
지역 밀집도 및 Hausdorff 거리를 이용한 영상기반 텍스트 매칭

Image Based Text Matching Using Local Crowdedness and Hausdorff Distance

손화정, 김수형, 김지수, 박미선, 유재명
전남대학교 전산학과

Hwa-Jeong Son(sonhj@iip.chonnam.ac.kr), Soo-Hyung Kim(shkim@chonnam.ac.kr),
Ji-Soo Kim(kimjisoo@iip.chonnam.ac.kr), Mi-Seon Park(sunnydix@hanmail.net),
Jae-Myeong Yoo(jmyoo@oracle.chonnam.ac.kr)

요약

본 논문에서는 영상의 유사성을 측정하는데 많이 이용되는 Hausdorff 거리 기법이 텍스트 영상을 검색하는 분야에도 효과적임을 입증하고자 한다. 즉, 시차를 두고 스캔된 임의의 텍스트 영상들의 동일성 여부를 판단할 수 있는 영상기반 텍스트 매칭 기법을 제안하고 이를 위해 지역 밀집도와 Hausdorff 거리를 이용한다. Hausdorff 거리 방법은 처리시간이 오래 걸리는 단점이 존재하는데, 본 논문에서는 지역 밀집도 알고리즘을 이용한 특징점 추출을 수행하여 이를 보완하였다. 우편 봉투에서 얻은 텍스트 영상으로 190개의 동일 영상 190개의 비동일 영상을 만들어 실험을 수행하였다. 기존에 영상 간의 유사도 매칭에 가장 일반적으로 이용되는 이진 상관도 및 Hausdorff 거리 방법과 본 논문에서 제안한 수정된 Hausdorff 방법의 실험 결과를 비교한 결과, 유사한 영역을 찾고 일치하는 정도를 얻는데 있어 다른 방법에 비해 약 2.7%에서 9.0%의 높은 정확률을 얻어 성능의 우수성을 입증하였다.

■ 중심어 : | 텍스트 영상 검색 | 영상-대-영상 매칭 | 전자 도서관 |

Abstract

In this paper, we investigate a Hausdorff distance, which is used for the measurement of image similarity, to see whether it is also effective for document retrieval. The proposed method uses a local crowdedness and a Hausdorff distance to locate text images by determining whether a pair of images scanned at different time comes from the same text or not. To reduce the processing time, which is one of the disadvantages of a Hausdorff distance algorithm, we adopt a local crowdedness for feature point extraction. We apply the proposed method to 190 pairs of the same class and 190 pairs of the different class collected from postal envelop images. The results show that the modified Hausdorff distance proposed in this paper performed well in locating the text region and calculating the degree of similarity between two images. An improvement of accuracy by 2.7% and 9.0% has been obtained, compared to a binary correlation method and the original Hausdorff distance method, respectively.

■ keyword : | Text Image Retrieval | Image-to-Image Matching | Digital Library |

I. 서론

영상과 영상을 비교하여 유사 여부를 결정하는 문제는 패턴 인식(pattern recognition) 및 컴퓨터 비전(computer vision), 객체 검출(object detection) 등 자동화 응용 분야에서 필수적인 역할을 차지하고 있다[1-3]. 또한, 전체 영상 또는 부분영상을 질의(query)로 하여 유사 영상을 매칭하는 영상 검색(image retrieval) 기법들도 현재 다양하게 연구되고 있다[4-6]. 그러나 이러한 연구의 대부분은 자연 영상을 대상으로 하고 있으며, 문서나 텍스트 영상에서는 적용하기 힘든 색상 정보를 기본적인 특징으로 이용한다. 본 논문에서는 자연 영상이 아닌 텍스트 영상에 효과적인 부분영상 매칭 기법을 제안하고자 한다.

두 텍스트 영상의 유사성을 측정하기 위한 방법은 인식기반 매칭 방법과 영상기반 매칭 방법으로 나누어 볼 수 있다. 인식기반 매칭 방법은 각 텍스트 영상의 내용을 인식한 후 인식한 내용을 기반으로 유사 여부를 판단하는 방법으로 인식을 위해 문자단위 분할 및 문자별 모델 생성 과정이 요구된다. 이와는 달리 영상기반 매칭 방법은 영상에서 다른 영상과 구별되는 특징을 추출하여 매칭을 시도하는 방법으로 인식기반 매칭 방법에서의 문자단위 분할 및 문자별 모델 생성 과정을 필요로 하지 않는다. 본 논문에서는 시차를 두고 스캔된 임의의 문서 영상들을 동일하다고 판단해 줄 수 있는 방법을 제안하였다. 이때 영상들은 동일 문서를 스캔하여 얻어졌을지라도 스캔 과정에서의 훼손, 잡영 및 기울어짐으로 차이가 발생할 수 있다. 그러나 이들 영상은 동일 문서로부터 얻어졌으므로 글자의 글꼴 및 크기, 문단의 모양 등이 모두 같다. 따라서 인식기반 매칭 방법보다는 영상기반 매칭 방법을 통하여 이들 영상의 일치성 여부를 판단함이 옳다.

텍스트 영상의 매칭에 관한 연구는 대부분 문서 영상 검색(document image retrieval) 분야에서 연구되고 있다[7-10]. Peng[7][8]은 문서 영상을 블록 단위로 분할하여 리스트(list)와 트리(tree)구조를 형성하고 그 블록의 크기나 위치 등을 특징으로 선택하여 검색을 수행하였다. 그러나 이 방법은 문서 영상의 전체적인 형태만을

살피기 때문에, 우편물이나 카드 전표 등과 같이 정해진 형식에 의해 작성된 문서의 경우에는 블록의 크기나 위치가 거의 유사하게 되어 이 연구에서 제안한 방법을 적용하기 힘들다. 또한 한글 문서에 대해 영상기반 매칭 방법을 적용한 예로 Oh[9]와 진영범[10]의 연구를 들 수 있다. 이 연구에서는 문서 내의 각 단어에 대한 4방향 프로파일 정보와 웨이블릿 특징을 이용하여 동일한 문서 영상을 검색하는 방법을 제안하였다. 그러나 문서 내에 있는 단어들을 낱자 단위로 분할하여 이들 분할된 영상에 대한 특징들을 이용하기 때문에, 만일 낱자 단위 분할 과정에서 오류가 발생하게 되면 이를 통해 추출된 특징도 실제 문자에 대한 특징과 일치하지 않게 된다. 따라서 낱자 단위 분할 과정을 생략하고 단어 단위 영상 매칭을 수행하기 위한 방법으로 본 논문에서 제안한 부분영상 매칭 방법을 적용하면 검색 오류나 수행 시간을 보다 효과적으로 줄일 수 있다.

본 논문에서는 지역 밀집도[11]와 Hausdorff 거리[2]를 특징 추출과 유사도 측정에 적용한 텍스트 영상 매칭 시스템을 제안한다. Hausdorff 거리를 이용한 텍스트 영상 매칭 방법은 일반적으로 두 영상간의 일치 영역을 정확하게 찾을 수 있지만 처리하는데 시간이 오래 걸리는 단점이 있다. 이러한 단점을 보완하기 위해, 지역 밀집도 알고리즘을 적용하여 영상으로부터 특징점을 추출한다. 지역 밀집도를 이용한 특징 추출과 Hausdorff 거리를 이용한 두 영상간의 유사도 측정을 통해, 텍스트 영상 매칭에 대한 정확도는 유지하면서 처리 시간을 효과적으로 줄일 수 있다. 제안 방법의 성능을 평가하기 위해, 이를 영상 매칭에 가장 일반적으로 적용되는 이진 상관(binary correlation) 함수를 이용한 방법 및 Hausdorff 거리만을 이용한 방법과 비교한다.

본 논문은 다음과 같이 구성된다. 2장에서는 텍스트 영상 매칭을 수행하기 위한 방법으로 일반적으로 적용하는 이진 상관 함수, Hausdorff 거리, 그리고 지역 밀집도에 대해 살펴본다. 3장에서는 실험을 통해 제안한 방법과 상관도 및 Hausdorff 거리 방법의 성능을 비교한다. 마지막 4장에서는 결론 및 향후 연구 방향에 대해 기술한다.

II. 텍스트 영상 매칭 시스템

두 텍스트 영상간의 유사도를 측정하기 위한 방법은 다음과 같이 정의된다. 즉, 질의 영상 Q 와 문서 영상 D 를 이진 영상으로 가정할 때 텍스트 영상 매칭은 Q 와 유사하다고 판단되는 D 의 부분영상 Q' ($Q' \subseteq D$)을 찾는 것을 말한다. 이는 어떤 특징 추출 방법과 유사도 측정 방법을 적용하느냐에 따라 다르며, 본 논문에서는 지역 밀집도와 Hausdorff 거리 방법을 특징 추출과 유사도 측정 방법에 적용하였다.

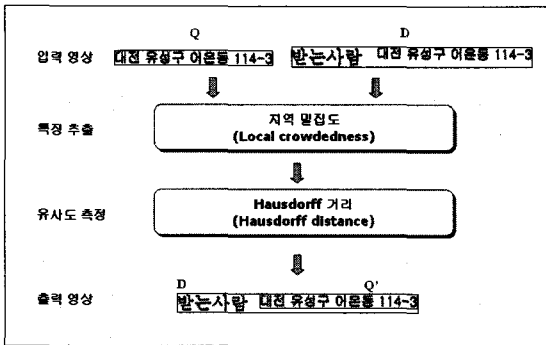


그림 1. 텍스트 영상 매칭 시스템 구성도

[그림 1] 본 논문에서의 제안 방법을 나타낸다. 또한 이를 가장 기본적인 영상의 유사도 측정 방법인 상관 함수 또는 Hausdorff 거리 방법을 이용한 알고리즘과 비교한다.

다음 1절과 2절에서는 이진 상관 함수 및 Hausdorff 거리 방법을 언급하고, 마지막 3절에서는 본 논문에서 적용한 특징 추출 방법인 지역 밀집도에 대해 서술한다.

1. 이진 상관(Binary Correlation, BC)

이진 상관 함수[12]는 두 영상간의 상관 정도를 나타내는 가장 기본적인 방법으로 일반적으로 원형 정합(template matching) 분야에 주로 이용되고 있다. 이 매칭 기법은 입력된 질의 영상과 가장 잘 일치하는 문서 영상의 위치를 찾는 방법으로, 한 픽셀씩 이동해가면서 계산된 상관 계수 중 가장 큰 값을 갖는 지점을 선택한다.

$X \times Y$ 의 크기를 갖는 문서 영상 D 와 $M \times N$ 의 크기를 갖는 질의 영상 Q 간의 상관도를 계산하는 이진 상관 함수는 다음과 같다.

$$C(x, y) = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N D(m+x, n+y) Q(m, n) \quad (1)$$

$$(-M \leq x \leq X, -N \leq y \leq Y)$$

여기서 " \circ "는 상관 연산자를 나타내며 상관 함수는 D 와 Q 가 서로 겹치는 영역에서의 곱에 대한 합 연산을 수행하여 (x, y) 에서의 값을 계산한다. 식 1과 같은 과정을 문서 영상의 모든 화소에 대해 반복하여 계산한다. 계산된 상관 계수 중 최대값을 갖는 좌표는 두 영상이 가장 잘 일치하는 위치를 나타내며, 일치하는 값이 클수록 두 영상의 일치 정도가 높다고 판단한다.

또한 FFT를 이용한 상관 함수는 아래 식 2와 같은 정리가 성립한다.

$$d(x, y) \circ q(x, y) \Leftrightarrow D(u, v) Q^*(u, v) \quad (2)$$

여기서 $D(u, v)$ 와 $Q(u, v)$ 는 각각 함수 $d(x, y)$ 와 $q(x, y)$ 를 Fourier 변환한 주파수 영역에서의 함수를 말하며, " \Leftrightarrow "는 식의 좌변과 우변, $d(x, y) \circ q(x, y)$ 와 $D(u, v) Q^*(u, v)$ 이 서로 Fourier 변환 쌍을 구성하는 것을 나타낸다. 즉, 공간 영역에서의 상관도 $d(x, y) \circ q(x, y)$ 는 주파수 영역에서 두 함수의 곱 $D(u, v) Q^*(u, v)$ 을 역 Fourier 변환함으로써 얻을 수 있다. 이는 식 3과 같이 반대의 경우도 성립한다.

$$d(x, y) q^*(x, y) \Leftrightarrow D(u, v) \circ Q(u, v) \quad (3)$$

본 논문에서는 두 영상간의 상관 정도를 계산하기 위해, 위에서 언급한 FFT 기반 상관 함수를 이용한다. 수행과정은 다음과 같다. 문서 영상을 d 질의 영상을 q (FFT 변환 결과와 비교하기 위해 소문자로 표기)라고 할 때 두 영상을 FFT 변환하여 D 와 Q 를 얻고, 테스트

영상에 대한 변환 값에 켈레 복소수를 적용하여 Q^* 을 계산한다. 식 2의 우변과 같이, 각각 Fourier 변환된 두 항을 서로 곱한 후 역 Fourier 변환을 취하게 되면 좌변과 동일한 값을 얻게 된다.

2. Hausdorff 거리(Hausdorff Distance, HD)

Hausdorff 거리[2]는 두 유한 집합 내의 모든 정점들 간의 비일치도를 측정하는 방법으로 일반적인 거리 함수(Euclidean, Manhattan, Chamfer 등)를 이용한 최대 최소(minmax) 기법이다. 두 텍스트 영상은 이진 영상이므로 m 과 n 개의 원소(점은 화소)를 갖는 두 유한집합 $D=\{d_1, \dots, d_m\}$ 와 $Q=\{q_1, \dots, q_n\}$ 로 간주할 수 있다. 두 집합 간의 Hausdorff 거리는 다음과 같다.

$$H(D, Q) = \max(h(D, Q), h(Q, D)) \quad (4)$$

$$h(D, Q) = \max_{d_i \in D} \min_{q_j \in Q} \|d_i - q_j\| \quad (5)$$

여기서 $\|d_i - q_j\|$ 는 문서 영상 내의 원소 d_i 와 질의 영상 내의 원소 q_j 사이의 기하학적 거리를 나타내는 것으로 계산 속도를 고려하여 Manhattan 거리를 이용한다. $h(D, Q)$ 는 먼저 문서 집합 내의 한 정점 $d_i \in D$ 와 질의 집합(Q) 내의 모든 정점을 비교하여 가장 가까이 있는 정점 $q_j \in Q$ 를 선택하고, 이러한 과정을 집합 D 내의 모든 정점에 대해 반복하여 계산한 후 이 중 가장 큰 값, 즉 가장 멀리있는 정점까지의 거리를 계산한다. $h(Q, D)$ 도 동일한 방법으로 계산한다. 계산된 두 값 중 최대값이 두 집합 간의 Hausdorff 거리이다. 두 집합이 유사할수록 거리는 작게 된다. 만일 질의 집합의 모든 점들을 한 픽셀씩 이동이 가능하다고 가정하면, 다음 식과 같이 거리가 가장 최소가 되는 좌표 t 를 계산할 수 있다.

$$H_T(D, Q) = \min_t H(D, Q \oplus t) \quad (6)$$

여기서 \oplus 는 Minkowski 합을 나타내는 것으로 $Q \oplus t = \{q + t | q \in Q\}$ 로 표현된다. 위의 식 6으로부터 질의 집합 Q 의 모든 정점들이 문서 집합 D 의 정점들에 가장 최소의 거리로 일치되는 지점 t 가 존재하면 이때의 $H_T(D, Q)$ 는 앞서 계산한 $H(D, Q)$ 보다 작게 되고, 이 값을 두 영상간의 비일치 정도로 판단한다.

앞서 언급한 식 5의 $h(D, Q)$ 는 최대(max) 연산의 특성으로 인해 집합 내의 원소들에 민감하게 반응하므로, 이상치가 발생한 경우에 큰 오류를 낼 수 있는 단점이 있다. 이에 최대 연산 대신 평균(average) 연산을 취하여 이상치의 영향을 줄이는 방법을 수정된 Hausdorff 거리(modified Hausdorff distance)라고 한다. 수정된 Hausdorff 거리는 다음과 같이 정의한다[13].

$$h(D, Q) = \frac{1}{n} \sum_{j=1}^n \min_{d_i \in D} \|d_i - q_j\| \quad (7)$$

이 방법은 문서 집합 내의 한 정점 $d_i \in D$ 에 대해 그 정점과 가장 가까이 있는 질의 집합 내의 정점 $q_i \in Q$ 를 찾고, 이러한 연산을 n 개의 모든 점에 대해 계산하여 이들 거리 값을 평균한다. 최대 연산이 아닌 평균 연산을 통해, 노이즈의 영향을 최소화 할 수 있다.

3. 지역 밀집도(Local Crowdedness, LC)

제안 알고리즘은 [그림 1]에 나타난 것처럼 특징 추출 단계와 유사도 측정 단계로 구성된다. 유사도 측정 단계는 2절에 언급한 Hausdorff 거리를 적용하였다. 그리고 특징 추출 단계에서 적용한 지역 밀집도는 다음과 같다.

Shiku[11]는 문서 영상에서 텍스트 영역을 추출하기 위해 지역 밀집도와 전역 밀집도 방법을 제안하였다. 지역 밀집도는 문서 영상에서 검은 픽셀의 연결요소(connected-component)들을 하나의 세그먼트(segment)로 보고 이들 세그먼트 내의 픽셀 수를 이용하여 계산하고, 전역 밀집도는 허프 변환(Hough transform)을 수행하여 일정한 직선을 지나는 모든 세그먼트들의 개수를 측정하여 계산한다. 본 논문에서는 텍스트 영상에서 글자의 분기점이나 끝점과 같은 특징점을 추출하기 위해

검은 화소의 밀집된 정도를 나타내는 지역 밀집도 방법을 응용한다. Shiku의 제안 방법에서는 세그먼트 단위로 밀집도를 계산하여 전체 문서 영상에서의 텍스트 영역을 검색하지만, 본 논문에서는 텍스트 영상을 픽셀단위로 고려하여 글자의 분기점과 같은 영역에서 높은 밀집도를 갖도록 하고, 일정한 임계값 이상의 높은 밀집도를 갖는 지점을 특징점으로 추출한다. 지역 밀집도 $D(x, y)$ 는 다음과 같이 정의한다.

$$D(x, y) = \sum_{i=1}^N e^{-\frac{d_i^2}{2\sigma^2}} \quad (8)$$

여기서, N 은 (x, y) 좌표를 원점으로 반지름 r ($=3\sigma$)인 원 영역 내에 포함된 검은 화소의 개수이며 $\exp(-d_i^2/2\sigma^2)$ 는 중심점 (x, y) 에서 반지름이 r 인 원 내에 포함된 i 번째 화소까지의 거리 d_i 에 대한 가중치이다. 즉, 지역 밀집도 $D(x, y)$ 는 이들 가중치들을 합하여 계산하는 것으로 특정한 원 영역 내에 얼마나 많은 검은 화소들이 밀집되어 있는지를 나타낸다. 해당 좌표에서의 지역 밀집도가 일정한 임계값보다 크면 특징점으로 선택한다.

[그림 2]는 텍스트 영상에 지역 밀집도를 계산한 결과를 보여준다. 본 논문에서는 실험을 통해 σ 는 1.5, 임계값은 100을 선택하였다.

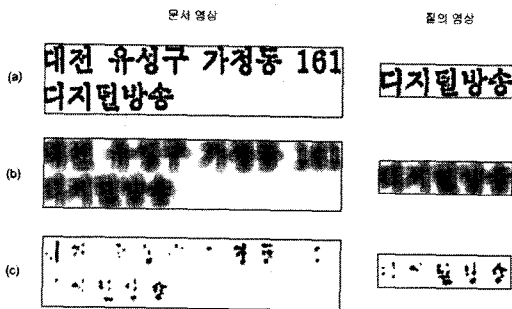


그림 2. 지역 밀집도를 이용한 특징 추출 :
 (a) 입력영상, (b) 지역 밀집도,
 (c) 추출된 특징

III. 실험 및 결과

본 절에서는 두 영상간의 유사도를 측정하기 위한 방법으로 제안 방법(LC+HD)과 BC, 그리고 HD를 우편물 자동화(postal automation) 문제에 적용하여 실험 결과를 비교한다. 본 논문에서 제안한 방법과 비교 가능한 텍스트 영상 매칭 기법에 관한 연구가 아직까지는 수행되지 않았기 때문에, 여기서는 가장 기본적인 영상 매칭 기법인 BC 및 HD의 성능과 비교하여 텍스트 영상 검출에 적절한 기법이 무엇인지를 알아보려고 하였다.

우편물 자동화란 우편물에 인쇄된 수신자 주소를 자동으로 인식하여 우편물 분류, 발송, 순로구분 등에 따른 각종 수작업을 최소화하는 기술을 의미한다. 현재 국내에 도입되어 사용되고 있는 우편배달 자동화 체제에서는 접수된 우편물을 집중국에서 일괄적으로 인식하여 배달점 코드(delivery point code)를 생성하고, 이를 바코드형태로 우편물상에 직접 인쇄한 후, 이들 바코드를 판독하여 배달지 우체국(배달국) 별로 분류한다. 우편물이 특정 배달국에 도착하면 집배원별로 자신의 이동경로에 따라 우편물들을 재정렬하게 되는데 이를 순로구분이라 하며, 이때에도 우편물에 인쇄된 바코드를 판독하여 순로구분 작업을 자동화할 수 있다[14].

[15]에서는 순로구분시스템으로부터 고가의 바코드 판독기를 제거하되 동일한 성능의 분류 능력을 갖도록 하기 위해 virtual ID 기반의 순로구분 방식을 제안하였다. Virtual ID를 이용한 우편물 검증은 순로구분시스템에서 우편물 검증을 위한 수단으로 사용되었던 바코드나 사용자 Tag를 사용하지 않고 우편물 영상에서 추출한 특징을 이용하여 우편물을 검증하는 방법이다. 즉, 집중국에서는 우편물의 주소를 인식하는 과정과 우편물에 대한 특징을 virtual ID 형태로 저장하는 과정을 거친다. 배달국으로 분류된 우편물은 순로구분시스템에 입력되어, 집중국에서와 동일한 방법으로 우편물의 특징을 추출한다. 추출된 특징을 저장된 virtual ID와 비교하여 동일한 우편물을 찾아낸다. Virtual ID에는 해당 우편물에 대한 주소정보가 있기 때문에 이동경로에 맞게 순서대로 재정렬할 수 있다.

Virtual ID 기반 순로구분 시스템에서 시차를 두고 스

캔된 두 우편물 영상의 동일성 여부를 판단하여 일치하는 우편물을 검증해내는 과정이 필요한데, 이러한 문제를 본 논문에서의 제안 방법을 적용하여 성능을 평가하였다.

1. 실험 영상

실험 영상은 우편 영상의 수신자 영역에서 서로 겹치는 정도에 따라 4가지 유형의 데이터 집합으로 분류하여 정의한다. 즉, 데이터베이스는 두 영상의 내용이 거의 유사한 경우, 문서 영상이 질의 영상을 각각 70%, 50%, 30% 포함하는 경우로 나뉜다. 실험 영상을 생성하기 위해, 스캔된 우편물 영상으로부터 수신자 주소 영역을 추출하고, 이 영역 내의 텍스트 영상을 박스 단위로 선택한다. 동일한 우편물 영상을 다시 스캔하여 새로운 영상을 얻고, 이와 동일한 방법으로 수신자 주소 영역 내의 텍스트 영상을 박스 단위로 추출한다. 첫 번째 선택된 박스 단위 영상과 두 번째 선택된 박스 단위 영상을 서로 비교하여 앞서 설명한 4가지 유형에 맞도록 선택하고, 선택된 두 영상을 질의 영상과 문서 영상으로 이용한다. 실험에 이용한 4가지 유형의 데이터 집합은 각각 50개, 38개, 58개, 44개 영상 쌍으로, 전체 190개 쌍으로 구성된다. [그림 3]은 4가지의 데이터 집합에 대한 영상 쌍의 예를 보여준다.

	질의 영상		문서 영상
(a)		←→	
(b)		←→	
(c)		←→	
(d)		←→	

그림 3. 실험 영상의 예: (a) 문서 영상이 질의 영상을 약 100% 포함하는 경우, (b) 문서 영상이 질의 영상을 약 70% 포함하는 경우, (c) 문서 영상이 질의 영상을 약 50% 포함하는 경우, (d) 문서 영상이 질의 영상을 약 30% 포함하는 경우

동일한 우편물을 스캔한 영상일지라도 약간의 왜곡이

나 잘림 현상이 발생하게 되는데, 본 논문에서는 이동에 의해 잘려진 경우만을 고려한다.

2. 실험 결과

실험을 위해 1절에서 언급한 실험 데이터를 다음 2개의 클래스로 구분하였다. $class_1$ 은 질의 영상 Q를 포함하는 문서 영상 D의 쌍에 대한 클래스이며, $class_2$ 은 질의 영상 Q와 문서 영상 D가 서로 관련이 없는 쌍에 대한 클래스이다. 각 클래스별로 190개씩, 전체 380개 영상 데이터베이스를 생성하고, 이들 중 190개는 훈련 데이터로 나머지 190개는 테스트 데이터로 사용한다.

두 영상에 대한 유사 여부를 판단하기 위해 본 논문에서는 베이즈 결정 규칙(Bayes' decision rule)을 이용한다. 즉, 훈련 단계에서는 2장에 언급한 유사도 측정 방법을 통해 얻은 클래스별 거리 및 상관도 값에 대한 확률 분포를 훈련 데이터에 대해 추정하고, 테스트 단계에서는 앞서 추정된 확률 분포에 따라 테스트 데이터 내의 임의의 두 영상 간에 계산된 거리 및 상관도 값이 어느 클래스에 속하는지를 판단한다. [그림 4]는 추정된 클래스별 확률 분포와 두 클래스를 구분 짓는 최적의 결정 경계면을 보여준다.

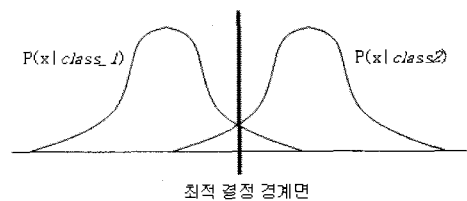


그림 4. 베이즈 결정규칙을 이용한 최적 결정 경계면

[그림 5]와 [표 1]은 텍스트 영상 매칭 방법에 대한 분류 정확률 및 처리 시간을 나타낸다. 여기서 BC와 HD는 각각 이진 상관 함수 및 Hausdorff 거리를 적용한 방법이며, LC+HD는 제안방법으로 지역밀집도와 Hausdorff 거리를 동시에 고려한 방법이다. [표 1]에 제시된 것처럼 실험 과정에서 발생할 수 있는 오류는 Type-I 오류와 Type-II 오류이다. Type-I 오류는

class_1에 속하는 영상을 class_2에 속한다고 잘못 분류하는 경우이고, Type-II 오류는 반대로 class_2에 속하는 영상을 class_1에 속한다고 잘못 분류하는 경우를 말한다. 실험 결과를 볼 때 제안 방법이 약 2.7%에서 9.0% 정도 다른 방법에 비해 우수함을 알 수 있다.

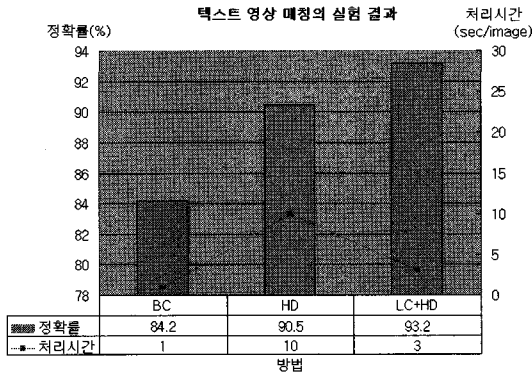


그림 5. 텍스트 영상 매칭 실험 결과

표 1. 텍스트 영상 매칭 실험 결과

방법	정확률 (%)	Type-I 오류 (%)	Type-II 오류 (%)	처리 시간 (sec/image)
BC	84.2	14.7	1.1	1
HD	90.5	0	9.5	10
LC + HD	93.2	2.1	4.7	3

[그림 6]은 class_1에 속하는 데이터에 대해 질의 영상이 문서 영상의 해당 위치를 정확하게 찾는지를 알아보고자 나타낸 것이다. 정확한 위치를 찾아 해당 좌표에 질의 영상을 겹쳐서 나타내었을 때 세 가지 방법 모두 유사한 부분을 찾아주긴 하지만, 그 중에서도 상관도 방법과 Hausdorff 거리 방법이 약간 더 우수함을 알 수 있다.

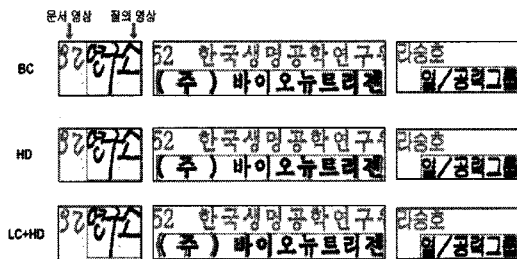


그림 6. 영상의 일치 영역 검색의 예

IV. 결론 및 향후 연구 방향

지금까지 영상과 영상을 매칭하는 방법으로 지역 밀집도와 Hausdorff 거리 방법을 이용한 알고리즘을 제안하고, 실제 우편 자동화 시스템에 제안 방법과 이전 상관 방법 및 Hausdorff 거리 방법을 적용하여 결과를 비교해보았다. 실험 결과, 이전 상관 방법은 190개 실험 영상 중 160개를 올바르게 인식하여 84.2%의 성능을 보였으며, Hausdorff 거리 방법은 172개를 정확히 인식하여 90.5%의 인식 성능을 나타내었다. 이에 반해 제안 방법은 177개의 영상을 정확히 인식하여 93.2%의 인식 성능을 나타내었으며, 특징추출을 적용하지 않은 HD에 비해 한 영상을 처리하는데 약 7초 정도를 단축할 수 있었다.

제안 방법은 영상기반 매칭 방법의 하나로 기존의 인식기반 매칭 방법에 필요한 문자별 모델을 생성하거나 저장하는 과정을 생략할 수 있기 때문에, 임의의 언어로 작성된 다양한 문서에 동일한 방법을 이용할 수 있는 언어 독립적(language independent)인 특징을 가지며, 시스템의 저장 공간을 효과적으로 줄일 수 있는 장점을 갖는다. 따라서 제안 방법은 앞서 실험한 우편물 자동분류 시스템에서의 우편물 검증 이외에도 다량의 문서 영상에서 전문검색(full-text retrieval)을 위한 핵심단어 인식(keyword spotting) 분야에 활용될 수 있다. 그리고 추가적으로 텍스트의 변형이나 왜곡에도 적용이 가능한 특징 추출 방법이나 영상 매칭 방법을 적용한다면 주어진 두 서명의 일치 여부를 판단하는 서명 검증(signature verification) 분야 또는 입력되는 하나의 필기체 영상에 대해 기존에 알고 있는 필기자 중 한명을 인식하는 필기자 인식(writer identification) 분야와 같이 필기체 영상을 매칭하는 문제에도 응용할 수 있다.

본 논문에서 다룬 텍스트 영상간의 매칭 시스템은 한 영상이 단순히 이동되어 질의 영상이 문서 영상에 일부 포함된 경우만을 고려하였다. 즉, 이동에 의해 영상의 일부분이 일치하는 경우 제안 방법이 효과적으로 적용될 수 있었다. 그러나 질의 영상이 축소되거나 확대되는 경우, 또는 회전되는 경우에는 제안한 방법을 적용하기 어렵다. 향후 스케일링(scaling)되거나 회전된 영상 매칭에 유용한 방법 및 텍스트 변형이나 왜곡에 불변하는 특

정을 추가하여 앞서 언급한 다양한 응용분야에도 적용해보고자 한다.

참고 문헌

- [1] A. Averbuch and Y. Keller, "FFT Based Image Registration," International Conference on Acoustics, Speech, and Signal Processing, Vol.4, pp.3608-3611, 2002.
- [2] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing Images Using the Hausdorff Distance," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15, No.9, pp.850-863, 1993.
- [3] N. Sebe, M. S. Lew, and D. P. Huijsmans, "Multi-Scale Sub-Image Search," Proceedings of the 7th ACM International Conference on Multimedia(Part II), pp.79- 82, 1999.
- [4] J. Luo and M. A. Nascimento, "Content Based Sub-Image Retrieval via Hierarchical Tree Matching," Proceedings of the 1th ACM International Workshop on Multimedia Databases, pp.63-69, 2003.
- [5] S. N. Srihari and Z. Shi, "Forensic Handwritten Document Retrieval System," Proceedings of the 1st International Workshop on Document Image Analysis for Libraries, pp.188-194, 2004.
- [6] Y. Chen, J. Z. Wang, and R. Krovetz, "Content-Based Image Retrieval by Clustering," Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval, pp.193-200, 2003.
- [7] H. Peng, F. Long, W. C. Sui, Z. Chi, and D. D. Feng, "Document Image Matching Based on Component Blocks," IEEE International Conference on Image Processing, Vol.2, pp.601-604, 2000.
- [8] H. Peng, F. Long, Z. Chi, and W. C. Siu, "Document Image Template Matching Based on Component Block Lists," Pattern Recognition Letters, Vol.22, No.9, pp.1033-1042, 2001.
- [9] Il-Seok Oh, Yoon-Sung Choi, Jin-Ho Yang, Soo-Hyung Kim, "A Keyword Spotting System of Korean Document Images," Proceedings of the 5th International Conference on Asian Digital Libraries, pp.530-534, 2002.
- [10] 진영범, 오일석, "대용량 인쇄 한글 문서 검색을 위한 영상기반 단어 매칭 방법", 한국정보과학회 추계 학술발표대회 논문집, 제27권, 제2호, pp.461-463, 2000.
- [11] O. Shiku, K. Kawasue, and A. Nakamura, "A Method for Character String Extraction Using Local and Global Segment Crowdedness," Proceedings of the 14th International Conference on Pattern Recognition, Vol.2, pp.1077-1080, 1998.
- [12] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, 2nd Edition, Prentice Hall, 2002.
- [13] M. P. Dubuisson, A. K. Jain, "A Modified Hausdorff Distance for Object Matching," Proceedings of the 12th International Conference on Pattern Recognition, Vol.1, pp.566-568, 1994.
- [14] 김수형, 박상철 외 4명, *Virtual ID 사용을 위한 우편 영상 검증 기술 연구 과제 중간보고서*, 연구보고서, 한국전자통신연구원, 2004.
- [15] 정창부, 박상철, 손화정, 김수형, "수취인 주소 영역의 영상매칭에 기반한 유사 우편물 추출 방법", 2004 한국정보과학회 추계 학술발표대회 논문집, 제31권, 제2호, pp.793-795, 2004.

저자소개

손 화 정 (Hwa-Jeong Son)

정회원

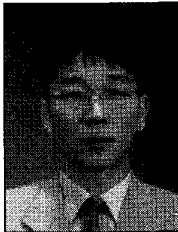


- 2001년 8월 : 전남대학교 통계학
과(이학사)
- 2004년 2월 : 전남대학교 전산학
과(이학석사)
- 2004년 3월~현재 : 전남대학교
전산학과 박사과정

<관심분야> : 패턴인식, 문서영상 정보검색

김 수 형 (Soo-Hyung Kim)

정회원

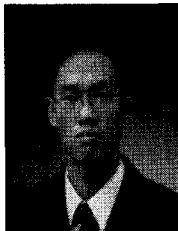


- 1986년 2월 : 서울대학교 컴퓨터
공학과(공학사)
- 1988년 2월 : 한국과학기술원 전
산학과(공학석사)
- 1993년 8월 : 한국과학기술원 전
산학과(공학박사)

- 1990년 9월~1996년 12월 : 삼성전자 멀티미디어 연
구소(선임연구원)
 - 2000년 12월~2002년 1월 : 캐나다 Concordia 대학
CENPARMI 연구소(방문교수)
 - 1997년 1월~현재 : 전남대학교 전자컴퓨터공학부 교수
- <관심분야> : 인공지능, 패턴인식, 문서영상 정보검
색, 유비쿼터스컴퓨팅

김 지 수 (Ji-Soo Kim)

정회원

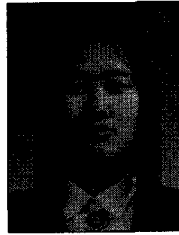


- 1996년 2월 : 광주대학교 컴퓨터
공학과(공학사)
- 2003년 2월 : 전남대학교 전산학
과(이학석사)
- 2003년 3월~현재 : 전남대학교
전산학과 박사과정

<관심분야> : 인공지능, 패턴인식, 문자인식

박 미 선 (Mi-Seon Park)

준회원



- 2005년 2월 : 전남대학교 정보통
신공학부 컴퓨터공학전공(공학
사)
 - 2005년 3월~현재 : 전남대학교
전산학과 석사과정
- <관심분야> : H.264 에러복원, 멀
티미디어

유 재 명 (Jae-Myeong Yoo)

정회원



- 1992년 2월 : 전남대학교 응용화
학공학부(공학사)
- 2003년 2월 : 전남대학교 전산학
과(이학석사)
- 2003년 3월~현재 : 전남대학교
전산학과 박사과정

<관심분야> : 멀티미디어 통신, 영상처리 및 복원