

이미지 데이터베이스 유사도 순위 매김 알고리즘

(A Similarity Ranking Algorithm for Image Databases)

차 광 호 †

(Guang-Ho Cha)

요 약 이 논문은 이미지 데이터베이스를 위한 유사도 순위 매김 알고리즘을 제시한다. 이미지 검색의 문제점 중 하나가 이미지로부터 자동적으로 계산한 하위 레벨 특성과 인간 지각과의 의미 차이이며, 검색 시에 이미지 유사도 측정을 위해 많은 알고리즘에서는 민코프스키 측정법(L_p -norm)을 사용하고 있다. 그러나 민코프스키 측정법은 인간 시각 시스템의 비선형적 특성과 문맥 정보를 반영하지 못한다. 본 알고리즘에서는 인간 지각의 비선형성과 문맥 정보를 반영하는 유사도와 탐색 알고리즘을 통해 이 문제를 해결한다. 본 알고리즘을 필기체 숫자 이미지 데이터베이스에 적용하여 성능의 우수성과 효과를 증명하였다.

키워드 : 이미지 데이터베이스, 내용기반 검색, 유사도 탐색, 유사도 순위 매김

Abstract In this paper, we propose a similarity search algorithm for image databases. One of the central problems regarding content-based image retrieval (CBIR) is the semantic gap between the low-level features computed automatically from images and the human interpretation of image content. Many search algorithms used in CBIR have used the Minkowski metric (or L_p -norm) to measure similarity between image pairs. However those functions cannot adequately capture the aspects of the characteristics of the human visual system as well as the nonlinear relationships in contextual information. Our new search algorithm tackles this problem by employing new similarity measures and ranking strategies that reflect the nonlinearity of human perception and contextual information. Our search algorithm yields superior experimental results on a real handwritten digit image database and demonstrates its effectiveness.

Key words : image database, content-based retrieval, similarity search, similarity ranking

1. 서 론

유사도 탐색(또는 k -최근접 탐색)은 멀티미디어 정보 검색에 필수적이다. 내용기반 이미지 검색에서 이미지는 아주 고차원 공간에서의 특성 벡터로 표현되고, 사용자는 입력 이미지와 가장 유사한 k 개의 이미지를 요구한다. 이미지 간의 유사도는 보통 유클리드 거리로 측정하는데, 때때로 유사도 비교에 사용된 하위 레벨 특성이

사용자 인식과 불일치를 나타낼 수 있다. 본 논문에서는 새로운 유사도와 탐색 알고리즘을 통해 이러한 의미 차이를 해결한다.

일반적인 유사도 측정법이 바람직한 결과를 낳지 못하는 실제 예로 보기 1은 필기체 숫자 이미지 데이터베이스에서의 실험 결과이다

보기 1. 필기체 숫자 이미지 데이터베이스에서 주어진 질의 이미지와 비슷한 두 개의 이미지를 찾고자 한다. 그림 1(a)는 질의 이미지이고, (b)와 (c)는 사람이 직접 유사한 두 개의 이미지를 선택한 결과이다. 그림 2는 유클리드 거리 함수를 사용하여 질의 이미지 (a)와 가장 가까운 거리에 있는 2 개의 이미지를 추출한 결과이다. 그림 1과 2의 결과에 따르면 인간의 지각과 민코프스키 거리 함수와 같은 유사도 측정법 사이에는 차이가 있으며, 이러한 차이를 메울 필요가 있다.

효과적인 내용기반 이미지 검색을 위해 이미지 내용을 충실히 표현하기 위한 많은 휴리스틱들이 개발되었으며, 검색에서 이미지 내용을 집합적으로, 즉, 문맥에

· 이 논문은 2008년도 교육과학기술부의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. D00309)

† 중신회원 : 서울산업대학교 컴퓨터공학과 교수

ghcha@snu.ac.kr

논문접수 : 2008년 12월 11일

심사완료 : 2009년 6월 24일

Copyright©2009 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지 : 데이터베이스 제36권 제5호(2009.10)

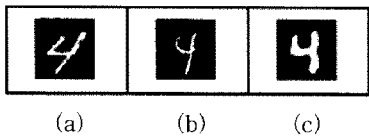


그림 1 사람이 직접 유사한 이미지를 선택한 경우, (a): 질의 이미지, (b), (c): 사람이 선택한 이미지

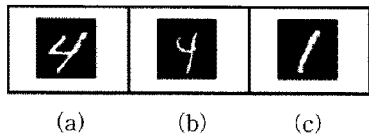


그림 2 유클리드 거리 함수를 사용하여 유사한 이미지를 선택한 경우, (a): 질의 이미지 (b), (c): 결과로 선택된 두 이미지

기반하여 해석하고 사용자 질의에 적합한 정도에 따른 순위 매김(ranking)이 중요하다.

[1-5]에서는 검색 효과를 높이기 위해 적합성 피드백(relevance feedback)의 사용을 제시하였다. 비록 적합성 피드백이 검색 효과를 높이기 위한 한 방법이지만 여전히 그 효과는 유사도 측정 방법과 순위 매김 알고리즘에 의해 크게 영향을 받는다. 적합성 피드백의 또 다른 문제는 여러 차례의 피드백이 필요하여 시간이 많이 소요된다.

보기 2. 그림 3은 본 연구의 동기가 되는 또 다른 예를 나타낸다. 두 개의 클러스터(cluster)로 이루어진 데이터 점들의 집합이 있다. 질의점(query point)은 '+'이고 사용자는 질의점에 가장 가까운 6 개의 점을 찾고자 한다. 만약 유클리드 거리 함수를 사용한다면 그 점들은 (a)에서 '+'를 중심으로 여섯 번째로 가까운 점까지의 거리를 반지름으로 하는 원 내부의 점들이 될 것이다. 그러나 주어진 데이터 집합의 분포를 고려할 때 (b)에 나타낸 것과 같이 클러스터 A에 있는 6 개의 점들이 클러스터 B에 있는 6 개의 점들보다 질의 점에 더욱 관련되어 있다고 할 수 있다. 왜냐하면 군집을 이루는 데이터들은 비슷한 성격을 지니며, 클러스터 A가 클러스터 B보다 질의 점에 더 가깝기 때문이다.

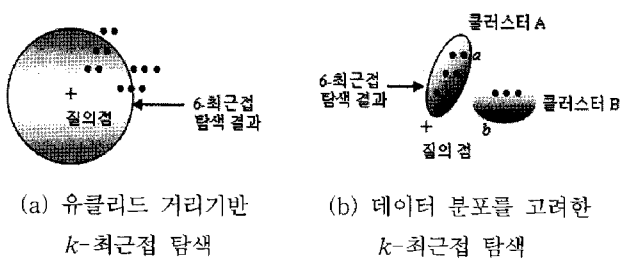


그림 3 두 개의 다른 유사도 모델에 따른 k-최근접 탐색 결과

이상의 두 예를 통해, 서로 인접한 점들이 유클리드 거리기반 순위에서는 떨어져 있는 다른 점들보다 후순위가 되지만, 내용기반 유사도 순위에서는 비슷한 순위를 갖도록 하는 것이 바람직하다고 가정한다. 이와 같은 개념에 기반하여 순위 매김을 수행하면 그림 3(b)의 클러스터 A의 점 a 가 클러스터 B의 점 b 보다 질의점 '+'에서 더 멀리 있지만 유사도 순위에서는 더 앞서게 되며, 이 개념은 [6]의 클러스터에 대한 가정과 일치한다.

이와 같은 관점에서 본 논문에서는 문맥 정보를 주어진 데이터 집합의 분포와 클러스터 구조에 관한 정보로 정의하고, 문맥 정보에 존재하는 비선형(nonlinear) 관계를 도출하기 위해 원 데이터 집합에 대해 비선형 변환을 수행한다. [7,8]에 따르면 가우스 함수가 우수한 비선형 근사 능력을 갖고 있으므로 본 논문에서는 비선형 변환을 위해 가우스 함수를 사용한다.

2. 관련 연구

많은 연구에서 이미지 내용을 표현하기 위해 색상, 질감, 모양 등의 시각적 특성들을 사용하였다. 그러나 이러한 특성들과 이미지 의미 간에는 차이가 존재하므로 많은 연구에서 자동적으로 추출한 시각적 특성과 인간의 해석 간의 의미 차를 좁히기 위해 적합성 피드백의 사용을 제안하였고, 적합성 피드백의 효과를 높이기 위해 지도 학습(supervised learning)[9]을 활용하는 연구도 진행되었다. Tong과 Chang은 훈련 데이터로 가장 유의한 샘플을 반복적으로 첨가하여 결정 경계(decision boundary)를 학습하기 위해 SVM_{active}를 제안했다[4]. Hoi와 Lyu는 결정 경계를 학습하는데 피드백 신뢰도를 고려하는 soft-label 지지 벡터 기계(support vector machine: SVM)을 제안했다[10]. 이상의 노력에도 불구하고, 비록 SVM[9]이 분류에는 효과적이지만 온라인 훈련을 위한 샘플의 수가 적을 때는 유용한 결정 경계가 불안정하다.

내용기반 이미지 검색의 최근 흐름은 적합한 이미지 공간에서 이미지 집합의 고유 구조를 찾는 쪽으로 전개된다. 즉, 전통적인 유클리드 공간에서 작업하는 대신 이미지들이 매니폴드(manifold)로 펼쳐져 있다고 가정하고, 데이터 집합을 형성하는 구조를 찾아서 그 특성을 나타낸다. 따라서 학습된 매니폴드 위에서 유사도를 계산할 수 있다. He 등[11]은 매니폴드를 따라 이미지 간의 거리를 근사하기 위해 측지선 거리(geodesic distance)를 사용하였다. 그러나 주된 약점은 학습된 데이터 집합에 대해서만 매핑이 정의되어 실험 데이터를 처리하기 위해서는 방사형 기저 함수(radial basis function) 네트워크같은 추가적인 구성이 필요하다. 비슷하게 Wu 등[12]은 문맥기반 거리 함수를 공식화하기 위한 방법을

제안하였는데, 커널 함수[9]를 사용하여 전통적인 거리를 변환된 특성 공간에서의 유사도로 비선형적으로 변환하였다. 그러나, 문맥 정보를 수집하기 위해 사람이 개입해야 하며, 문맥 정보를 훈련 데이터 형태로 요구하고 있다. [13]에서는 민코프스키 측정법보다 인간 지각 유사도에 더 가까운 거리 함수로 비계량(non-metric) 거리 함수인 동적 부분 함수(dynamic partial function)를 제안하였다. 그러나, 이 함수의 거리 계산에 사용할 특성을 동적으로 선택하는 것이 어렵다.

3. 비선형 유사도 모델

3.1 비선형 유사도 함수

이미지 유사도 평가를 위해 인간의 지각을 시뮬레이션하기 위해 본 논문에서는 제일 먼저 비선형 모델을 수립한다. 비선형 접근법의 가정은 인간 지각의 경우 같은 길이의 거리가 항상 같은 정도의 유사도를 제공하지는 않는다는 것이다[14]. 비선형 모델은 입력 이미지 x 의 특성 값을 사용하여 주어진 질의에 대한 유사도를 계산하는 입출력 사상 함수에 의해 형성되고, 본 논문에서는 입출력 사상 함수로 다음 유사 가우스 함수를 채택한다:

$$G(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{d_i d_j}\right) \quad (1)$$

여기서 d_i 와 d_j 는 그 이하에서는 각각 두 점 x_i, x_j 와 유사하다고 판단할 수 있는 참조 거리이다. 이 참조 거리를 계산하는 한가지 방법은, 해당 점에서 k 번째 최근 점 이웃까지의 거리로 지정할 수 있다. 이 원칙에 따라 본 논문에서는 각 점 x_i 마다 k 번째로 가까운 점까지의 거리를 d_i 로 지정한다. 한 점 x_i 에서 다른 점 x_j 까지의 거리를 $\|x_i - x_j\| / d_i$ 로 나타낼 때, 거리 $\|x_j - x_i\| / d_j$ 는 x_j 에서 x_i 까지의 거리이다. 그러면 두 점 x_i, x_j 간의 거리의 제곱은 $\|x_i - x_j\|^2 / d_i d_j$ 로 표현되고, 그 둘 간의 유사도를 $\exp(-\|x_i - x_j\|^2 / d_i d_j)$ 로 표현할 수 있다.

공식 (1)에서 함수 G 는 참조 거리 d_i, d_j 를 제외하고는 가우스 함수 $\exp(-\|x_i - x_j\|^2 / \sigma^2)$ 와 비슷하다. 본 논문에서는 각 점의 지역성(locality)을 포함하기 위해 유사도 함수에 개별적인 참조 거리 d_i 와 d_j 를 포함하였다. 이것은 유사도 함수가 질의점에 따라 가변적임을 의미하고 이것은 질의점 주변의 지역적 통계적 특성을 이용한다는 점에서 내용기반 이미지 검색에서 바람직하다. 더욱이, 지수적(exponential) 유사도 함수는 지역적 변화에 더 민감하고 인간의 지각적인 관점에서 더 나은 성능을 보인다.

3.2 비선형 변환 구조

점 들의 분포를 나타내는 그림 3(a)에서 점들 간의 상호 위치는 전역(global) 정보(즉, 데이터 집합이 두 개의 클러스터로 구성됨)를 포함한다. 그림 3(b)와 같이 분류하기 위해서는 이 전역 정보를 도출해야 하는데 이것은 어려운 문제이므로 본 논문에서는 점 들 간의 지역(local) 정보를 도출하여 전역 정보를 유추한다. 즉, 점 들 간의 상호 유사도를 통해 데이터 집합의 전역 정보를 도출하며 이 아이디어는 Shi와 Malik의 연구[15]에 기반한다.

만약 데이터 집합이 m 개의 이미지를 포함한다면 공식 (1)에 따라 두 이미지 간의 유사도를 나타내는 $m \times m$ 유사도 행렬 S 를 다음과 같이 생성한다:

$$S_{ij} = \exp(-\|x_i - x_j\|^2 / d_i d_j)$$

두 이미지가 비슷하다면 S_{ij} 는 큰 값을 갖고, 그렇지 않다면 S_{ij} 는 0에 가까워진다. 이 유사도 함수는 다음의 성질을 갖는다:

$$\forall i, j: S_{ij} \in [0, 1], S_{ii} = 0, S_{ij} = S_{ji}.$$

첫번째 특성은 S 에 대한 정규화(normalization)를 의미하고, 두번째 특성은 자신과의 유사도는 0임을 나타낸다. 세번째 특성은 대칭성을 나타낸다.

데이터 집합으로부터 전역 정보를 도출하고 데이터 집합의 잡음을 없애기 위해 유사도 행렬 S 에 주성분 분석(principal component analysis) 같은 고유 분해(eigen-decomposition) 기법을 적용하고, 이것은 데이터 집합에 고유한 클러스터 구조에 관한 특성을 도출한다.

본 논문의 비선형 변환을 요약하면, 공식 (1)의 유사도 사상 함수 G 를 통해 원 데이터를 가우스 특성 공간(feature space)으로 사상시키고, 사상된 데이터에 대해 고유 분해를 적용한다. 모든 유사도 비교는 특성 공간에서 수행되고, 유사도는 이 특성 공간에서의 유클리드 거리로 정의한다. 변환된 특성 공간에서 고유 벡터(eigenvector)를 추출하는 아이디어는 커널 주성분 분석(kernel PCA)[16]과 스펙트럴 클러스터링(spectral clustering)[17]에 기반한다.

보기 3. 그림 4는 분명한 클러스터 구조를 갖는 3 개의 표본 데이터 집합을 사용하여 비선형 변환을 수행한 결과를 보여준다. 비선형 변환의 목적은 원 데이터 집합의 고유 구조를 나타내게 하는데 있다.

4. 유사도 순위 매김 알고리즘

4.1 단순 유사도 순위 매김 알고리즘

본 논문의 첫번째 순위 매김 알고리즘(Algorithm 1)은 그림 4에 나타난 것처럼 데이터 간의 관계가 확실하게 나타나는 특성 공간에서 유사도를 비교한다. 이러한 결과를 얻기 위해, 유사도 행렬 S 에서 고유 벡터들을 추출한다.

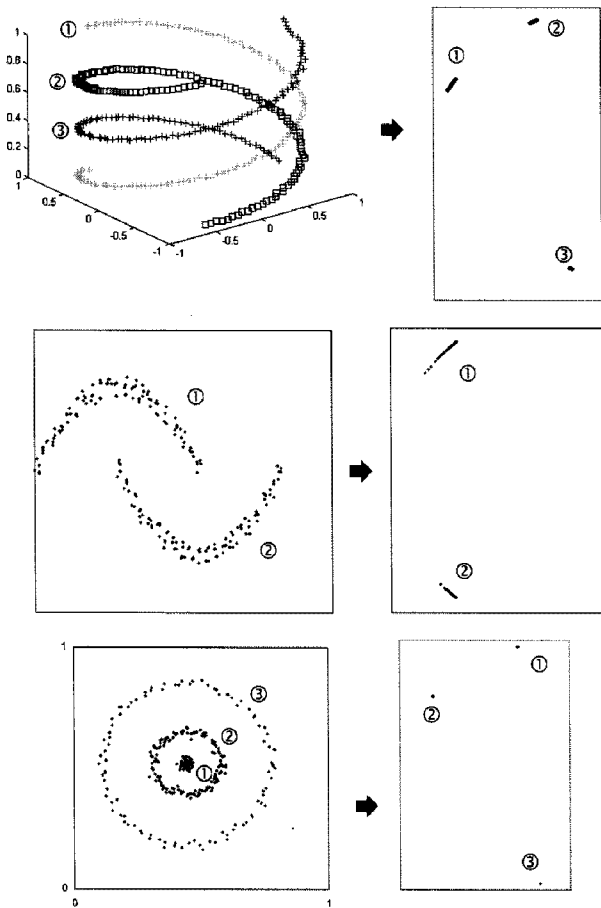


그림 4 3 개의 표본 데이터 집합을 사용하여 비선형 변환을 수행한 실제 예: 입력 데이터는 좌측, 변환된 데이터는 우측에 표시하였다.

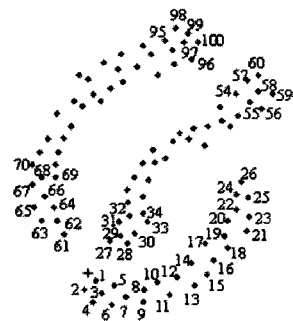


그림 5 질의 점 '+'에 대해 100 개의 점에 대한 유사도 순위

순위 매김 알고리즘은 다음과 같다. 주어진 데이터 점들의 집합 $X = \{x_1, \dots, x_m\} \in R^n$ 을 특성 벡터에 의해 변환된 특성 벡터 공간으로 사상시키고, 이 공간에서 순위 매김을 실행한다. 여기서 R^n 은 n 차원 실수 공간이다.

Algorithm 1

- 아래와 같이 정의된 유사도 행렬 $S \in R^{m \times m}$ 를 생성한다.

$$S_{ij} = \exp(-\|x_i - x_j\|^2 / d_i d_j) \text{ if } i \neq j, \text{ and } S_{ii} = 0$$

- 여기서 d_i 와 d_j 는 각각 x_i 와 x_j 에 대한 10 번째 최근 점 이웃 간의 거리이다.
- (i, i) -요소가 S 의 i -번째 행의 합인 대각 행렬 D 를 형성한다.
- 정규화된 유사도 행렬 $S' = D^{-1/2} S D^{-1/2}$ 를 형성한다.
- S' 로부터 K 개의 가장 큰 특성 벡터 v_1, v_2, \dots, v_K 를 도출하고, 이 특성 벡터를 열 단위로 정렬한 행렬 $V = [v_1 v_2 \dots v_K] \in R^{m \times K}$ 를 형성한다.
- 행렬 V 에서 각 행이 단위 길이를 갖도록 정규화시킨 행렬 Z 를 형성한다. 즉, $Z_{ij} = V_{ij} / (\sum_j V_{ij}^2)^{1/2}$.
- Z 의 각 행은 특성 공간 R^K 의 한 점이다.
- 이 특성 공간 R^K 에서 유클리드 거리를 사용하여 질의 점에 대한 이미지 순위를 계산한다.

위 단계 1에서 각 점 들간의 유사도로 구성되는 행렬 S 를 형성한다. 즉, $S_{ij}(i \neq j)$ 는 두 점 x_i 와 x_j 간의 유사도를 나타낸다. 자신과의 유사도 S_{ii} 는 0로 주어진다. 중요한 한가지는 유사도 행렬 S 와 마지막 데이터 집합 Z 는 탐색 전에 미리 계산되므로 실제 탐색 시에 부담이 되지 않는다. 단계 2-3과 5는 정규화 과정이다.

보기 4. Algorithm 1의 효과를 검증하기 위해 그림 5에서 100 개의 표본 데이터에 대해 순위 매김을 수행하였다. 데이터 집합은 3 개의 클러스터 구조로 되어있다. '+'로 표시한 것이 질의점이고, 각 점 옆에 표시한 숫자가 질의점에 대한 유사도 순위를 나타낸다. 그림 5에서 보듯이 본 알고리즘은 데이터 집합의 문맥을 매우 잘 반영하고 있다.

Algorithm 1은 그림 4와 5에서와 같은 저차원 데이터에 대해서는 매우 잘 동작한다. 그러나, 본 연구와 같이 고차원의 이미지 데이터 집합에 대해서는 제한적인 성공을 나타낸다. Algorithm 1의 단계 4에서 유사도 행렬의 고유 분해를 통해 차원 축소가 일어난다. 그러나 이미지 데이터 집합과 같이 아주 고차원 데이터에 대해서는 차원의 저주(dimensionality curse)로 인해 우수한 K 개의 특성 벡터를 구하기 어렵고, K 의 결정도 힘들다. 따라서, 고차원 데이터를 위해 차원 축소를 수행하지 않는 유사도 분포에 기초한 순위 매김 알고리즘과 새로운 유사도 측정법을 제시한다.

4.2 유사도 분산에 기초한 순위 매김 알고리즘

데이터 집합이 나타내는 고유 구조를 고려하기 위해 본 논문에서는 유사도 분포(similarity distribution)의 개념을 도입한다. 순위를 정할 점들의 집합을 $X = \{x_1, x_2, \dots, x_m\} \in R^n$ 로 표현하고, 질의점은 x_q 라 하자. 벡터 s_i 를 다음과 같이 정의한다: $s_i = [S_{iq}, S_{i1}, S_{i2}, \dots, S_{im}]^T$, 여기서 S_{ij} 는 두 점 x_i, x_j 간의 유사도이다. 유사도 S_{ij} 는 공식 (1)로 계산된다. 본 논문에서는 벡터 s_i 를 한 점 x_i

와 질의점을 포함하는 데이터 집합에 있는 모든 다른 점들 간의 유사도 분포로 간주한다. 벡터 s_q 는 질의점 s_q 와 질의점 자체를 포함하여 모든 다른 점들 간의 유사도 분포로 간주한다: $s_q = [s_{qq}, s_{q1}, s_{q2}, \dots, s_{qm}]^T$. 본문에서 s_{qq} 는 1.0으로 정의한다. 그러나 s_{ii} ($i \neq q$), 즉, 자신과의 유사도는 0.0으로 정의한다.

한 점 x_i 와 질의점 x_q 간의 유사도를 가우스 특성 공간에서 x_i 에 대한 유사도 분포와 x_q 에 대한 유사도 분포 간의 내적(dot product)으로 정의한다. 그러면, 한 점 x_i 의 질의점 x_q 에 대한 유사도 S_{iq} 는 다음과 같이 계산된다:

$$s_{iq} = s_i^T \cdot s_q = s_{iq} s_{qq} + \sum_{j=1}^m s_{ij} \cdot s_{qj} = s_{iq} + \sum_{j=1}^m s_{ij} \cdot s_{qj} \quad (2)$$

공식 (2)에서 s_i 와 s_q 는 각각 x_i 와 x_q 에 대한 유사도 분포이고, 유사도 s_{ij} 와 s_{qj} 는 공식 (1)로 계산된다. 공식 (2)로 주어진 유사도는 질의점과 점 x_i 간의 실제 유사도와 점 x_i 와 그 이웃 점들 간의 유사도를 이웃 점들의 질의 점에 대한 유사도를 가중치로 하여 선형 결합한 결과를 더한 것을 나타낸다. 따라서 한 점의 유사도는 그 이웃 점들의 유사도에 영향을 주고, 어떤 두 점에 대해서 질의점에 대한 그 점의 유사도에 두 점 간의 유사도를 선형 결합하기 때문에 만약 두 점이 가깝다면 서로 더 많은 영향을 준다. 유사도 분포에 기초한 본 유사도 측정법에서는 질의점에 가까이 군집화되어 있는 점들이 순위 매김에서 유리하게 된다. 구체적인 순위 매김 알고리즘은 다음과 같다:

Algorithm 2

[입력] 데이터 점들의 집합 $X = \{x_1, \dots, x_q, x_{q+1}, \dots, x_m\} \in R^n$, 여기서 x_1, \dots, x_q 는 질의점들이고, 나머지 x_{q+1}, \dots, x_m 는 순위를 매길 데이터 점들이다.
 [출력] 질의점에 대한 데이터 점들의 유사도 순위

1. 아래와 같이 정의되는 유사도 행렬 $S \in R^{m \times m}$ 를 생성한다: $S_{ij} = \exp(-\|x_i - x_j\|^2 / d\alpha_j)$ if $i \neq j$, and $S_{ii} = 0$
2. 대각 행렬 D 를 생성한다. D 의 (i, i) -요소는 S 의 i -번째 행의 합이다.
3. 정규화된 유사도 행렬 $S' = D^{-1/2} S D^{-1/2}$ 를 생성한다.
4. 각 점 x_i 와 질의점 x_j 간의 초기 유사도를 다음과 같이 계산한다, 여기서, $1 \leq j \leq q, q+1 \leq i \leq m$.

$$S_{ij} = \begin{cases} 1 & \text{for } 1 \leq i \leq q, \text{ 즉, } x_i \text{와 } x_j \text{ 둘 다 질의점이다.} \\ S'_{ij} & \text{for } q+1 \leq i \leq m. \end{cases}$$

5. **for** $i = q+1$ to m **do** // 각 데이터 점에 대해 수행
6. **for** $j = 1$ to q **do** // 각 질의점에 대해 수행

7. $s_{ij} = s_{ij} + \sum_{k=q+1}^m S'_{ik} S'_{kj}$

8. **end for**

9. **end for**

10. q 개의 질의 점에 대한 x_i 의 유사도 s_i 를 다음과 같이 계산한다: $s_i = \max_{1 \leq j \leq q} \{s_{ij}\}$.

11. 집합 $SS = \{s_{q+1}, s_{q+2}, \dots, s_m\}$ 를 내림 차순으로 정렬하고, 최상위 k 개의 점을 k -최근점 질의의 결과로 반환한다.

단계 1에서, S_{ii} 를 0로 지정하였는데 그 이유는 다음과 같다: 한 점 x_i 의 질의점 x_q 에 대한 유사도 S_{iq} 를 계산하기 위해 공식 (2)를 전개해보자: $S_{iq} = S_{iq} \cdot S_{qq} + (S_{i1} \cdot S_{q1} + S_{i2} \cdot S_{q2} + \dots + S_{ii} \cdot S_{qi} + \dots + S_{im} \cdot S_{qm}) = S_{iq} + (S_{i1} \cdot S_{q1} + S_{i2} \cdot S_{q2} + \dots + S_{qi} + \dots + S_{im} \cdot S_{qm})$. 여기서 만약 S_{ii} 를 1로 지정한다면 유사도 S_{iq} 를 계산하는데 S_{iq} (= S_{qi})가 두 번 포함됨을 의미한다. 이것을 피하기 위해 자신과의 유사도 S_{ii} 를 0로 지정한다. 그러나 질의점 간의 유사도 S_{qq} 는 1로 지정한다.

두 순위 매김 알고리즘 Algorithm 1과 Algorithm 2의 복잡도를 비교해보면, Algorithm 2에서 단계 1-3은 Algorithm 1과 마찬가지로 실제 탐색 전에 미리 수행된다. Algorithm 1과 비교할 때, Algorithm 2의 추가 비용은 단계 7에서 $\sum_k S'_{ik} \cdot S'_{kj}$ 의 추가 계산에 있다. 이 추가 비용은 데이터 집합에 있는 데이터 수가 늘어나면 같이 증가하지만 유사도 S'_{ik} 와 S'_{kj} 가 미리 계산되어 유사도 행렬 S' 에 저장되어 있으므로 큰 부담이 되지는 않는다.

보기 4. 유사도 분포에 기초한 순위 매김 알고리즘의 효과를 보기 위해 74 개의 2 차원 점들로 구성된 데이터 집합을 사용하여 실험을 하였고, 그림 6에 그 결과를 나타내었다. 데이터 집합은 두 개의 클러스터 구조를 가지고 있고, 두 개의 질의 점은 '+'와 'x'로 표시하였다. 각 데이터 점 옆에 있는 숫자는 그 점에 할당된 순위이다. 그림 6에서 보는 것과 같이 본 알고리즘은 데이터 집합의 전체적 클러스터 구조를 반영하고 있다.

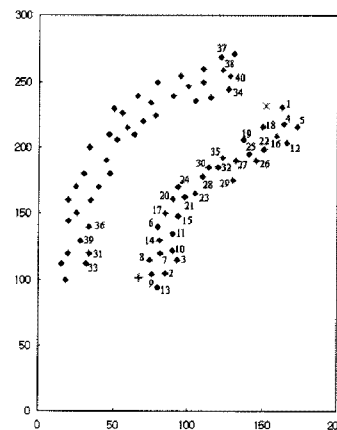


그림 6 두 개의 질의 점 '+'와 'x'에 대한 유사도 순위 매김

5. 성능 평가

본 논문의 접근법에 대한 실험적 평가를 위해 28 × 28 화소의 공간 해상도를 갖는 120,000 개의 필기체 숫자 이미지를 포함하는 MNIST 데이터베이스[18]를 활용하였다. MNIST 데이터베이스는 AT&T와 Bell 연구소 등에서 벤치마크 데이터로 많이 활용하며, 다수의 연구에서도 실험에 사용하는 데이터베이스이다. 각 이미지의 특성은 784(=28×28) 차원의 벡터로 표현된다. 본 실험에서는 120,000 개의 이미지 중에서 처음 6,000 개만 사용하였고, 주어진 질의 이미지와 가장 유사한 k 개의 이미지를 반환하는 k -최근접 질의를 수행하였다.

본 논문에서는 성능에 대한 객관적인 측정을 위해 질의의 개념을 '0', '1', ..., '9'의 10 개의 숫자 이미지 카테고리 중 하나로 설정하였고, k -최근접 질의의 정확도 (precision) 성능을 평가하였다. 여기서 k 는 10에서 100

까지를 사용하였고 정확도는 질의의 결과로 반환된 k 개의 이미지 중에서 질의 이미지 카테고리에 속하는 것의 비율로 계산하였다.

실험을 위해 MNIST 데이터베이스에서 무작위로 선정한 100 개의 이미지에 대한 k -최근접 질의를 수행하였고 그 평균 성능을 계산하였다. Algorithm2의 성능에 대한 직관을 제시하기 위해 그림 7-10에 질의처리 결과를 나타내었다.

그림 7과 8은 하나의 질의 이미지를 사용한 결과이고, 좌상단의 이미지가 질의 이미지이다. 그림 8의 유클리드 거리에 기반한 순위 매김의 결과에서는 질의 이미지의 카테고리인 '9'에 속하지 않는 많은 이미지가 도출된 반면에, 그림 7의 본 알고리즘에 기반한 순위 매김에서는 대부분이 질의 이미지 카테고리에 속하는 이미지들이다. 그림 9-10까지의 결과는 두 개의 이미지를 질의 이미지로 사용한 결과이다. 좌상단의 두 개의 이미지가 질의

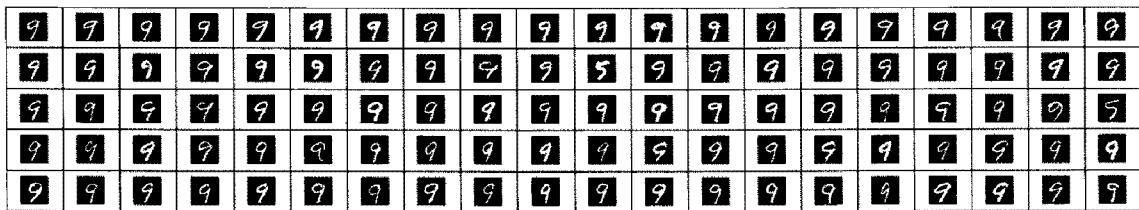


그림 7 본 알고리즘의 100-최근접 질의 처리 결과(좌상단이 질의 이미지)(정확도: 98%)

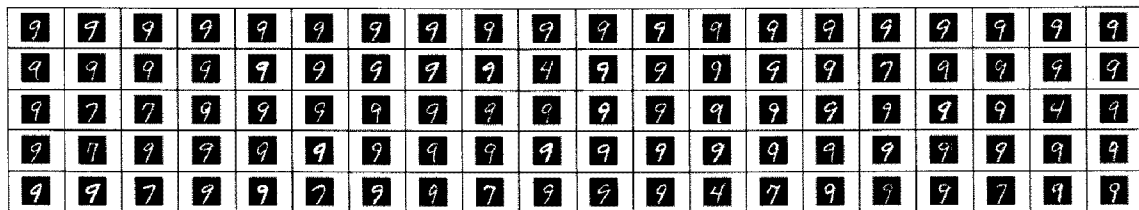


그림 8 유클리드 거리 기반 100-최근접 질의 처리 결과(좌상단이 질의 이미지)(정확도: 88%)

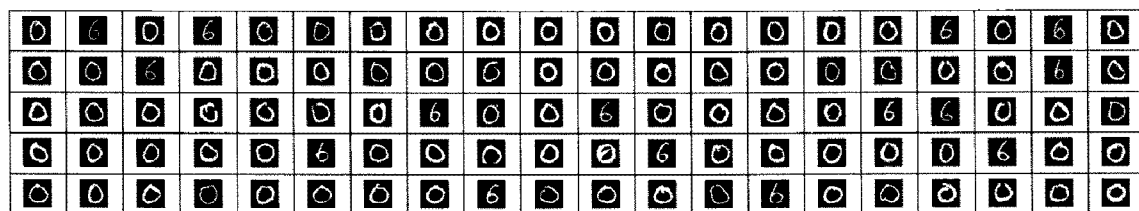


그림 9 본 알고리즘의 100-최근접 질의 처리 결과(좌상단 두 개가 질의 이미지)(정확도: 100%)

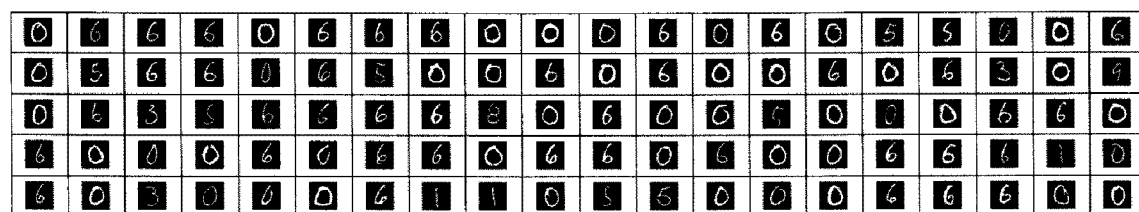


그림 10 Falcon의 100-최근접 질의 처리 결과(좌상단이 질의 이미지)(정확도: 84%)

이미지로 사용되었다. 다수의 질의 이미지를 사용하는 (즉, OR에 기반한 결합적(disjunctive)) 질의 처리를 위해서는 상수 $\alpha = -3$ 을 사용하는 Falcon[5]의 결합적 비유사도 측정법(aggregate dissimilarity measure)을 사용하였다. 그림 9,10에서 보면 Falcon의 결합적 비유사도 측정법을 사용하는 경우에는 질의 이미지의 카테고리리와 다른 이미지가 많이 도출되는 반면에, 본 논문의 방법은 균일한 결과를 산출한다. 이 실험 결과는 본 논문의 방법이 우수함을 간접적으로 증명하고 있다.

그림 11은 본 논문의 Algorithm 2, 유클리드 거리 기반 기법, SVM_{active} 기법[4], 그리고 적합성 피드백을 사용하지 않는 일반 지지 벡터 기계(support vector machines; SVM) 기반 기법 간의 k -최근접 질의의 정확도 성능을 비교하고 있다. [4]에서 SVM_{active}가 질의 정제 기법의 범주에 드는 질의 재가중치(reweighting) 방법인 MARS[3], 질의점 이동 방법인 MARS[2]와 MindReader[1] 뿐만 아니라 질의 확장 방식인 Falcon [5]보다 우수하다고 보고되었다. 따라서 SVM_{active}를 비교 대상으로 선정하였다. SVM_{active}는 지지 벡터 기계와 동적 학습(active learning)에 기반한 적합성 피드백 방법으로써 몇 번의 적합성 피드백 과정 후에 가장 유사하다고 판단하는 k 개의 이미지를 반환한다. 매번 적합성 피드백 과정 중에 SVM_{active}는 이미지가 질의 이미지와 같은 영역에 속하는지 결정한다. SVM_{active}의 실험에서 적합성 피드백 과정을 4 번 실행하였고, 매 과정마다 100 개의 훈련 이미지를 사용하였다. 그림 11은 이상의 네가지 기법에 대한 평균 최상위 k 개의 이미지에 대한 정확도를 나타낸다. SVM_{active}가 최악의 성능을 보이고 있다. SVM_{active} 학습기는 이미지 카테고리에 대한 어떠한 사전 정보도 갖고 있지 않으며 오직 적합성 피드백을 통해서만 질의 개념을 학습한다. 그러나 많은 경우에 고차원 공간에서 적은 수의 훈련 데이터만으로는 좋은 분류를 하기 어려우며, SVM_{active}의 경우 30% 정확도도 이루지 못하고 있다.

본 논문의 도메인에서 SVM의 분류 효과를 살펴보기 위해 SVM을 가지고 수행한 실험에서는 6,000 개의 이미지로 SVM을 훈련시킨 후에 하나의 질의 이미지와 두 개의 질의 이미지를 사용한 실험에서 각각 86.1%와 82.9%의 정확도를 얻었다. 그러나 SVM_{active}에서는 이러한 사전 훈련 과정 없이 동적으로 적은 수의 훈련 데이터를 통해 질의 개념을 습득하고자 하였다.

그림 11에서 보듯이 본 논문의 방법은 최상위 k 개의 이미지에 대한 정확도에서 90% 이상의 정확도를 나타낸다. 반면에 다른 방법들은 이와 같은 정확도를 이루지 못하고 있다.

그림 12는 다수의 질의 이미지를 사용한 경우의 정확

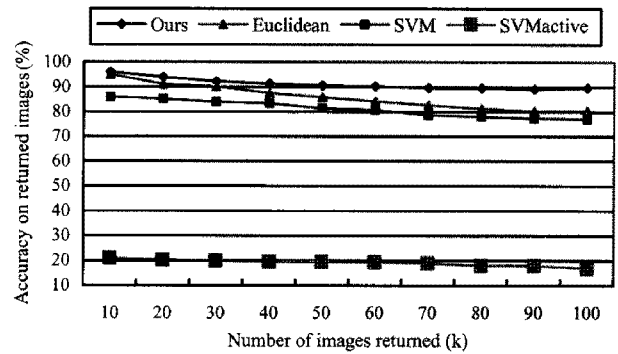


그림 11 하나의 이미지를 사용한 질의의 k 개 결과의 평균 정확도

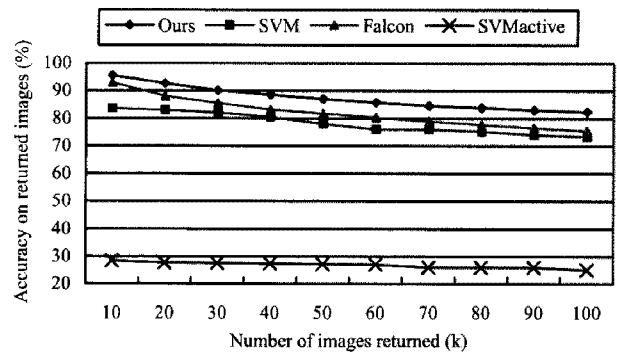


그림 12 다수의 이미지를 사용한 질의의 k 개 결과의 평균 정확도

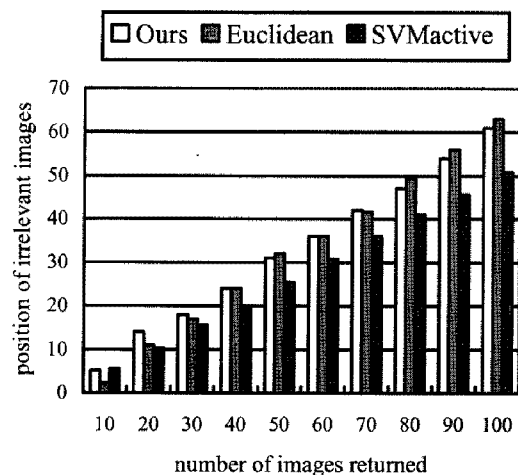


그림 13 k 개의 결과 이미지에서 관련 없는 이미지의 위치

도에 대한 실험 결과를 나타낸다. 본 논문의 방법은 어떠한 경우에도 80% 이상의 정확도를 나타내는 반면에 다른 방법들은 이 정도의 성능을 보이지 못하고 있다.

그림 13은 k 개의 결과 이미지에서 관련 없는 이미지의 평균 위치를 나타내고 있으며, 표 1에 그 수치를 표현하였다. 이 위치가 뒤에 위치할수록 바람직하다. 본

표 1 결과로 나온 k 개의 이미지에서 관련 없는 것들의 평균 위치

k	10	20	30	40
본 알고리즘	5.2	14.0	18.0	24.0
유클리드 거리 함수	2.3	11.0	17.0	24.0
SVM _{active}	5.5	10.4	15.6	20.0

논문의 방법에서는 k 의 크기가 작을 때 다른 방법에 비해 관련 없는 이미지의 위치가 훨씬 뒤에 위치함을 볼 수 있다. 많은 경우에 사용자들은 많은 수의 이미지를 원하지 않으므로(즉, k 의 크기가 대체로 작다) 이것은 바람직한 현상이다.

6. 결론

내용기반 이미지 검색을 위해 본 논문에서는 비선형 유사도 모델에 기반한 새로운 유사도 척도와 함께 새로운 문맥기반 순위 매김 알고리즘을 제시하였다. 이 새로운 유사도 척도와 순위 매김 알고리즘은 이미지 유사도를 평가하기 위해 데이터 집합 고유의 분포와 구조를 고려한다. 실험에 따르면 본 논문에서 제시한 알고리즘은 SVM_{active}, Falcon, 유클리드 거리기반 기법, SVM 분류기반 기법에 비해 우수한 성능을 보였다. 본 논문의 연구 동기는 내용기반 이미지 검색에서 유클리드 거리를 포함하여 민코프스키 측정법에 따른 거리가 인간이 인지하는 유사도와 항상 일치하지는 않는다는 관점에서 출발하였다.

참고 문헌

[1] Ishikawa, Y., Subramanya, R., and Faloutsos, C., "MindReader: Querying databases through multiple examples," *Proc. VLDB*, pp.218-227, 1998.

[2] Rui, Y., et al., "Relevance feedback: A Power tool for interactive content-based image retrieval," *IEEE Trans. Circuits and Video Technology*, vol.8, no.5, pp.644-644, 1998.

[3] Rui, Y., Huang, T.S., and Mehrotra, S., "Content-based image retrieval with relevance feedback in MARS," *Proc. ICIP*, pp.815-818, 1997.

[4] Tong, S. and Chang, E., "Support Vector Machine Active Learning for Image Retrieval," *Proc. ACM Multimedia*, pp.107-118, 2001.

[5] Wu, L., Faloutsos, C., Sycara, K., and Payne, T.R., "FALCON: Feedback Adaptive Loop for Content-Based Retrieval," *Proc. VLDB*, pp.297-306, 2000.

[6] Zhou, D. et al., "Learning with Local and Global Consistency," *Advances in Neural Information Processing Systems*, 16, MIT Press, Cambridge, MA, 2004.

[7] Haykin, S., *Neural Networks: A Comprehensive*

Foundation, Maxmillan, NY, 1994.

[8] Schölkopf, B. et al., "Comparing Support Vector Machines with Gaussian Kernels to Radial Basis Function Classifiers," *IEEE Trans. Signal Processing*, vol.45, pp.2758-2765, 1997.

[9] Vapnik, V.N., *Statistical Learning Theory*, Wiley, NY, 1998.

[10] Hoi, C.-H. and Lyu, M., "A novel log-based relevance feedback technique in content-based image retrieval," *Proc. ACM Multimedia*, pp.24-31, 2004.

[11] He, X., Ma, W.-Y., and Zhang, H.-J., "Learning an image manifold for retrieval," *Proc. ACM Multimedia*, pp.17-23, 2004.

[12] Wu, G., Chang, E.Y., and Panda, N., "Formulating Context-dependent Similarity Functions," *Proc. ACM Multimedia*, pp.725-734, 2005.

[13] Goh, K.-S., Li, B., and Chang, E., "DynDex: A Dynamic and Non-metric Space Indexer," *Proc. ACM Multimedia*, pp.466-475, 2002.

[14] Valois, R.L.De and Valois, K.K.De, *Spatial Vision*, Oxford Science Pub., Oxford, 1988.

[15] Shi, J. and Malik, J., "Normalized Cuts and Image Segmentation," *IEEE Trans. PAMI*, vol.22, no.8, pp.888-905, 2000.

[16] Schölkopf, B., Smola, A., and Müller, K., "Nonlinear Component Analysis as a Kernel Eigenvalue Problem," *Neural Computation*, vol.10, pp.1299-1319, 1998.

[17] Ng, A.Y., Jordan, M.I., and Weiss, Y., "On Spectral Clustering: Analysis and Algorithm," *Advances in Neural Information Processing Systems*, 14, MIT Press, MA, 2001.

[18] LeCun, Y. et al., "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol.86, no.11, pp.2278-2324, 1998.



차 광 호

1984년 부산대학교 계산통계학과 학사
 1989년 KAIST 전산학과 석사. 1997년
 KAIST 정보및통신공학과 박사. 1989년~
 1996년 데이콤 선임연구원. 1999년~
 2000년 IBM Almaden 연구소 방문과학
 자. 1997년~2002년 동명정보대학교 멀티
 미디어공학과 조교수. 2002년~2005년 숙명여자대학교 멀티
 미디어학과 교수. 2005년~현재 서울산업대학교 컴퓨터공학과 교수