

부분 영상 매칭에 기반한 텍스트 검증

손 화 정[†] · 정 선 화^{††} · 김 수 형^{†††}

요 약

한 영상이 다른 영상을 포함하고 있는 경우, 이들 영상의 일치 여부를 판단하는 부분 영상 매칭 방법은 대부분 자연 영상을 대상으로 연구되고 있다. 본 논문에서는 자연 영상이 아닌 텍스트 영상을 매칭하는데 효과적인 두 가지 기법, 즉 메쉬 기반 방법과 상관성 기반 방법을 제안하고자 한다. 메쉬 기반 방법은 두 영상의 일치하는 모서리를 찾은 후 겹치는 영역에 대한 메쉬 특징을 이용하여 유사 여부를 판단하는 것으로, 일치 영역 검색 단계와 유사성 측정 단계로 구성된다. 상관성 기반 방법은 FFT를 이용하여 두 영상의 상관성을 계산함으로써 유사도를 측정한다. 우편 자동화 시스템에서 텍스트 영상을 검증하는 분야에 제안 방법을 적용한 결과, 메쉬 기반 방법은 90.1%, 상관성 기반 방법은 92.7%의 성능을 나타내었다.

Text Verification Based on Sub-Image Matching

Hwa Jeong Son[†] · Seon Hwa Jeong^{††} · Soo Hyung Kim^{†††}

ABSTRACT

The sub-image matching problem, in which one image contains some part of the other image, has been mostly investigated on natural images. In this paper, we propose two sub-image matching techniques: mesh-based method and correlation-based method, that are efficiently used to match text images. Mesh-based method consists of two stages, box alignment and similarity measurement by extracting the mesh feature from the two images. Correlation-based method determines the similarity using the correlation of the two images based on FFT function. We have applied the two methods to the text verification in a postal automation system and observed that the accuracy of correlation-based method is 92.7% while that of mesh-based method is 90.1%.

키워드 : 부분영상 매칭(Sub-Image Matching), 텍스트 인식(Text Recognition), 우편 분류 자동화(Postal Automation)

1. 서 론

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 부분 영상을 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 부분 영상 매칭 기법을 제안하고자 한다.

두 텍스트 영상의 유사성을 측정하기 위한 방법은 인식

기반 매칭 방법과 영상 기반 매칭 방법으로 나누어 볼 수 있다. 인식 기반 매칭 방법은 각 텍스트 영상의 내용을 인식한 후 인식한 내용을 기반으로 유사 여부를 판단하므로 엄밀히 말하면 텍스트 영상의 내용이 동일한지를 판별하는 방법이다. 예를 들어 두 텍스트 영상의 글자의 글꼴과 글자의 크기가 다르더라도 인식결과 내용이 동일하면 동일영상으로 판별하게 된다. 이와는 달리 영상 기반 매칭 방법은 영상에서 다른 영상과 구별되는 특징을 추출하여 매칭을 시도하는 방법으로 두 텍스트 영상의 글자의 내용뿐만 아니라 글꼴과 글자의 크기가 동일한 경우에 동일 영상으로 판별한다.

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 인식 기반 매칭 방법보다는 영상 기반 매칭 방법을 통하여 이들 영상의 일치성 여부를 판단함이 옳다.

※ 본 연구는 한국전자통신연구원의 "Virtual ID 사용을 위한 우편 영상 검증 기술 연구" 과제의 지원으로 수행되었음.

† 준 회 원 : 전남대학교 전산학과 박사과정

†† 정 회 원 : 한국전자통신연구원 우정기술연구센터 선임연구원

††† 정 회 원 : 전남대학교 전산학과 교수

논문접수 : 2004년 12월 10일, 심사완료 : 2005년 3월 14일

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 영상 기반 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 부분 영상 매칭 방법을 적용한다면, 검색 오류나 수행 시간을 보다 효과적으로 줄일 수 있을 것으로 판단된다.

본 논문에서는 영상간의 일치성 여부를 판단하기 위하여 영상 기반 매칭 방법인 메쉬 기반 방법과 상관성 기반 방법을 제안한다. 메쉬 기반 방법은 일치 영역 검색 단계와 유사성 측정 단계로 구성된다. 즉, 입력된 두 영상의 일치하는 위치를 찾기 위해 4개의 모서리 중 가장 유사한 모서리를 선택하는 단계와 선택된 모서리를 이용하여 두 영상의 겹치는 부분을 계산하고 화소들의 밀도를 고려한 메쉬 특징을 추출하여 두 영상의 일치 여부를 판단하는 단계로 이루어진다. 상관성 기반 방법은 원형 정합(template matching) 분야에 널리 이용되는 방법으로 두 영상에 대한 상관성을 이용하여 일치 여부를 판단하기 위해 사용한다. 또한, Fast Fourier Transform(FFT)을 이용하여 주파수 영역(frequency domain)에서 두 영상의 상관성을 계산하기 때문에 공간 영역(spatial domain)에서 계산하는 것보다 더 빠른 시간에 효과적으로 수행할 수 있다.

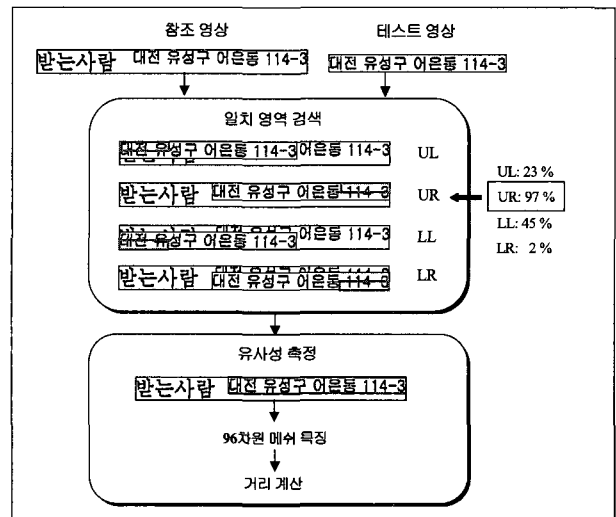
제안 방법은 메쉬 특징 또는 상관성 함수를 이용한 단순한 알고리즘으로 기존의 인식 기반 매칭 방법에서 필요한 각 문자에 대응되는 모델을 설정하거나 저장하는 과정을 생략할 수 있다. 따라서 임의의 언어로 작성된 다양한 문서에도 동일한 방법을 이용하여 적용할 수 있으며, 시스템의 저장 공간을 효과적으로 줄일 수 있는 장점이 있다. 또한 제안 방법은 다량의 문서 영상 데이터베이스에 대한 전문검색(full-text retrieval)을 위한 핵심단어 인식(keyword spotting) 분야 또는 우편물 자동 분류 시스템에서 영상 매칭 기반으로 우편물을 검증하는데 활용될 수 있으며, 두 서명의 일치 여부를 판단하는 서명 검증(signature verification) 분야, 입력되는 하나의 필기체 영상에 대해 기존에 알고 있는 필기자 중 한명을 인식하는 필기자 인식(writer identi-

fication) 분야와 같은 필기체 영상을 매칭하는 문제에도 응용할 수 있다.

본 논문은 다음과 같이 구성된다. 2절에서는 메쉬 기반 방법을 설명하고, 3절에서는 상관성 기반 방법을 다룬다. 4절에서는 제안한 두 가지 방법의 성능을 평가하기 위해 우편 자동화의 응용분야에 적용한 결과를 보여준다. 마지막 5절에서는 결론 및 향후 연구 방향에 대해 기술한다.

2. 메쉬 기반 방법

메쉬 기반 방법은 일치 영역 검색 단계와 유사성 측정 단계로 구성된다. (그림 1)과 같이 일치 영역 검색 단계에서는 영상의 네 모서리 중 가장 일치하는 모서리를 선택하고 유사성 측정 단계에서는 메쉬 특징을 이용하여 두 영상의 일치 정도를 계산한다. 두 영상 중 하나의 영상은 상하좌우가 일부 잘려져 있을 수 있다. 그러나 동시에 상하가 잘리거나 동시에 좌우가 잘려져 있지는 않다. 이 때문에 네 모서리 중 하나의 모서리와 일치하는 영역이 찾아질 수 있다.



(그림 1) 메쉬 기반 방법의 수행 과정

2.1 일치 영역 검색

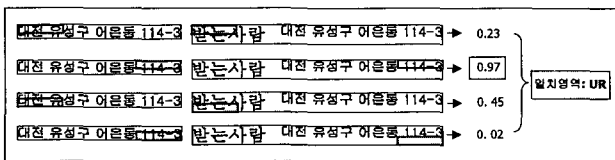
일치 영역을 검색하기 위하여 먼저 두 영상을 4군데 모서리, 즉 좌상(Upper Left: UL), 우상(Upper Right: UR), 좌하(Lower Left: LL), 우하(Lower Right: LR)를 기준으로 정렬한다. 그 다음 4군데 모서리로부터 블록을 선택한다. 영상 전체를 비교하지 않고 블록을 사용하여 일치 여부를 결정하는 이유는 처리 속도를 줄이기 위해서이다. 블록의 너비(w)와 높이(h)는 식(1)과 같이 계산된다. w_1 과 h_1 는 첫 번째 영상의 너비와 높이를 나타내며, w_2 와 h_2 는 두 번째 영상의 너비와 높이를 나타낸다.

$$\begin{aligned} W &= \text{MIN}(w_1, w_2) \times 0.4 \\ H &= \text{MIN}(h_1, h_2) \times 0.4 \end{aligned} \quad (1)$$

두 영상에서 각 모서리에 대해 블록을 선택한 후, 서로 일치하는 흑화소의 개수를 계산하고 네 개의 모서리에 해당하는 블록 중 일치하는 흑화소의 수가 가장 많은 블록의 모서리를 일치 모서리로 선택한다. 계산식은 다음과 같다.

$$Corner = \arg \text{MAX}_{1 \leq i \leq 4} \left\{ \frac{2 \times n_i}{t_{i1} + t_{i2}} \right\} \quad (2)$$

n_i 는 두 영상에서 i 번째 모서리에 해당하는 블록간에 일치하는 흑화소의 개수를 말하고, t_{i1} 과 t_{i2} 는 각 영상에서 선택된 블록 내에 있는 흑화소의 수를 의미한다. (그림 2)는 제안 방법을 임의의 두 영상에 적용하여 가장 일치하는 모서리를 선택하는 과정을 보여준다. 식 (1)과 식 (2)에 의해 각 모서리로부터 선택된 블록의 일치 정도를 계산한 결과, UR에 해당하는 블록의 일치도가 가장 높다. 따라서 두 영상의 일치 영역을 측정하기 위한 기준점이 되는 모서리를 UR로 결정한다.

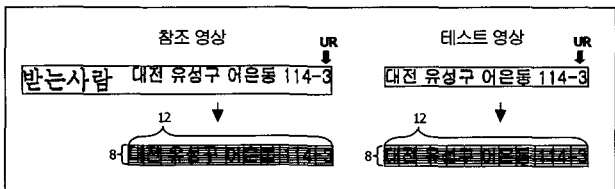


(그림 2) 일치 영역 검색 단계의 예

2.2 유사성 측정

2.1의 방법에 의해 선택된 모서리에서 두 영상이 겹치는 영역을 선택한 후, 식 (3)과 같이 96차원(8×12) 메쉬 특징을 추출한다. 메쉬 특징을 추출하는 과정은 (그림 3)과 같다.

$$F(k) = \frac{k\text{번째 셀의 흑화소수}}{k\text{번째 셀의 전체크기}} \quad (1 \leq k \leq 96) \quad (3)$$



(그림 3) 메쉬 특징 추출 과정

두 영상에 대한 유사도를 측정하기 위해, 식 (4)와 같이 앞서 계산된 두 영상의 메쉬 특징간의 Manhattan 거리를 계산한다.

$$Dist = \sum_{k=1}^{96} |F_{Ref}(k) - F_{Test}(k)| \quad (1 \leq k \leq 96) \quad (4)$$

여기서 F_{Ref} 와 F_{Test} 는 각각 참조 영상(reference image)과 테스트 영상(test image)으로부터 추출된 96차원 메쉬 특

징을 말한다. 거리가 작을수록 두 영상간의 유사한 정도는 커지게 된다.

3. 상관성 기반 방법

상관성은 두 영상간의 상관 정도를 나타내는 기본적인 방법으로 원형 정합(template matching) 분야에 주로 이용되고 있다. 이 매칭 기법은 기존에 알고 있는 영상의 집합 중에서 입력된 임의의 영상과 가장 잘 일치하는 영상을 찾아내는 것으로, 영상간의 상관성을 이용한다. 즉, 가장 큰 상관성 함수 값을 갖는 영상이 가장 일치하는 영상이라고 판단한다.

3.1 상관성 함수

두 영상간의 상관성을 계산하는 함수는 다음과 같다[11].

$$f(s, t) \circ w(s, t) = \sum_x \sum_y f(x, y)w(x - s, y - t) \quad (5)$$

여기서 " \circ "는 상관성 연산자를 나타내며 $f(x, y)$ 는 $A \times B$ 크기의 영상, $w(x, y)$ 는 $C \times D$ 크기의 영상을 나타낸다 ($A \geq C, B \geq D$ 라고 가정). 상관성 함수는 f 와 w 가 서로 겹치는 영역에서의 곱에 대한 합 연산을 수행하여 (s, t) 에서의 값을 계산한다. 식 (5)와 같은 과정을 모든 좌표 (s, t) 에 대해 반복하여 계산한다. 계산된 상관 계수 값 중 최대 값을 갖는 좌표는 $f(x, y)$ 와 $w(x, y)$ 가 가장 잘 일치하는 위치를 나타내며, 일치하는 값이 클수록 두 영상의 일치 정도가 높다고 판단한다.

3.2 FFT 기반 상관성 함수

3.1에 언급된 상관성 함수는 영상 w 의 크기가 커질수록 계산 복잡도가 커지는 단점을 갖는다. 이러한 경우 FFT를 통해 주파수 영역에서 상관성을 계산하는 것이 훨씬 효율적이다. Campbell[12]이 수행한 비교에서는 영상 w 에서 0이 아닌 값을 갖는 화소들의 개수가 132(약 13×13 픽셀)보다 클 경우 식 (5)를 수행하는 것보다 FFT를 이용하는 방법이 훨씬 빠르다는 것을 보여준다. 참고적으로 함수 $f(x, y)$ 와 $w(x, y)$ 가 동일한 영상인 경우를 자기상관성(autocorrelation) 함수라고 하며 두 함수가 동일하지 않은 영상인 경우를 교차상관성(cross correlation) 함수라고 말한다. FFT를 이용한 두 영상에 대한 상관성은 다음과 같이 정의된다[11].

$$f(s, t) \circ w(s, t) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)w^*(x + s, y + t) \quad (6)$$

여기서 " $*$ "는 켈레 복소수(complex conjugate)를 나타낸다. $f(x, y)$ 와 $w(x, y)$ 를 이용한 상관성 함수가 주기성을 갖는다고 가정할 경우, 겹침 오차(wraparound error)를 피하기 위해 제로 패딩(zero padding) 방법을 이용하여 두 영상을

$M \times N$ 의 크기로 확장하는 과정이 필요하다. 즉, 두 영상이 $M \times N$ 의 크기가 되도록 0 값을 대입한다. 여기서 M 과 N 은 다음 식 (7)을 만족해야 한다.

$$\begin{aligned} M &\geq A + C - 1 \\ N &\geq B + D - 1 \end{aligned} \quad (7)$$

또한 FFT를 이용한 상관성 함수는 아래와 같은 상관성 정리가 성립한다.

$$f(s, t) \cdot w(s, t) \Leftrightarrow F(u, v) W^*(u, v) \quad (8)$$

여기서 $F(u, v)$ 와 $W(u, v)$ 는 각각 함수 $f(s, t)$ 와 $w(s, t)$ 를 Fourier 변환한 주파수 영역에서의 함수를 말한다. “ \Leftrightarrow ”는 $f(s, t) \cdot w(s, t)$ 와 $F(u, v) W^*(u, v)$ 가 서로 Fourier 변환 쌍을 구성하는 것을 나타낸다. 즉, 공간 영역에서의 상관성 $f(s, t) \cdot w(s, t)$ 은 주파수 영역에서 두 함수의 곱 $F(u, v) W^*(u, v)$ 을 역 Fourier 변환함으로써 얻을 수 있다는 것을 알 수 있다. 이는 식 (9)와 같이 반대의 경우도 성립한다.

$$f(s, t) w^*(s, t) \Leftrightarrow F(u, v) \cdot W(u, v) \quad (9)$$

본 논문에서는 두 영상간의 상관 정도를 계산하기 위해, 위에서 언급한 FFT 기반 상관성 함수를 이용한다. 수행과정은 다음과 같다. 참조 영상을 f , 테스트 영상을 w 라고 할 때 두 영상을 FFT 변환하여 F 와 W 를 얻고, 테스트 영상에 대한 변환 값에 켈레 복소수를 적용하여 W^* 를 계산한다. 식 (8)의 우변과 같이, 각각 Fourier 변환된 두 항을 서로 곱한 후 역 Fourier 변환을 취하게 되면 좌변과 동일한 값을 얻게 된다. 즉, 공간 영역에서의 두 영상 f, w 에 대한 상관성을 얻게 된다. 만일 영상의 크기가 $M \times N$ 이라 하면 식 (5)를 이용한 상관성 방법은 $(M \times N)^2$ 시간이 걸리는데 반해, 식 (6)을 이용한 FFT 기반 상관성 방법은 $(M \log N) \times (N \log N)$ 시간이 걸린다. 따라서 매칭하고자 하는 영상의 크기가 클수록 FFT 기반 상관성 함수를 이용하는 것이 효과적이다.

위와 같은 방법에 의해 계산된 상관도를 이용하여 참조 영상과 테스트 영상의 유사성을 측정한다. 상관도가 클수록 두 영상의 일치 정도가 크고, 이때의 값이 최대인 좌표가 두 영상이 가장 일치하는 지점이다. 따라서 상관도의 최대 값을 두 영상의 유사도로 간주하고 이 값이 클수록 두 영상간의 일치 정도 역시 크다고 판단한다.

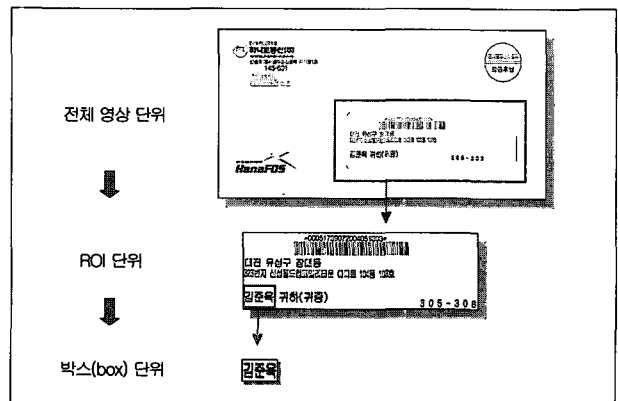
4. 실험 및 결과

이 절에서는 두 영상간의 유사도를 측정하기 위한 방법으로 제안한 메쉬 기반 방법과 상관성 기반 방법을 우편 자동화(postal automation) 문제에 적용한 결과를 보여주고자 한다. 본 논문에서 제안한 방법과 비교 가능한 부분 영상 매칭 기법에 관한 연구가 아직까지는 수행되지 않았기 때문에,

여기서는 본 논문에서 제안한 두 가지 기법의 성능과 수행 시간을 비교하여 텍스트 영상의 검증에 적절한 부분 영상 매칭 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, 이 때에도 우편물에 인쇄된 바코드를 판독하여 순로구분 작업을 자동화할 수 있다[13].

본 연구에서는 순로구분시스템으로부터 고가의 바코드 판독기를 제거하되 동일한 성능의 분류 능력을 갖도록 하기 위해 virtual ID 기반의 순로구분 방식을 제안한다. Virtual ID를 이용한 우편물 검증은 순로구분시스템에서 우편물 검증을 위한 수단으로 사용되었던 바코드나 사용자 Tag를 사용하지 않고 우편물 영상에서 추출한 특징을 이용하여 우편물을 검증하는 방법이다[13]. 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. 즉, 시차를 두고 스캔된 두 우편물 영상의 동일성 여부를 판단하여 일치하는 우편물을 검증해내는 과정이므로 글꼴이나 글자의 크기, 문단의 모양 등이 동일한 두 영상을 비교하는 문제이다. 그러므로 본 연구에서는 인식 기반 매칭 방법이 아닌 영상 기반 매칭 방법을 적용하는 것이 바람직하다.



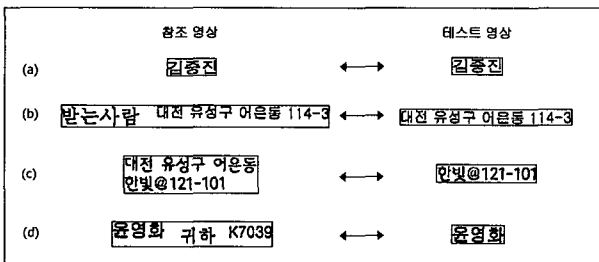
(그림 4) Virtual ID 기반의 단계별 우편물 검증

Virtual ID를 이용한 우편물 자동화 시스템은 성능 및 처리 속도를 고려하여 (그림 4)와 같이 3단계의 계층적인 과

정(우편물 전체 영상 단위 검증, 수신자주소 영역(ROI, Region of Interest)의 영상 단위 검증, ROI 내의 텍스트 박스 영상 단위 검증)으로 구성된다. 임의의 우편물 영상을 검증하기 위해 최대 7개의 참조영상을 대상으로 각 단계별로 두 우편물 영상이 비교되고, 유사하지 않다고 판단되는 영상은 걸러내게 된다[14]. 또한 전체 영상, ROI 영상, ROI 내의 박스 영상으로부터 추출된 특징들을 각 단계별로 달리 적용함으로써 효율적인 분석을 수행할 수 있도록 한다. 여기서, 우편물 자동화 시스템의 세 번째 단계에서는 두 번째 단계에서 선택된 ROI 영역에서 박스 단위로 영상을 추출하여 검증을 수행한다. 그러나 우편 영상이 스캔될 때 이동, 기울어짐 또는 회전 등과 같은 변형이 발생하게 되면 동일한 우편 영상이라 할지라도 ROI 영역으로부터 추출된 박스 영상은 그 일부가 잘려지거나 훼손되는 현상이 발생하게 된다. 이러한 경우에 본 논문에서 제안하는 부분영상 매칭 기법이 우편물 자동화 시스템의 박스 단위 영상 검증 단계에 유용하게 이용될 수 있다. 다음 4.1에서는 실제 이 시스템에 적용한 실험 방법 및 결과를 서술한다.

4.1 실험 영상

우편물 자동화 시스템의 박스 영상 단위 검증에 본 논문에서 제안한 방법들에 대한 성능을 평가하기 위해, 참조 영상과 테스트 영상의 겹치는 정도에 따라 4가지 유형의 데이터 집합을 정의한다. 즉, 실험 영상은 두 영상의 내용이 거의 유사한 경우와 참조 영상이 테스트 영상을 각각 70%, 50%, 30% 포함하는 경우로 나뉘어 구성된다. 실험 영상을 생성하기 위해, 스캔된 우편물 영상으로부터 수신자 주소 영역 즉 ROI를 추출하고, ROI 내의 텍스트 영상을 박스 단위로 선택한다. 동일한 우편물 영상을 다시 스캔하여 새로운 영상을 얻고, 위와 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 참조 영상과 테스트 영상으로 이용한다. 실험에 이용한 4가지 유형의 데이터 집합은 각각 50개, 39개, 58개, 44개 영상 쌍으로, 전체 191개 쌍의 실험 영상으로 구성된다. 거의 유사한 경우의 쌍을 “Ideal”로 나타내며, 70%, 50%, 30% 정도 일치하는 경우의 쌍을 각각 “Subset70”, “Subset50”, “Subset30”으로 표기한다. (그림 5)는 4가지의



(그림 5) 실험 영상의 예: (a) Ideal, (b) Subset70, (c) Subset50, (d) Subset30

데이터 집합에 대한 영상 쌍의 예를 보여준다.

4.2 메쉬 기반 방법의 실험 결과

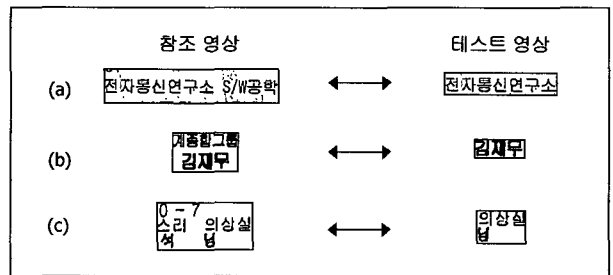
메쉬 기반 방법은 일치 영역 검색 단계와 유사성 측정 단계로 구성된다. 전체적인 인식 성능을 평가하기 전에 각 단계별 성능을 독립적으로 계산하고 에러를 분석함으로써, 개별적으로 각 단계가 전체 인식 성능에 미치는 영향을 알아보고자 한다.

4.2.1 일치 영역 검색 단계의 성능

일치 영역 검색 단계의 성능을 평가하기 위해 4가지 유형의 데이터 집합 - Ideal, Subset70, Subset50, Subset30을 이용하였다. <표 1>은 실험 결과를 보여준다. 2.1에서 설명한 방법으로 각 모서리마다 겹치는 흑화소의 수를 고려하여 가장 일치하는 모서리를 찾은 결과, 실험 영상 191개 중 175개를 정확히 맞추어 91.6%의 성능을 나타내었다. Ideal 영상의 경우 50개 영상 중 50개 모두를 정확히 맞추어 100%의 정확률을 보였다. Subset70 영상은 39개 영상 중 38개, Subset50 영상은 58개 중 51개, Subset30 영상은 44개 중 36개를 맞추어 각각 97.4%, 87.9%, 81.8%의 성능을 나타내었다. 이는 겹치는 부분이 적을수록 정확한 모서리를 찾기 가능성이 낮아지는 것을 보여준다.

<표 1> 일치 영역 검색 단계의 성능

결과	전체 영상	맞춘 영상	정확률
Ideal	50	50	100%
Subset70	39	38	97.4%
Subset50	58	51	87.9%
Subset30	44	36	81.8%
전체	191	175	91.6%

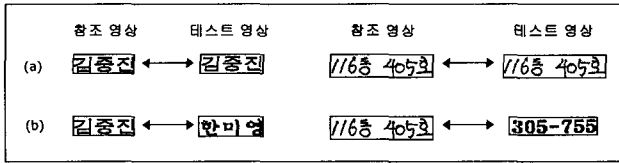


(그림 6) 일치 영역 검색 단계에서 오류 영상의 예

(그림 6)은 오류 영상 중 일부를 보여준다. (a)와 (b)는 제안 방법이 두 영상의 네 모서리만을 고려하기 때문에 영상의 가운데 부분에서 겹치는 경우에 발생하게 되는 오류이다. (c)는 일치하는 모서리의 대부분의 화소가 0값을 갖기 때문에 제안 방법에서 고려한 검은 화소를 이용한 특징으로는 검증할 수 없다는 한계점을 보여준다.

4.2.2 유사성 측정 단계의 성능

이 실험은 메쉬 특징의 유용성을 알아보기 위한 것으로 Ideal 영상 50개를 이용하여 동일한 영상의 쌍 50개 (Intra-pair)와 서로 다른 임의의 영상의 쌍 50개(Inter-pair)를 구성하였다. (그림 7)은 실험에 사용한 일부 영상을 보여 준다.

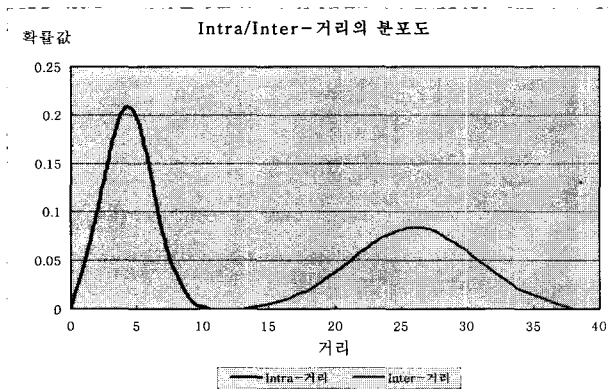


(그림 7) 실험 영상의 예: (a) Intra-pair, (b) Inter-pair

Intra-pair와 Inter-pair 내에 있는 실험 영상에 대해, 2.2에서 설명한 방법을 이용하여 두 영상에 대한 96차원 메쉬 특징을 추출하고 이들 간의 Manhattan 거리를 계산한다. <표 2>는 계산된 거리에 대한 평균 및 표준편차를 나타낸다. (그림 8)은 이에 대한 정규분포 곡선을 보여주고 있는데, 두 집합간의 메쉬 특징에 대한 거리의 분포가 서로 잘 분리됨을 알 수 있다. 따라서 메쉬 기반 방법에서 채택한 특징 추출 방법이 효과적이라 판단된다.

<표 2> Intra/Inter-거리에 대한 평균 및 표준편차

	Intra-거리	Inter-거리
평균(μ)	4.33	26.02
표준편차(σ)	1.91	4.75



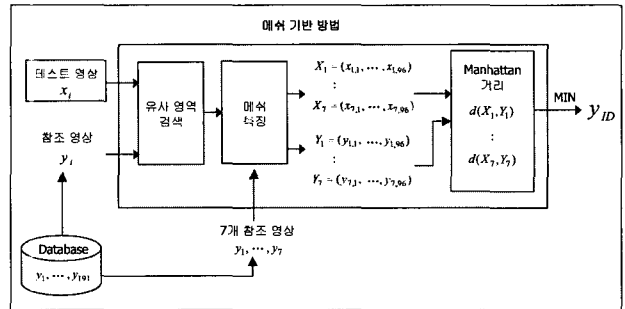
(그림 8) Intra/Inter-거리의 분포 곡선

4.2.3 메쉬 기반 방법의 전체 성능

앞에서 언급된 우편물 자동화 시스템은 하나의 우편 영상이 입력되었을 때 데이터베이스에 있는 7개의 참조 영상들과 비교하여 입력 영상과 가장 유사한 영상을 인식한다. 제안 방법의 성능을 평가하기 위한 실험에서도 우편물 자동화 시스템의 인식 방법과 동일하게 하나의 입력 영상과 데이터베이스 내의 7개 참조 영상을 비교하여 유사한 영상을 찾아

낸다. 선택된 7개의 참조 영상에는 입력된 테스트 영상과 동일한 쌍(4가지 데이터 집합 내의 영상 쌍 중 하나)이 포함되어 있다고 가정한다.

실험 과정은 (그림 9)와 같다. 메쉬 기반 방법은 일치 영역 검색 단계와 유사성 측정 단계로 구성된다. 첫 번째 단계는 입력된 테스트 영상(x_i)과 데이터베이스에 저장된 동일한 참조 영상(y_j) 쌍을 비교하여 일치하는 모서리를 계산한다. 데이터베이스 내의 7개 참조 영상(y_1, \dots, y_7)과 7개의 쌍을 구성하고, 식 (3)을 이용하여 이들 각각에 대한 96차원 메쉬 특징을 추출한다($X_1, Y_1, X_2, Y_2, \dots, X_7, Y_7$). 추출된 메쉬 특징에 대해 식 (4)를 이용하여 특징들 간의 거리를 계산하고, 이들 중 거리가 가장 최소인 참조 영상이 테스트 영상과 가장 유사하다고 판단한다.



(그림 9) 메쉬 기반 방법을 이용한 부분 영상 매칭 시스템

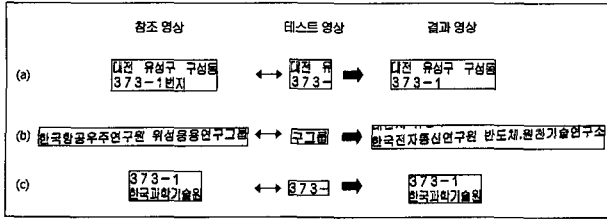
<표 3>은 4가지 유형의 데이터 집합을 이용하여 실험한 결과를 보여준다. 일치 영역 검색 단계에서는 전체 191개 중 175개를 정확히 인식하였고, 유사성 측정 단계에서는 175개 영상 중 172개를 정확히 인식하였다. 전체적으로는 191개 중 172개의 영상을 정확히 인식하여 90.1%의 성능을 나타내었다. 또한 메쉬 기반 방법을 이용할 경우, 1개 영상을 처리하는데 약 1.759초 정도가 소요되었다. 본 논문의 모든 실험은 Pentium IV 1.7GHz 프로세서를 이용하여 수행되었다.

<표 3> 패턴 매칭 방법의 실험 결과

수행 단계	결과	전체 영상	맞춘 영상	정확률	수행시간 (sec/image)
1단계 (일치 영역 검색)		191	175	91.6%	0.319
2단계 (유사도 측정)		175	172	98.3%	1.440
전체		191	172	90.1%	1.759

(그림 10)은 유사성 측정 단계에서의 오류 영상을 보여준다. (a)와 (c)는 테스트 영상이 기준에 쌍으로 구성된 참조 영상에 포함되지만, 인식 결과로 나온 결과 영상에도 역시 포함되는 것을 알 수 있다. 다시 말해, 답이 2개 있는 경우의 오류이다. (b)는 잘못 인식한 경우로서 화소의 밀도만을 고려한 메쉬 특징보다는 모서리나 윤곽선 같은 세부적인 특

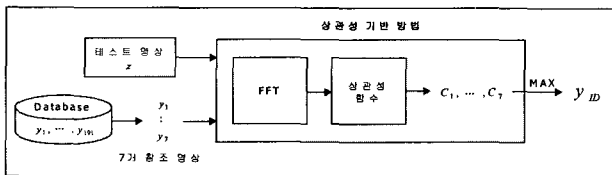
징을 고려한다면, 좀 더 좋은 성능을 얻을 수 있을 것이다.



(그림 10) 유사도 측정 단계에서 오류 영상의 예

4.3 상관성 기반 방법의 실험 결과

두 영상간의 일치 정도를 계산하기 위해 상관성 함수를 이용하는 상관성 기반 방법의 성능을 평가하기 위한 과정은 (그림 11)과 같다.



(그림 11) 상관성 기반 방법을 이용한 부분 영상 매칭 시스템

테스트 영상 x 는 데이터베이스 내의 7개 참조 영상들과 각각 쌍을 구성하여 상관성을 계산한다. 이 때에도 선택된 7개의 참조 영상에는 입력된 테스트 영상과 동일한 쌍이 포함되어 있다고 가정한다. 7개 쌍에 대한 상관 계수 값 중 최대값을 갖는 참조 영상을 테스트 영상과 가장 유사한 영상이라고 판별한다. FFT를 이용하여 상관성을 계산하는 방법은 3.2 및 <표 4>에 설명되어 있다. 단, FFT를 수행하기 전에 두 영상의 크기를 2의 지수승이 되도록 0 값을 설정해주는 제로 패딩 과정을 수행한다.

<표 4> 상관성 기반 방법을 구현하기 위한 MATLAB 프로그램

```

Reflmg = 참조 영상
Testlmg = 테스트 영상

M = 2^nextpow2( max(size(Reflmg,1),size(Testlmg,1)) ); % 제로 패딩
N = 2^nextpow2( max(size(Reflmg,2),size(Testlmg,2)) );
Reflmg(M, N) = 0;
Testlmg(M, N) = 0;

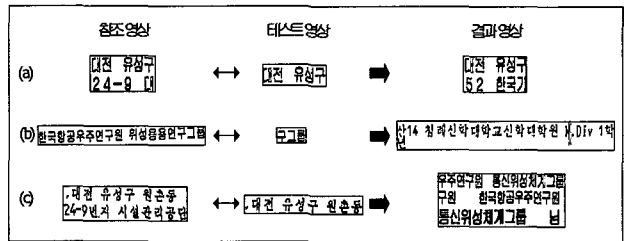
Fref = fft(Reflmg);           % 참조 영상에 대한 FFT
Ftest = conj(fft(Testlmg));   % 테스트 영상에 대한 FFT와 켈레
                               % 복소수 적용
Corr = real( ifft( Fref * Ftest ) ); % 상관성
Max_corr = max(Corr);        % 최대 상관성
    
```

<표 5>는 4.2에서 수행된 실험과 동일한 데이터를 이용하여 실험한 결과를 보여준다. 전체 191개의 영상 중 176개의 영상을 올바르게 인식하여 92.7%의 성능을 나타내었으며, 테스트 영상 1개를 처리하는 데 약 0.895초가 소요되었

다. 이러한 실험 결과를 통해, FFT를 이용한 상관성 기반 방법이 메쉬 기반 방법에 비해 약 2.6% 정도 인식 성능이 향상되었고, 수행시간 또한 1.759초에서 0.895초로 2배 정도 빨라지는 것을 알 수 있었다.

<표 5> 상관성 기반 방법의 인식 성능

	전체 영상	맞춘 영상	정확률	수행시간 (sec/image)
인식 결과	191	177	92.7%	0.895



(그림 12) 상관성 기반 매칭 방법에서의 오류 영상의 예

(그림 12)는 오류 영상 중 일부를 보여준다. (a)는 그림 10의 (a), (c)와 같이 테스트 영상이 참조 영상과 결과 영상 둘 다 매칭이 가능한 경우로, 답이 2개인 경우 발생하는 오류이다. (b)는 테스트 영상이 참조 영상의 가장자리 부분에 위치한 경우에 발생한 오류로서 영상의 가장자리 부분에서 정확도가 낮아지는 단점이 있다는 것을 알 수 있다. (c)는 테스트 영상이 참조 영상에 비해 상대적으로 큰 경우, 발생하는 오차의 양도 이에 비례하게 된다는 점을 보여준다. 이러한 단점을 극복하기 위해 영상에 대한 정규화 과정을 수행하여 특징 값을 -1에서 1 사이의 범위를 갖도록 변환하는 정규화된 상관성 함수를 이용하는 방법 또는 FFT 변환 후 계산된 진폭(amplitude) 성분과 위상(phase) 성분 중 위상 성분만을 취하여 연산함으로써 영상의 변화량(variation)에 좀더 로버스트(robust)하도록 만드는 위상 상관(phase correlation) 방법 등을 추가로 고려하는 것도 유용할 것으로 판단된다[15].

5. 결론 및 향후 연구 방향

지금까지 영상과 영상을 매칭하는 방법으로 메쉬 기반 방법과 상관성 기반 방법의 두 가지 영상 기반 매칭 방법을 제안하고 실제 우편 자동화 시스템에 이들 방법을 적용한 결과를 살펴보았다. 실험 결과, 메쉬 기반 방법은 191개 실험 영상 중 172개를 올바르게 인식하여 90.1%의 성능을 보였으며, 상관성 기반 방법은 191개 중 176개를 정확히 인식하여 92.7%의 인식 성능을 나타내었다. 또한 한 영상을 처리하는데 걸리는 수행시간도 약 0.864초 정도 단축할 수 있었다. 이는 FFT를 이용한 상관성 방법이 영상과 영상을 매칭하는 데 더 효과적임을 알 수 있다.

본 논문에서 다룬 텍스트 영상간의 매칭 시스템은 한 영

상이 단순히 이동되어 테스트 영상이 참조 영상에 일부 포함된 경우만을 고려하였다. 즉 이동에 의해 영상의 일부만이 일치하는 경우, 모서리를 찾아 겹치는 영역을 탐색하는 방법, 두 영상간의 상관성을 계산하는 방법이 효과적으로 적용될 수 있었다. 그러나 테스트 영상이 축소되거나 확대되는 경우, 또는 회전되는 경우에는 제안한 방법을 적용하기 어렵다. 향후 스케일링(scaling)되거나 회전된 영상을 매칭하는 데 유용한 새로운 방법을 연구하게 되면, 우편 자동화 시스템뿐만 아니라 핵심 단어 인식을 이용한 문서 영상 검색, 서명 검증, 웹 기반 문서영상 검색 분야에도 다양하게 이용할 수 있다.

참 고 문 헌

[1] A. Averbuch, 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, 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, D. P. Huijismans, "Multi-Scale Sub-Image Search," Proc. 7th ACM International Conference on Multimedia(Part II), pp.79-82, 1999.

[4] J. Luo, M. A. Nascimento, "Content Based Sub-Image Retrieval via Hierarchical Tree Matching," Proc. 1th ACM International Workshop on Multimedia Databases, pp.63-69, 2003.

[5] S. N. Srihari, Z. Shi, "Forensic Handwritten Document Retrieval System," International Workshop on Document Image Analysis for Libraries, pp.188-194, 2004.

[6] Y. Chen, J. Z. Wang, R. Krovetz, "Content-Based Image Retrieval by Clustering," Proc. 5th ACM SIGMM International workshop on Multimedia information retrieval, pp.193-200, 2003.

[7] H. Peng, F. Long, W. C. Sui, Z. Chi, D. D. Feng, "Document Image Matching Based on Component Blocks," Proc. 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," Proc. 5th International Conference on Asian Digital Libraries, pp.530-534, Dec. 2002.

[10] 진영범, 오일석, "대용량 인쇄 한글 문서 검색을 위한 영상 기반 단어 매칭 방법," 2000 한국정보과학회 추계 학술발표대회 논

문집, 제27권, 제2호, pp.461-463, 2000.

[11] R. C. Gonzales, R. E. Woods, "Digital Image Processing, 2nd Edition," Prentice Hall, 2002.

[12] J. D. Campbell, "Edge Structure and the Representation of Pictures," Ph.D. dissertation, Dept. of Elec. Eng., University of Missouri, Columbia, 1969.

[13] 김수형, 박상철 외 4명, Virtual ID 사용을 위한 우편 영상 검증 기술 연구 과제 중간보고서, 연구보고서, 한국전자통신연구원, 2004.

[14] 정창부, 박상철, 손화정, 김수형, "수취인 주소 영역의 영상매칭에 기반한 유사 우편물 추출 방법," 2004 한국정보과학회 추계 학술발표대회 논문집, 제31권, 제2호, pp.793-795, 2004.

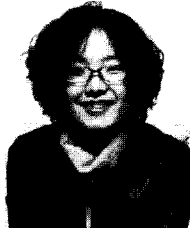
[15] J. P. Lewis, "Fast Normalized Cross-Correlation," Vision Interface, pp.120-123, 1995.

손 화 정



e-mail : sonhj@iip.chonnam.ac.kr
 2001년 전남대학교 통계학과 졸업(학사)
 2004년 전남대학교 전산학과 졸업(이학석사)
 2004년~현재 전남대학교 전산학과 박사 과정
 관심분야: 패턴인식, 문자인식, 영상처리

정 선 화



e-mail: sh-jeong@etri.re.kr
 1996년 전남대학교 통계학과 졸업(학사)
 1998년 전남대학교 전산통계학과 졸업(이학석사)
 2001년 전남대학교 전산통계학과 졸업(이학박사)

2001년~현재 한국전자통신연구원 우정기술연구센터 선임연구원
 관심분야: 패턴인식, 문자인식, 영상처리, 신경망 학습, 우정자동화

김 수 형



e-mail : shkim@chonnam.ac.kr
 1986년 서울대학교 컴퓨터공학과 졸업(학사)
 1988년 한국과학기술원 전산학과 졸업(공학석사)
 1993년 한국과학기술원 전산학과 졸업(공학박사)

1993년~1996년 삼성전자 멀티미디어연구소 선임연구원
 1997년~현재 전남대학교 전자컴퓨터정보통신공학부 부교수
 관심분야: 인공지능, 패턴인식, 문서영상 정보검색, 유비쿼터스 컴퓨팅