

효율적인 내용 기반 이미지 검색을 위한 근사 Earth Mover's Distance

장 민 희* · 김 상 욱**

요 약

정확한 내용 기반 이미지 검색을 위하여 Earth mover's distance와 Optimal color composition distance와 같은 거리함수들이 제안되었다. 이 거리함수들은 정확도가 높은 검색 결과를 가져오지만 검색 시간이 매우 크기 때문에 대용량 데이터베이스에서 사용하기 어렵다는 문제점을 가지고 있다. 본 논문에서는 이러한 문제를 해결하기 위하여 선형 시간에 근사 Earth mover's distance를 구하기 위한 새로운 거리 함수를 제안한다. 제안하는 방법은 선형 시간에 두 이미지의 거리를 계산하기 위하여 공간 채움 곡선을 이용한다. 다양한 실험을 통하여 본 논문에서 제안하는 방법의 우수성을 검증한다. 실험 결과, 제안하는 기법이 Earth mover's distance에 비해 약 160배 정도의 검색 속도 향상 효과를 보이면서도 매우 유사한 결과를 검색하는 것으로 나타났다.

키워드 : 내용 기반 이미지 검색, Earth Mover's Distance, 거리 함수

Approximate Earth Mover's Distance for the Efficient Content-based Image Retrieval

Min-Hee Jang^{*} · Sang-Wook Kim^{**}

ABSTRACT

For content-based image retrieval, the earth mover's distance and the optimal color composition distance are proposed to measure the dissimilarity. Although providing good retrieval results, both methods are too time-consuming to be used in a large image database. To solve the problem, we propose a new distance function that calculates an approximate earth mover's distance in linear time. To calculate the dissimilarity in linear time, the proposed approach employs the space-filling curve. We have performed extensive experiments to show the effectiveness and efficiency of the proposed approach. The results reveal that our approach achieves almost the same results with the EMD in linear time.

Keywords : Content-based Image Retrieval, Earth Mover's Distance, Distance Function

1. 서 론

내용 기반 이미지 검색 (Content-based image retrieval, CBIR)은 이미지를 질의로 던져 그 이미지의 색 분포, 형태, 그리고 질감 등을 분석하여 유사한 이미지를 검색하는 기술이다[1]. 본 논문에서는 CBIR의 색 기반 검색에 대해 논의한다.

정확한 색 기반 검색을 수행하기 위해서는 이미지를 구성하는 색들의 분포를 어떻게 표현할 것인가가 매우 중요하다. 이를 위해 주요색 표현 방법(dominant color descriptor, DCD)이 제안되었다[1][2]. DCD는 색의 양자화(color quantization)를 통해 유사한 색 간의 통합을 수행하여 이미지를 구성하는 지배적인 몇몇 색과 그 색이 이미지에서 차지하는 비율로 이미지를 표현한다. 여러 논문의 실험 결과, DCD를 이용한 이미지 검색이 전통적인 색 표현 방법인 히스토그램 방법보다 뛰어난 검색 결과를 보이는 것으로 나타났다[1].

DCD를 기반으로 주어진 이미지와 유사한 이미지를 찾기 위해서는 이미지간의 거리를 측정할 수 있는 거리 함수가 필요하다. 이러한 거리 함수를 세우기 위해서는 다음과 같

* 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초 연구사업 지원을 받아 수행된 것입니다(2008-0061006). 또한, 지식경제부 및 정보통신 산업진흥원의 IT융합 고급인력과정 지원사업(NIPA-2011-C6150-1101-0001)과 두뇌한국 21 사업의 부분적인 지원을 받아 수행됨.

† 준 회 원 : 한양대학교 전자컴퓨터통신공학과 박사과정

** 종신회원 : 한양대학교 정보통신대학 정보통신학부 교수

논문접수 : 2011년 6월 15일

수정일 : 1차 2011년 9월 8일

심사완료 : 2011년 9월 15일

은 두 가지 조건이 고려되어야 한다[2]. 첫째, 두 이미지 내에서 대응되는 색이 얼마나 유사한가. 둘째, 대응되는 색의 비율은 얼마나 유사한가. 위의 두 조건을 만족하는 거리함수로서 비교하고자 하는 두 이미지 중 한 이미지의 색-비율 분포를 다른 이미지의 분포로 옮기는데 들어가는 비용인 minimum work를 사용할 수 있다[1]. work는 한 이미지에서 다른 이미지로 옮겨진 색들의 양과 옮겨진 거리의 곱을 의미한다. 이러한 minimum work를 구하는 거리함수를 Earth mover's distance (EMD)라고 한다[1].

그러나 한 이미지의 색-비율 분포를 다른 분포로 옮길 수 있는 경우의 수가 매우 많기 때문에 minimum work를 계산하기 위해서는 큰 계산량이 요구된다. n 개의 색이 있을 때 EMD는 $O(n^3 \log n)$ 의 복잡도를 가진다. 대용량 데이터베이스에서는 이미지 비교 횟수가 많이 요구된다. 각 이미지 비교마다 위와 같은 계산복잡도가 필요하다면 그 거리함수는 대용량 데이터베이스에서 사용하기 어렵다.

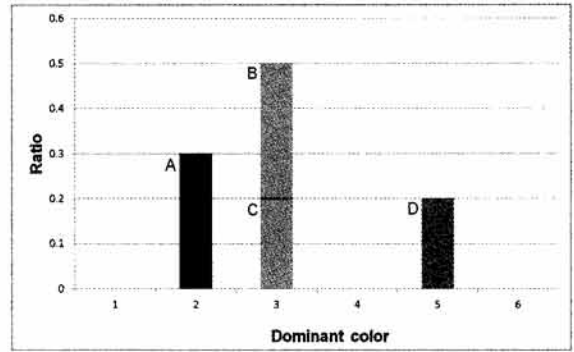
본 논문에서는 선형 시간(linear time)에 근사 EMD를 계산하기 위한 거리 함수에 대해 논의한다. 기존의 연구에서도 근사 EMD에 대해 연구된 적이 있다. Wavelet EMD는 두 이미지의 근사 EMD를 linear time에 계산하는 연구이지만 히스토그램 기반 색 표현 방법을 위한 것으로서 DCD에서는 사용하기 어렵다[3].

본 논문에서 제안하는 방법은 먼저 비교하고자 하는 두 이미지의 색들을 1차원으로 표현하여 순서화한다. 그 후 1차원의 시작점에서 가장 가까운 두 색을 골라 매치시키는 과정을 앞에서부터 순서대로 수행함으로써 1차원 공간 상 두 이미지의 minimum work를 선형 시간에 계산한다. 그러나 이미지의 색 공간은 RGB나 HSV처럼 3차원이기 때문에 이러한 3차원 공간을 1차원과 같이 순서화할 수 있는 방법이 요구된다. 본 논문에서는 공간 채움 곡선(space-filling curve)을 이용하여 3차원 공간상의 두 이미지의 색들을 순서화한다. 공간 채움 곡선은 k 차원 공간을 1차원으로 사상(mapping)해주는 기법이다[4].

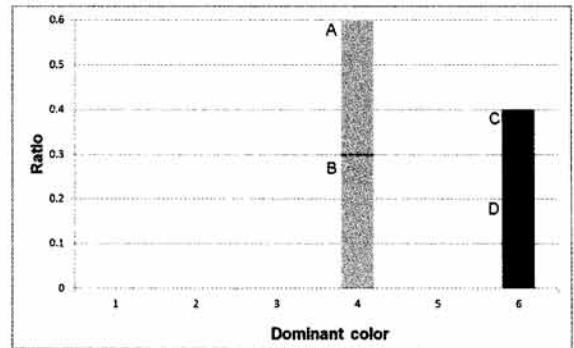
제안하는 기법은 공간 채움 곡선을 이용하여 3차원 공간의 색들을 순서화한 후 시작점에서 가장 가까운 두 색을 선택해 3차원 실제 공간에서의 거리를 비교하는 것을 반복한다. 선수에 따라 색들을 한번만 스캔하면 되기 때문에 $O(n)$ 의 계산 복잡도에 두 이미지의 거리를 구할 수 있다. 이에 더해 본 논문에서는 좀 더 정확한 근사 EMD를 구하기 위해 공간 채움 곡선의 시작점을 달리하여 다양한 곡선을 생성한다. 그 후 각 곡선을 통해 만들어진 색들의 순서에 따라 두 이미지의 실제 거리를 계산한 후 그 중 최솟값을 두 이미지의 거리로 사용한다. 본 논문에서는 다양한 실험을 이용한 성능 분석을 통하여 제안하는 기법의 우수성을 검증한다.

2. 관련 연구

본 장에서는 먼저 minimum work의 개념에 대해 자세히 설명한 후 minimum work를 구하기 위한 기존 연구들에 대



(a) 이미지 P의 색들



(b) 이미지 Q의 색들

(그림 1) 두 이미지의 minimum work 예제

해 논의한다. 두 이미지 $P=\{(p_1, w_{p1}), \dots, (p_m, w_{pm})\}$, $Q=\{(q_1, w_{q1}), \dots, (q_l, w_{ql})\}$ 가 있을 때 p_i 와 q_j 는 이미지의 색을, w_{pi} 와 w_{qj} 는 각 색의 비율을 의미한다. P 와 Q 의 비율의 총합은 동일하다고 가정한다. 이 때, P 와 Q 의 거리는 minimum work를 계산함으로써 측정될 수 있다. minimum work는 P 의 색-비율 분포를 Q 의 분포로 옮기는데 들어가는 최소 work의 양을 의미한다. work는 한 이미지에서 다른 이미지로 옮겨진 색들의 양 f 와 ground distance d 의 곱으로 구할 수 있다. ground distance란 색들 간의 거리를 측정하기 위해 사용된 Euclidean distance나 L_1 distance와 같은 기본 거리를 의미한다[2]. work를 공식으로 표현하면 다음과 같다.

$$WORK(P, Q, F) = \sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij}$$

$F=[f_{ij}]$ 는 p_i 에서 q_j 로 옮겨진 색의 양을 의미하고 $D=[d_{ij}]$ 는 옮겨진 색의 거리를 의미한다. (그림 1)은 minimum work를 계산하는 예제이다. (그림 1)에서 이미지 P 는 3개의 색, Q 는 2개의 색을 가지고 있다. 그 중 이미지 P 의 색들을 Q 의 같은 알파벳으로 옮기면 두 이미지의 minimum work를 구할 수 있다: $0.3*2$ (A 부분) + $0.3*1$ (B 부분) + $0.2*3$ (C 부분) + $0.2 * 1$ (D 부분). 이처럼 정확한 minimum work를 구하는 거리 함수를 Earth mover's distance (EMD)라고 한다[1]. 그러나 EMD를 구하기 위해서는 매우 큰 계산량이 요구되기 때문에 이러한 계산량을 줄이기 위한 여러 연구가 진행되었다.

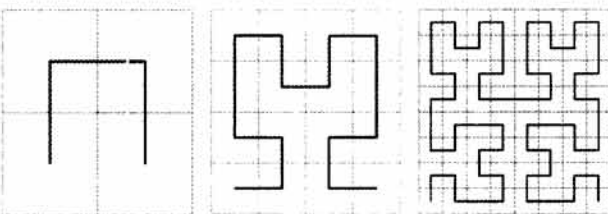
Optimal color composition distance (OCCD)는 EMD의 계산복잡도를 줄이기 위한 근사 기법이다[3]. OCCD는 색들의 비율을 m 개의 color unit으로 나눈다. 각각의 color unit들은 모두 같은 영역 비율 p 를 가지고 있으며, $m \times p=1$ 이다. OCCD는 이 color unit들을 m 개의 node를 가지는 비방향 그래프로 가정하고 minimum cost graph matching이라는 방법을 활용하여 두 이미지의 근사 minimum work를 구한다. OCCD는 $O(m^3)$ 의 계산 복잡도를 가진다. OCCD 논문에서는 color unit의 개수를 20개로 설정한다[3].

EMD 논문에서는 minimum work를 계산하기 위해 transportation problem을 이용한다[1]. transportation problem은 공급자들이 가지고 있는 상품을 소비자들에게 공급하기 위해 들어가는 최소 비용을 계산하는 문제이다. EMD는 한 이미지의 색 분포를 공급자들로 다른 한 이미지의 색 분포를 소비자들로 가정하고 minimum work를 계산한다. EMD는 $O(n^3 \log n)$ 의 복잡도를 가진다. 이와 같이 EMD와 OCCD의 계산복잡도가 매우 크기 때문에 이미지 검색 시 매우 많은 검색 시간을 요구한다.

3. 제안하는 기법

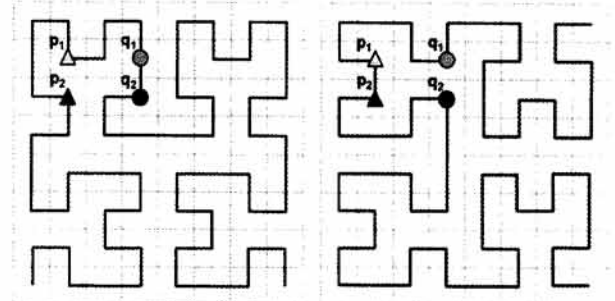
본 장에서는 선형 시간에 두 이미지의 근사 EMD를 구하기 위한 거리 함수에 대하여 논의한다. 1차원 색 공간에서는 greedy 알고리즘을 적용하면 n 개의 색에 대하여 $O(n)$ 에 두 이미지의 minimum work를 쉽게 구할 수 있다. 2장의 (그림 1)은 1차원 색 공간상 두 이미지 P 와 Q 의 색-비율 분포를 나타낸다. (그림 1)에서 두 이미지의 minimum work를 구하기 위해서는 시작점에서 가장 가까운 두 색을 선택하여 최대 공통비율만큼의 거리를 비교해야 한다.

최대 공통비율이란, 비교하고자 하는 두 색들에서 비율이 작은 색의 비율을 의미한다. 예를 들어 (그림 1)의 이미지 P 에서 시작점에 가장 가까운 색의 비율은 0.3이고 Q 에서 시작점에 가장 가까운 색의 비율은 0.6이다. 이 두 색 중 비율이 작은 색의 비율이 0.3이기 때문에 이 두 색의 최대 공통비율은 0.3이다. 비교된 최대 공통비율 부분을 제거한 후 다시 시작점에서 가장 가까운 두 색을 선택하여 최대 공통비율만큼의 거리를 비교한다. 이 과정을 앞에서부터 순서대로 모든 색에 대해 수행한 후 그 결과들을 모두 합하면 1차원 공간 상 두 이미지의 minimum work를 구할 수 있다.



(a) level 1 (b) level 2 (c) level 3

(그림 2) 힐버트 커브의 예제



(a) 좌하점 시작 HC (b) 우상점 시작 HC

(그림 3) 2차원 공간의 힐버트 커브들

본 논문에서는 이러한 방법을 이미지 검색에 적용하고자 한다. 그러나 색은 RGB나 HSV와 같은 3차원 공간의 데이터이기 때문에 3차원 데이터를 1차원으로 순서화할 수 있는 방법이 요구된다. 본 논문에서는 다차원 공간을 1차원의 곡선으로 표현하는 기법인 공간 채움 곡선 방법 중 힐버트 커브를 이용하여 다차원 공간의 데이터를 1차원으로 표현한다[4]. (그림 2)는 2차원 공간의 힐버트 커브들이다.

(그림 2)에서 보는 바와 같이 힐버트 커브는 다차원의 공간을 1차원의 곡선으로 채운다. 이 곡선에 따라 다차원 데이터들을 순서화할 수 있다. 다차원 공간 상 각 입자들의 근접성을 반영하여 순서를 생성하기 때문에 힐버트 커브를 이용하면 실제 minimum work에 근접한 값을 구할 수 있다. 이미지 공간은 RGB나 HSV같은 3차원 공간이기 때문에 본 논문에서는 3차원 힐버트 커브를 사용하여 이미지의 색들을 순서화한다. 그 후, 순서화 된 색들을 대상으로 색들의 거리를 비교하는데 각 색들의 거리 비교 시에는 실제 다차원 상 거리를 이용한다. 이를 통해 3차원 공간 상 두 이미지의 근사 EMD를 $O(n)$ 에 계산할 수 있다.

그러나 힐버트 커브를 통해 생성된 순서가 다차원 공간 상 색들의 근접성을 정확히 반영하기는 어렵다. 실제 다차원 공간에서는 가까이 있는 두 색들이 힐버트 커브를 통해 멀어지는 경우가 발생한다. 이 결과 색들의 순서가 엉뚱하게 배열되어 EMD의 결과와 오차를 발생할 수 있다.

(그림 3)은 2차원 색 공간에 힐버트 커브를 적용한 예이다. (그림 3)에서 이미지 $P=\{p_1, p_2\}$, 이미지 $Q=\{q_1, q_2\}$ 이며 모든 색의 비율은 동일하다. 이미지 P 와 Q 의 minimum work를 만족하기 위해서는 p_1 은 q_1 과 p_2 은 q_2 와 비교해야 한다. 그러나 (그림 3(a))의 힐버트 커브를 통해 색들을 순서화한 후 그 순서대로 비교하면 p_1 은 q_2 와 p_2 은 q_1 과 비교하게 되어 minimum work를 만족하지 못한다. 이처럼 힐버트 커브가 다차원 공간의 상대적인 근접성을 만족하지 못하는 경우 색들의 순서가 엉뚱하게 배열된다. 그러나 하나의 다차원 공간에는 시작점을 달리하여 다양한 힐버트 커브들을 생성할 수 있다. (그림 3(b))와 같이 힐버트 커브의 시작점을 달리하여 다른 힐버트 커브를 그리면 (그림 3(a))의 힐버트 커브와는 다른 순서가 그려져 p_1 은 q_1 과 p_2 은 q_2 와 비교되어 두 이미지의 minimum work를 계산할 수 있다. 본

논문에서는 k 차원 공간의 각 모서리 (corner)를 시작점으로 가지는 2^k 개의 힐버트 커브를 생성한다. 그 후 각 힐버트 커브를 통해 만들어진 순서에 따라 두 이미지의 실제 거리를 계산한 후 그 중 최솟값을 두 이미지의 거리로 사용한다.

4. 실험

4.1 실험 환경

제한한 기법의 성능 평가를 위해 다양한 실험을 수행하였다. 실험에 사용된 데이터 집합은 SIMPLicity이다. SIMPLicity는 1,000개의 이미지로 구성되어 있는 이미지 데이터로 한 개의 카테고리에 100개의 이미지씩 10개의 카테고리로 나누어져 있다[5].

SIMPLicity에서 색들을 추출하기 위하여 $32 \times 32 \times 32$ 의 RGB 공간에서 MPEG7 DCD를 이용하였다[5]. MPEG7 DCD는 이미지에 따라 최대 8개의 색들을 자동적으로 추출한다. 제안한 기법의 비교대상은 EMD와 OCCD(20)이다. OCCD(20)은 [2]에서 추천한 바와 같이 color unit의 개수를 20개로 설정한 것이다. 정확도의 척도로 mean absolute percentage error(MAPE)를 사용하였다.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right|$$

A_i 는 두 이미지의 실제 EMD, F_i 는 제안한 기법 또는 OCCD로 구한 근사 EMD이다. 성능의 척도로는 파일 접근

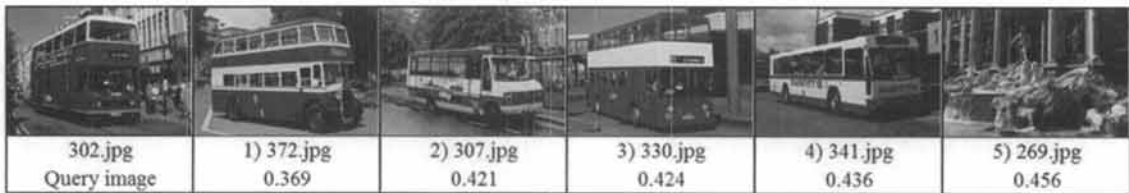
시간을 제외한 검색 시간을 이용하였다. 정확도와 성능의 결과로 1000개 이미지와 유사도를 계산하는 임의의 질의 100개를 던진 결과의 평균값을 사용하였다. 모든 실험들은 윈도우즈7에서 2GH의 메모리 크기와 Intel Core 2.80GHZ PC를 이용하여 수행되었다.

4.2 실험 결과

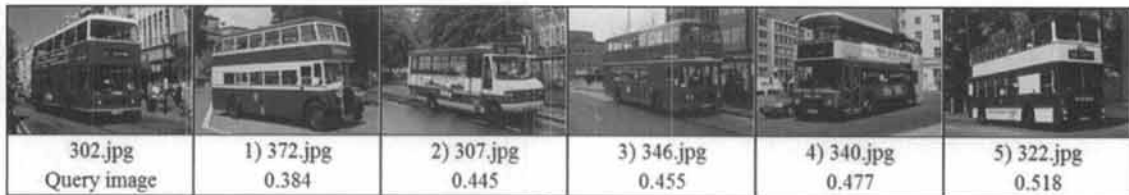
<표 1>은 EMD, OCCD 그리고 제안한 기법 (MHC)의 정확도와 성능을 비교한 것이다. 표에서 보듯이 제안한 기법은 작은 오차만을 발생하면서 검색 시간은 크게 향상되었다. EMD에 비하여 8.7%의 오차가 발생하며 검색 시간은 약 160배 정도로 향상되었다. 즉, 제안한 기법은 매우 빠른 시간 안에 EMD와 유사한 minimum work를 계산할 수 있다. OCCD(20)은 제안한 기법보다 낮은 오차를 보이지만 검색 시간이 EMD와 유사하다. OCCD는 color unit의 비율과 실제 색들의 비율이 정확히 맞지 않기 때문에 오차가 발생한다. 또한 비교하고자 하는 두 이미지의 color unit들 중 minimum work를 만족하는 color unit의 쌍을 찾는 데 시간이 걸리기 때문에 검색 시간이 크다.

<표 1> 검색 정확도 및 성능

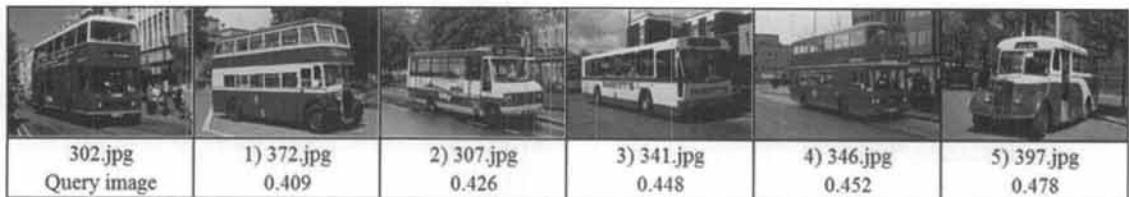
	MHC	EMD	OCCD(20)
MAPE (%)	8.7	-	4.3
검색시간 (ms)	4.7	789.3	684.9



(a) EMD



(b) MHC



(c) OCCD(20)
(그림 4) 검색 예제

<표 2> Precision과 recall
(a) precision

	EMD	MHC	OCCD(20)
10	86.8%	83.8%	85.4%
20	74.5%	72.0%	75.5%
30	65.8%	64.1%	65.4%

(b) recall

	EMD	MHC	OCCD(20)
10	8.7%	8.4%	8.5%
20	14.9%	14.4%	15.1%
30	19.9%	19.0%	19.7%

다음 실험은 이러한 오차가 실제 검색 결과에 어떠한 영향을 미치는지 알기 위해 SIMPLicity를 대상으로 정확도(precision)과 재현율(recall)을 측정하였다. 10, 20, 그리고 30개의 이미지를 검색하는 임의의 질의들을 던져 각 질의에 대해 정확도와 재현율을 구한 후 그 평균을 사용하였다. <표 2>에서 보는 바와 같이 모든 거리 함수가 유사한 결과를 도출하였다. 발생하는 오차가 적기 때문에 이 오차가 실제 검색에 미치는 영향 또한 적다. 즉, 제한한 기법은 검색 시간을 크게 줄이면서도 EMD와 거의 동일한 검색 결과를 가져온다.

(그림 4)는 각 거리 함수에 대한 실제 검색 결과의 예제이다. SIMPLicity의 버스 이미지 302.jpg를 질의로 던졌을 때 검색된 8개의 최근접 유사 이미지로서 이미지 이름 밑에 명시되어 있는 숫자는 질의 이미지와의 거리를 의미한다. 그림에서 보는 바와 같이 모든 거리 함수가 질의와 유사한 버스 이미지를 검색하는 것으로 나타났다.

5. 결 론

본 논문에서는 선형 시간에 EMD를 계산하기 위한 새로운 거리함수를 제안하였다. 제한한 거리함수는 힐버트 커브를 기반으로 색들의 순서를 생성함으로써 선형 시간에 두 이미지의 근사 minimum work를 계산할 수 있다. 실험 결과, 제한한 기법은 EMD에 비해 약 160배의 검색 속도 향상 효과를 보이면서도 유사한 검색 결과를 가져오는 것으로 나타났다.

참 고 문 헌

[1] Y.Rubner, C. Tomasi, and J. Guibas, "The Earth Mover's Distance as a Metric for Image Retrieval," *International*

Journal of Computer Vision, Vol.40, No.2, pp.99-121, 2000.
 [2] A. Mojsilovic, J. Hu, and E. Soljanin, "Extraction of Perceptually Important Colors and Similarity Measurement for Image Matching, Retrieval, and Analysis," *IEEE Transactions on Image Processing*, Vol.11, No.11, pp.1238-1248, 2002.
 [3] S. Shirdhonkar and D. Jacobs, "Approximate Earth Mover's Distance in Linear Time," In *Proc. IEEE Int'l Conf. on Computer Vision and Pattern Recognition, CVPR*, 2008.
 [4] H. Jagadish, "Analysis of the Hilbert Curve for Representing Two-Dimensional Space," *Information Processing Letters*, Vol.62, No.1, pp.17-22, 1997.
 [5] J. Wang, J. Li, and G.Wiederhold, "SIMPLicity: Semantics-Sensitive Integrated Matching for Picture Libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.23, No.9, pp.947-963, 2001.



장 민 희

e-mail : zzmini@agape.hanyang.ac.kr
 2003년 2월 홍익대학교 신소재공학과(학사)
 2006년 8월 한양대학교
 전자컴퓨터통신공학과(석사)
 2010년 3월~2010년 4월 미국 Carnegie Mellon University, Computer Science Department, 방문 연구원
 2006년 9월~현 제 한양대학교 전자컴퓨터통신공학 박사과정
 관심분야: 멀티미디어 데이터 검색, 사회연결망분석, 인터넷 포털 데이터 분석, 데이터 마이닝, 공간 데이터베이스/GIS



김 상 욱

e-mail : wook@agape.hanyang.ac.kr
1989년 서울대학교 컴퓨터공학과(학사)
1991년 한국과학기술원 전산학과(석사)
1994년 한국과학기술원 전산학과(박사)
1991년 7월~1991년 8월 미국 Stanford
University, Computer Science

Department, 방문 연구원

1994년 3월~1995년 2월 KAIST 정보전자연구소 전문 연구원
1999년 8월~2000년 8월 미국 IBM T.J. Watson Research
Center, Post-Doc.
1995년 3월~2003년 2월 강원대학교 정보통신공학과 부교수
2009년 1월~2010년 2월 미국 Carnegie Mellon University,
Visiting Scholar
2003년 3월~현 재 한양대학교 정보통신대학 정보통신학부 교수
관심분야: 데이터베이스 시스템, 저장 시스템, 트랜잭션 관리,
데이터 마이닝, 멀티미디어 정보 검색, 공간 데이터베
이스/GIS, 주기억장치 데이터베이스, 이동 객체 데이
터베이스/텔레매틱스, 사회 연결망 분석, 웹 데이터
분석