq h$

이 식을 이용하여 만든 에지이미지를 그림3에 나타낸다. 여기서 문제가 되는 것이 웨이블릿 변환의 분해도지만, 분해도를 높이면 정확한 에지이미지를 작성할 수 없고, 분해도를 낮추면 계산시간이 걸린다. 본 연구에서는 분해도 1의 웨이블릿을 이용하여 에지이미지를 작성했다.

그림 3 에지 이미지

3.3 고차국소 자기상관함수를 이용한 특징벡터의 생성

에지정보의 특징량을 추출하는 방법으로써 고차 국소 자기상관함수를 이용한다[7].

이 방법에 의한 에지 특징추출의 장점은,

- (i) 이미지 내에 있는 물체의 위치에 관하여 불변한 특징 추출이 수행된다.
 - (ii) 특징벡터의 차원 수를 줄일 수 있다.
- 는 점에 있다. 이러한 장점은 유사검색을 수행하는 데에서 중요한 성질중의 하나이다. 벡터의 차원 수는 이미지사이즈에 관계없이 일정한 크기로 할 수 있다.

먼저 국소 마스크를 P 라고 한다. P 는 P 내에서 정의된 함수 $I(r) \geq 0$ (where $r \in P, \text{Supp}(I) = \{I(r) \geq 0\} \subset P$)

로 표현된다. $I(r)$ 의 P 내에서의 평행이동은

$$T(a)I(r) = I(r+a) \quad (3)$$

로 나타낸다. 다만, 여기에서의 변위 $a \in R^2$ 는 P 를 넘지 않는 범위이다. 본 기법에서는 $I(r)$ 이 밝기정보가 된다. N 차 자기상관함수는 참조점(reference point) r 주변의 N 개의 변위 a_1, a_2, \dots, a_N 대하여

$$R^N(a_1, \dots, a_N) = \sum_r I(r)I(r+a_1) \dots I(r+a_N) \quad (4)$$

으로 얻어진다. 물론, N 차 국소 마스크의 크기에 따라 이 벡터의 차원 수는 커질 수 있다. 그러나, 벡터의 차원 수를 줄이고, 계산 효율을 높임과 동시에 검색정도를 일정한 수준에 달하게 해야한다. 따라서 N 을 $0 \leq N \leq 2$ 로 하고, 국소 마스크의 크기를 3×3 으로 한정했다. 이때, 평행이동에 의한 동가패턴(equivalent pattern)을 제외하면 25개까지 마스크 패턴을 줄일 수가 있다.

이미지의 특징의 계산은, 국소 패턴의 대응하는 화소값의 곱을 합하면 된다. 각 패턴에 대응하는 특징을 f_1, \dots, f_{25} 라 하면, 이미지에 대한 고차국소 자기상관특징 벡터를

$$f = (f_1, \dots, f_{25}) \quad (5)$$

로 나타낸다. 이와 같이 계산된 특징은 이미지내의 오브젝트의 위치에 관하여 불변이다.[8] 이것은 매우 유효한 특징이다. 이상과 같이 제한된 25의 벡터에 $I(x,y)$ 있어서 참조점 r 을 (x,y) 그 점에 있어서의 밝기정보를 $I(x,y)$ 로 하여 f_i 를 다음과 같이 정의한다.

$$\begin{cases} f_1 = \sum_x \sum_y I(x,y) \\ f_2 = \sum_x \sum_y I(x,y)I(x,y+1) \\ \dots \\ f_5 = \sum_x \sum_y I(x,y)I(x-1,y-1) \\ f_6 = \sum_x \sum_y I(x,y)I(x-1,y)I(x+1,y) \\ \dots \\ f_{25} = \sum_x \sum_y I(x,y)I(x-1,y-1)I(x+1,y-1) \end{cases} \quad (6)$$

참조점 r 을 이미지의 가장 왼쪽위(첫번째)의 화소에 적용시켜 계산하는 경우 마스크 패턴에 따라서는 변위가 이미지에서 벗어나 버리는 경우가 있다. 본 기법에서는 변위가 이미지에서 벗어나 버리는 경우에는 그 계산을 하지 않는다.

만약, 이미지의 모서리의 화소를 모두 참조점으로 간주하지 않으면 무시되어, 이미지의 정확한 정보를 잃어버린다. 범위에서 벗어난 부분을 평균치 등으로 계산을 하려고 해도 그 값은 정확하지 않다. 다만, 마스크 패턴에 의하여 계산하기 위한 가로세로의 픽셀수가 달라지기 때문에, 다음에 논할 너비, 높이의 정규화에서는 이미지의 너비, 높이가 아니라 각각의 마스크 패턴에서 계산할 수 있는 너비, 높이를 의미한다.

이 특징벡터는 차수 N 이 $0 \leq N \leq 2$ 의 고차국소 자기상관 함수를 이용하고 있다. 그러나 이 차수의 차이는 유사검색에 나쁜 영향을 주게 된다. 또, 이미지 전체의 밝기가 밝을수록 고차국소 자기상관 함수의 값은 커진다는 성질을 가지므로, 이미지의 밝기에 대하여 불변한 검색을 행하기 위한 정규화를 했다.

3.4 특징 벡터의 정규화

이미지의 크기는 각각 다르고, 고차국소 자기상관 함수에서는 이미지의 크기에 따라 특징 벡터의 크기가 비례한다. 그래서, 함수를 높이 h , 너비 w 로 정규화 한다. 단, 여기에서의 높이 h , 너비 w 은 앞에서 논한 바와 같이 마스크 패턴을 적용할 수 있는 범위의 높이와 너비를 의미한다. 예를 들어 그림 4의 2개의 자동차의 사진은 육안으로는 매우 유사한 이미지이다.

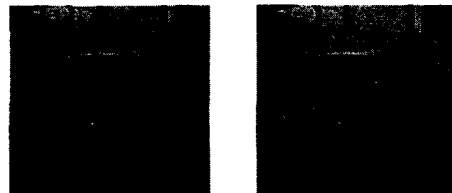


그림 4 크기가 다른 유사한 이미지 오른쪽:축소

그러나 고차국소 자기상관 함수에 의하여 분석해 보면 상호간의 차이를 발견할 수가 있다. 이 원인은 고차국소 자기 상관함수의 차수 N 에 있다. 원시 이미지의 밝기정보의 레벨을 S 라고 할 때, $N=0$ 에서는 $f_1 \propto S$ 임에 대하여 $N=1$ 에서는

$$f_2 \cdots f_5 \propto S^2, N=2 \text{에서는 } f_6 \cdots f_{25} \propto S^3 \text{ 이 된다.}$$

이러한 경우, $N=3$ 일때 거의 무시되어 버리므로 N 층의 근을 취하여 정규화 한다. 이와 같이해서 정규화 된 식은 아래와 같이 된다.

$$R^N = \frac{1}{wh} \sum_{r=1}^N \sqrt{I(r)I(r+a_1) \cdots I(r+a_N)} \quad (7)$$

이 식에 의하여 추출한 특징벡터를 이용하면 두 개의 유사 이미지의 차이를 구별할 수 있다.

또 하나의 문제로서 이미지 전체의 밝기정보의 차를 들 수가 있다. 동일한 사진이라도 빛의 정도에 따라 인간의 눈으로는 판단할수 없을 정도의 미묘한 차이지만, 화면전체의 밝기의 차가 생긴다. 그림 4는 같은 사진처럼 보이지만, 실제로 밝기가 조금 다르다. 그러므로 밝기정보의 변위에서 그 차이를 분석할 수 있다. 이를 위해 표준편차를 사용할 수 있지만, 보다 효과적인 결과를 얻을 수 있는 기법을 다음절에서 논하기로 한다.

3.5 거리 척도

멀티미디어 데이터 베이스에서의 거리계산에서 가장 자주 이용되고 있는 거리척도는 유클리드 거리이다. 그러나 우리의 예지 특징에는 적합하지 않다. 이미지의 밝기에 차가 있는 경우, 인간이 보고 달랐다고 생각되는 이미지를 유클리드 거리로 계산하면 부적절한 결과가 나온다. 그래서 방향여현(direction cosine) 거리를 사용한다.

그림 5는 3개의 이미지의 예지 특징벡터를 모식적으로 표현한 것이다. 인간이 보고 달랐다고 생각되는 이미지는 b 와 c 다. 밝기가 다르므로 벡터의 크기는 다르지만 방향은 거의 같다.

a 와 c 는 그다지 달지 않은 이미지이지만, 유클리드 거리를 이용하면 $|a-d| \leq |b-c|$ 가 되며 달랐다고 판단된다. 방향여현 거리에서는 벡터간의 각도가 작은 쪽이 서로 가까운(달은) 이미지이다.

$\theta_{bc} \leq \theta_{ac}$ 이므로 b 와 c 가 가까운 것으로 계산된다. 벡터 x, y 의 방향여현 거리는 다음 식으로 정의된다.

$$\text{dist}(x, y) = \arccos \frac{x \cdot y}{|x||y|} \quad (8)$$

이상과 같이 거리척도로서 방향여현 거리를 사용하는 것이 예지특징의 유사성을 계산하는 데는 매우 유효하다.

4. 시스템의 구현과 실험

본 시스템은 SUN Microsystem 사의 Ultra10(OS: Solaris 5.8, memory:128MB)의 워크스테이션에서 구현했다. 프로그램은 C++와 ODQL(Object Database Query Language)에 의하여 코딩했으며 clustering 은 SOM을 이용했다.

실험에서 사용한 이미지데이터는 Stanford 대학에서 다운로드 한 이미지[9]에서 매우 비슷한 이미지 5개씩 16조, 합계 80개를 이용했다. 이미지의 장르는, 자연이미지, 인공이미지, 자동차 등 여러 갈래로 분류된다. 데이터베이스로서는 작은 규모이지만, 비교실험을 수행하기 위해서는 모든 이미지를 인간이 파악하고 있어야 할 필요가 있다. 이 단계에서의 분류는 인간이 모두 수작업으로 행한다. 그 후, 이미지에서의 특징추출, SOM을 이용한 클러스터링을 행한다. 질의이미지에는 처음에 분류한 16개의 조 가운데에서 10개의 이미지를 무작위로, 동시에 같은 조에서 여러 개 선택되지 않도록 결정했다. SOM의 맵층에도 4x4의 맵을 이용한다. 평균하면 하나의 클러스터에 5개의 이미지가 분류되게 된다. 다른 기법과 비교를 하기 위해 다음의 두가지 방법에 대해 Precision ratio(적합률)과 Recall ratio(재현율)를 구했다.

(i) COS

고차국소 자기상관함수(식7)에 , 거리척도는 방향여현 거리를 이용한 것. 본 논문에서의 제안 기법

(ii) NORM

고차국소 자기상관함수(식4)에 , 거리척도는 유클리드 거리를 이용한 것, 기존에 사용되는 기법

Precision, Recall을 각각 다음과 같이 정의한다. 데이터베이스 내에서의 적합 이미지를 $I_{re\ lewant}$, 질의 이미지에 대한 검색 결과를 $I_{re\ trievied}$ 로 하면,

$$Precision = \frac{I_{re\ trievied} \cap I_{re\ lewant}}{I_{re\ trievied}}, Recall = \frac{I_{re\ trievied} \cap I_{re\ lewant}}{I_{re\ lewant}}$$

로 나타낼 수 있다. 본 실험에서는 $I_{re\ lewant} = 5$ (사용자가 검색하고자 하는 이미지수)가 된다. 즉 검색결과로서 얻은 이미지의 개수가 10개, 그 중에서 적합 이미지가 3개일 때 Precision 은 $3/10 = 0.3$, Recall 은 $3/5 = 0.6$ 이 된다. 실험결과 Precision은 본 기법에서 약 49%, 기존 기법에서 약 48%로 어느 쪽도 별로 차이가 없으나, Recall에서는 본 기법에서는 88%, 기존의 기법에서는 66%로 본 기법은 기존의 기법에 비교하여 압도적으로 높은 수치를 보여주고 있다.

표1 SOM 에서의 Recall과 Precision의 비교

	Recall	Precision
COS	0.88	0.49
NORM	0.66	0.48

5.결 론

본 논문에서는 예지 특징에 의한 유사이미지 검색에 대하여 논하였다. 예지 특징 추출을 위하여 웨이블릿 변환과 고차국소 자기상관함수를 적용했으며, 클러스터링에는 SOM을 이용했다. 예지 특징정보는 유사이미지 검색에 있어서 중요한 요소이며, 본 논문에서 제안한 예지 특징은 색상, 오브젝트의 위치, 이미지 사이즈, 이미지의 밝기에 불변한 것이다. 현재의 시스템에서는 이미지에 의한 검색만 가능하지만, 인터페이스 상에서 스케치한 이미지에 의한 검색기능을 추가할 예정이다.

실험에서는 기존의 방법과 본 제안기법을 비교했다. 각각의 Precision, Recall을 비교하였고, 높은 검색율을 얻었다. 또, SOM을 이용하여 검색시간의 고속화를 실현했으며 높은 검색 정도를 얻을 수 있었다.

SOM을 이용하면 검색결과가 반드시 순차검색의 결과와 일치하지 않는 경우가 있다. 그러나 본 연구와 관련하여 완전탐색이 가능하며, 동시에 고속의 최근접 탐색 알고리즘인 SOM-based K-NN search[10][11]를 고안했다. 향후 본 시스템에 이러한 알고리즘을 구현할 예정이다.

참고문헌

- [1] T.Kohonen, Self-Organizing Maps Third Edition. Springer-Verlag, 2000.
- [2] M. Flickner, H. Sawhney, and W. Niblack, "Query by Image and Video Content:The QBIC System," IEEE Computer Magazine, Vol.28, Sep 1995.
- [3] C.E. Jacobs, A. Finkelstein, and D.H. salesin, "Fast multiresolution image quering," in Proc. SIGGRAPH95, 1995.
- [4] J.R. Smith and S.-F. Chang, "VisualSEEK:A Fully Automated Content-Based Image Query System," ACM Multimedia, 1996.
- [5] J.Z. Wang, "SIMPLcity:A region-based image retrieval system for picture libraries and biomedical image databases." ACM Multimedia Conference, 2000.
- [6] M. Koskela, J. Laaksonen, S. Laakso, and E. Oja, "The picsom retrieval system: Description and evaluation," CIR2000, May 2000.
- [7] T. Kurita, N. Otsu, and T. Sato, "A face recognition method using higher order local autocorrelation and multivariate analysis," in Proc. PATTERN RECOGNITION, 1992.
- [8] M. Kobayakawa, M. Hoshi, T. Ohmori, and T. Terui, "Interactive Image Retrieval Based on Wavelet Transform and Its Application to Japanese Historical Image Data," IPSJ Journal Vol.40, No.3, pp. 899-911, Mar 1999(printed in Japanese).
- [9] "http://www-db.stanford.edu/~wangz/image.vary.jpg.tar"
- [10] Z. Aghbari, K.-S. Oh, Y. Feng, and A. Makinouchi, "SOM-Based K-Nearest Neighbors Search In Large Image Databases," The 6th IFIP Working Conference on Visual Database Systems(VDB-6), Kluwer Academic Publishers, 2002.
- [11] K.-S. Oh, Z. Aghbari, P.-K. Kim, "Fast k-NN Image Search with Self-Organizing Maps." CIVR 2002.