

저화질 영상 인식을 위한 화질 저하 모델 기반 다중 인식기 결합

류 상 진[†] · 김 인 중^{**}

요 약

본 논문에서는 화질 저하 모델에 기반한 다중 인식기 결합을 이용하여 저화질 영상에 대한 인식 성능을 개선하기 위한 방법을 제안한다. 제안하는 방법은 화질 저하 모델을 이용해 특정 화질에 각각 특화된 복수의 인식기들을 생성한다. 인식 과정에서는 인식기들의 결과를 가중 평균에 의해 결합함으로써 최종 결과를 결정한다. 이 때, 각 인식기의 가중치는 입력 영상의 화질 추정 결과에 따라 동적으로 결정된다. 입력 영상의 화질에 특화된 인식기에는 큰 가중치를, 그렇지 않은 인식기에는 작은 가중치를 지정한다. 그 결과, 입력 영상의 화질 변이에 효과적으로 적응할 수 있다. 뿐만 아니라, 복수의 인식기를 사용하기 때문에 저화질 영상에 대하여 단일 인식 시스템보다 더욱 안정적인 성능을 나타낸다. 제안하는 다중 인식기 결합 방법은 화질을 고려하지 않은 다중 인식기 결합 방법이나, 화질을 고려한 단일 인식 방법과 비교하여 더 높은 인식률을 보였다.

키워드 : 저화질 영상 인식, 다중 인식기 조합, 화질 저하 모델, 문자 인식

Multiple-Classifier Combination based on Image Degradation Model for Low-Quality Image Recognition

RYU, SANG-JIN[†] · KIM, IN-JUNG^{**}

ABSTRACT

In this paper, we propose a multiple classifier combination method based on image degradation modeling to improve recognition performance on low-quality images. Using an image degradation model, it generates a set of classifiers each of which is specialized for a specific image quality. In recognition, it combines the results of the recognizers by weighted averaging to decide the final result. At this time, the weight of each recognizer is dynamically decided from the estimated quality of the input image. It assigns large weight to the recognizer specialized to the estimated quality of the input image, but small weight to other recognizers. As the result, it can effectively adapt to image quality variation. Moreover, being a multiple-classifier system, it shows more reliable performance than the single-classifier system on low-quality images. In the experiment, the proposed multiple-classifier combination method achieved higher recognition rate than multiple-classifier combination systems not considering the image quality or single classifier systems considering the image quality.

Keywords : Low-Quality Image Recognition, Multiple Classifier Combination, Image Degradation Model, Character Recognition

1. 서 론

휴대용 카메라의 보급이 크게 증가함에 따라 카메라로 촬영한 문서 영상을 인식하기 위한 시도가 많이 이루어지고 있다. 그러나, 카메라로 취득한 영상의 경우 제한된 해상도나 통제되지 않은 촬영 환경 등의 이유로 영상의 화질이 저하되는 경우가 많다. 이를 극복하기 위해서는 저화질 영상 인식

※ 이 논문은 2007년 정부(교육과학기술부)의 재원으로 한국연구재단(구 학술진흥재단)의 지원을 받아 수행된 연구임(KRF-2007-331-D00420)

※ 본 연구는 교육과학기술부와 한국산업기술재단의 지역혁신인력양성사업으로 수행된 연구결과임

† 준 회원 : 한동대학교 정보통신공학과 석사과정

** 중신회원 : 한동대학교 교수

논문접수 : 2010년 1월 8일

수정일 : 1차 2010년 2월 16일, 2차 2010년 3월 23일

심사완료 : 2010년 4월 27일

을 위한 기술의 발전이 요구된다. 저화질 문서 영상을 인식하기 위한 방법은 크게 두 가지로 나눌 수 있다. 저화질 입력 영상으로부터 고화질 영상을 복원한 후 인식하는 방법[1, 2]과 저화질 영상을 인식기 학습에 참여시킴으로써 화질 저하에 대한 인식기의 적응력을 높이는 방법[3-5]이다.

본 논문에서는 다중 인식기 결합을 통해 화질 변이에 효과적으로 적용할 수 있는 인식 방법을 제안한다. 다중 인식 시스템은 많은 연구에서 단일 인식기보다 우수한 성능을 보여왔다[6, 7]. 제안하는 시스템은 동일한 방법론을 사용하되, 서로 다른 화질의 데이터로 학습됨으로써 차별화 된 복수의 인식기로 구성된다. 이러한 인식기들은 동일한 범위와 스케일의 클래스별 점수, 또는 거리를 출력하면서도 서로 강한 상보성을 갖는다. 인식 과정에서 인식기들의 결과를 가중 평균하여 최종 인식 결과를 결정한다. 이 때, 입력 영상과 유사한 화질의 데이터로 학습된 인식기에는 높은 가중치를, 그렇지 않은 인식기에는 낮은 가중치를 지정한다. 입력 영상의 화질에 따라 인식기들의 가중치를 동적으로 결정하기 때문에 화질 변이에 효과적으로 적용할 수 있다.

본 논문의 2 장에서는 저화질 문서 영상 인식 및 화질 저하 모델에 관한 기존 연구들을 소개한다. 3 장에서는 본 논문에서 제안하는 화질 저하 모델 기반 다중 인식기 결합 방법에 대해 설명하고, 4 장에서는 실험 결과를 제시한다. 마지막으로 5 장에서는 결론을 맺는다.

2. 관련 연구

2.1 저화질 문서 영상 인식

저화질 영상에 대한 인식기의 적응력을 향상시키기 위해 많이 사용되는 방법은 저화질 영상을 인식기 학습에 참여시키는 것이다. Ishida는 고화질의 명도 영상에 다양한 강도의 블러링과 서브샘플링을 적용하여 다양한 화질의 학습 영상을 생성하였다[3]. 그리고 이를 이용하여 인식기를 학습시킴으로써 저화질 영상에 대한 인식기의 적응력을 개선하였다. 고화질 영상만을 학습에 참여시켰던 기존의 방법에 비해 상당한 개선 효과가 있었다. 그러나, 단일 인식기만으로 다양한 화질 변이를 모두 흡수하기에는 한계가 있다. 따라서, 저화질 영상에 대한 적응력을 강화하기 위하여 저화질 영상을 학습에 많이 참여시킬 경우 고화질 영상에 대한 성능이 저하될 수 있다.

Omachi는 이진 영상에 대한 화질 저하의 영향을 모델링함으로써 화질 저하에 대한 인식기의 적응력을 개선하였다[4]. Omachi는 입력 영상과 각 클래스의 유사도를 Mahalanobis 거리를 이용하여 계산하였는데, 이진 영상의 화질에 따라 공분산 행렬을 조정함으로써 화질 저하에 대한 적응력을 개선하였다. Omachi의 방법은 인식 대상이 이진 영상인 인식 시스템에 적합하다. 그런데, 인식 대상이 명도 영상일 경우 화질 저하가 있더라도 영상에 남아있는 정보의 양이 이진 영상보다 많다. Omachi의 방법은 이와 같은 명도 영상의 장점을 활용하지 않는다.

Liu는 명도 영상의 화질을 명도 히스토그램 분석을 통하여

상중하 세 단계로 분류한 후 각 화질에 특화된 인식기로 인식하는 방법을 제안하였다[5]. 학습 단계에서는 학습 영상을 화질에 따라 세 단계로 나누어 세 개의 서로 다른 인식기를 학습하였다. 인식 단계에서는 입력 영상의 화질을 추정한 후 그 결과에 따라 해당 인식기를 선택하여 인식하였다. Liu의 시스템은 특정 화질에 특화된 복수의 인식기로 구성되었으나, 실제 인식에는 하나의 인식기만을 사용하고 그 외의 인식기들은 사용하지 않기 때문에 다중 인식 시스템의 장점을 충분히 활용하지 않았다.

2.2 화질 저하 모델

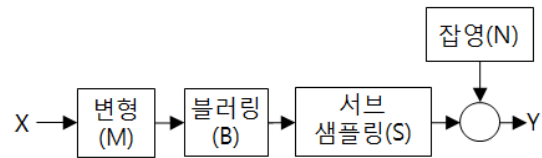
카메라를 이용해 영상을 취득하는 과정에는 다양한 왜곡 및 블러링에 의한 화질 저하가 발생한다. 이와 같은 과정은 화질 저하 모델에 의해 표현할 수 있다. 널리 사용되는 화질 저하 모델은 (그림 1)과 같다[1]. 촬영 대상이 되는 물체, 또는 장면을 X 라 하고 카메라에 의해 취득되는 영상을 Y 라 하자. X 로부터 Y 가 생성되는 과정에는 변형, 블러링, 서브샘플링 등의 화질 저하가 발생하며, 잡영도 포함된다.

(그림 1)의 과정을 수식으로 나타내면 식 (1)과 같다.

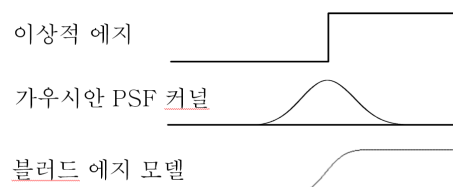
$$Y = S \cdot B \cdot M \cdot X + N \tag{1}$$

여기에서 M 은 왜곡, B 는 블러링, S 는 서브샘플링, N 은 잡영을 각각 나타낸다. 그런데, 이와 같이 다양한 형태의 화질 저하 중 인식 성능 저하에 가장 큰 원인이 되는 것은 블러링이다. 따라서, 본 연구에서는 블러링에 의한 화질 저하에 중점을 두었다.

블러링은 주로 PSF(point spread function) 커널에 대한 컨볼루션으로 모델링된다[2, 8]. Chiang은 블러링을 (그림 2)과 같이 가우시안 커널을 이용한 컨볼루션으로 모델링하였다[8]. 원 영상이 이상적인 에지로 이루어졌다고 가정할 후 블러링이 포함된 입력 영상을 이상적 에지와 가우시안 PSF 커널의 컨볼루션으로 생성된 블러드 에지 모델(blurred edge model)로 나타내었다. 이 때 화질 저하의 정도는 가우시안 커널의 표준편차에 의해 표현된다.



(그림 1) 화질 저하 모델



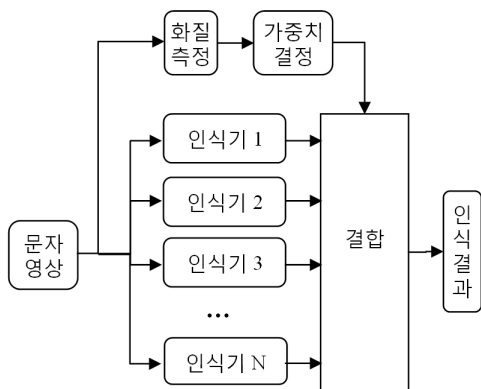
(그림 2) Chiang의 블러링 모델

3. 화질 저하 모델 기반 다중 인식기 결합에 의한 저화질 영상 인식

3.1 시스템 개요

인식기를 특정 화질의 데이터로 학습시키면 학습 데이터와 유사한 화질의 입력 영상에 대한 적응력을 개선할 수 있다. 그러나, 단일 인식기만으로 다양한 화질 문자 영상을 모두 인식하기에는 한계가 있다. 또한, 화질 저하가 심한 영상의 경우 해당 화질에 적용된 인식기라 해도 좋은 성능을 얻기는 어렵다. 이와 같은 문제는 각각 특정 화질에 적용된 복수의 인식기를 결합함으로써 개선될 수 있다. 따라서, 본 논문에서는 다중 인식기 결합을 이용한 저화질 영상 인식 방법을 제안한다.

제안하는 방법은 (그림 3)와 같다. 인식 시스템은 특정 화질에 특화된 다수의 인식기로 구성된다. 영상이 입력되면 각 인식기를 이용하여 인식을 수행한 후, 그 결과를 결합하여 최종 인식 결과를 출력한다. 각 인식기는 특정 화질에 특화되어 해당 화질의 영상에 대해서는 우수한 성능을 갖는다. 결합 과정에서는 다중 인식기의 클래스 별 인식 결과들을 조합하여 입력 영상의 화질을 추정한 후, 이를 이용하여 각 인식기에 가중치를 지정한다. 입력 영상의 화질과 유사한 화질에 특화된 인식기에는 높은 가중치를, 그렇지 않은 인식기에는 작은 가중치를 지정한다. 제안하는 방법은 결합에 참여하는 모든 인식기가 동일한 방법론을 사용하며 학습 가능한 방법일 경우 적용 가능하며, 인식기의 종류나 개수 N 에 제한을 받지 않는다.



(그림 3) 화질 저하 모델 기반 다중 인식기 결합

3.2 화질 별 인식기 생성

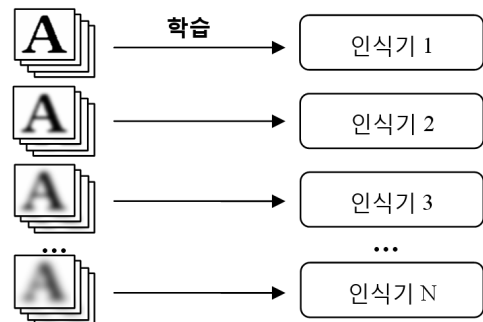
최근에는 인식기를 구현하기 위하여 신경망, 통계적 인식기, SVM 등 학습 데이터로부터 기계 학습이 가능한 알고리즘들이 많이 사용된다. 이러한 인식기들의 특성 및 성능은 그 인식기의 학습 과정에 사용한 샘플 데이터의 성질에 크게 영향을 받는다. 본 연구에서는 이러한 점을 이용하여 (그림 4)과 같이 각 인식기들을 각각 서로 다른 화질의 학습 영상으로 학습함으로써 특정 화질에 특화된 인식기들을 생성하였

다. 결합에 참여하는 인식기들은 모두 동일한 학습 및 인식 방법론을 사용한다. 그러나, 서로 다른 화질의 데이터로 학습되어 서로 다른 특성을 갖는다. 즉, 고화질 영상으로 학습된 인식기는 고화질 영상에 특화된 성능을 갖고, 저화질 영상으로 학습된 인식기는 저화질 영상에 특화된 성능을 갖는다.

다양한 화질의 학습 영상을 수집하는 것은 현실적으로 쉽지 않은 일이다. 그러나, 화질 저하 모델을 이용하면 고화질 영상으로부터 다양한 화질의 학습 데이터를 생성할 수 있다 [2]. 가우시안 PSF 커널을 이용한 화질 저하 모델을 사용할 경우 커널의 표준편차를 이용하여 화질 저하의 정도를 조절할 수 있다. 그런데, 인식에 있어서 실질적인 화질 저하의 영향은 커널 뿐 아니라 입력 영상의 크기와도 관계가 있다. 동일한 커널을 사용할 때 실질적인 화질 저하의 영향은 입력 영상의 크기에 반비례한다. 즉, 동일한 정도로 화질이 저하된다 하더라도 입력 영상이 작은 경우는 인식 성능이 급격히 저하되지만, 입력 영상이 큰 경우는 별로 영향을 받지 않는다. 따라서, 본 연구에서는 가우시안 커널의 표준편차대신 표준편차와 입력 영상의 크기에 대한 비를 이용하여 화질을 나타내었다. 이는 다음과 같은 식으로 표현된다.

$$b_Y = \alpha/h_X \tag{2}$$

b_Y 는 입력 영상 Y 의 실질적 화질 저하도를, α 와 h_X 는 가우시안 커널의 표준편차와 원본 영상 X 의 높이를 각각 나타낸다. 고화질 학습 영상 X 에 대하여 b_Y 만큼 화질을 저하시키기 위한 커널의 표준편차 α 는 b_Y 와 h_X 의 곱으로 구할 수 있다.



(그림 4) 화질별 인식기 생성

3.3 다중 인식기 결합

신경망, 통계적 인식기 등 학습 가능한 인식기들은 입력 영상이 주어졌을 때 각 클래스별로 점수, 또는 거리의 의미를 갖는 측정치(measurement)를 출력한다. 단일 인식 시스템에서는 이 값이 점수일 경우 최대값을 갖는 클래스를, 거리일 경우 최소값을 갖는 클래스를 인식 결과로 선택한다.

다중인식기 결합 방법은 크게 추상화(abstraction) 수준의 결합, 순위(rank) 수준의 결합, 측정치 수준의 결합으로 분류할 수 있다[6]. 이 중 측정치 수준의 결합은 성능이 가장 우수하다. 그러나, 인식기들의 측정치가 유사한 범위 및 스케일

을 가질 경우에 한해 직접 적용할 수 있으며 그렇지 않을 경우 별도의 측정치 정규화 과정을 필요로 한다. 제안하는 방법과 같이 동일한 방법론을 사용하되 학습 데이터를 차별화함으로써 상보성을 확보한 인식기들은 유사한 범위와 스케일을 갖는 측정치를 출력하기 때문에 별도의 정규화 과정 없이도 측정치 수준의 결합이 가능하다.

다중 인식 시스템에 영상이 입력되면 각 인식기는 출력한 클래스 별 점수(또는 거리)를 출력한다. j 번째 인식기가 출력한 c 번째 클래스의 점수를 $s_j(c)$ 라고 하자 ($1 \leq j \leq N, 1 \leq c \leq C$). N 은 결합에 참여하는 인식기의 수, C 는 인식의 대상이 되는 클래스의 수를 나타낸다.

입력 영상이 주어졌을 때 각 인식기의 정확도는 입력 영상과 그 인식기의 학습 데이터 간 화질의 유사도에 비례한다. 따라서, 본 연구에서는 각 인식기에 입력 영상의 추정된 화질과 그 인식기가 학습된 화질 간의 유사도에 의해 가중치를 지정한 후 그 가중 평균을 통해 인식기들의 결과를 결합하였다. 이 때, 가중치는 가우시안 가중치 함수를 이용하여 결정하였다[9]. 입력 영상 Y_i 의 화질을 b_i 라 하고, j 번째 인식기가 특화된 화질을 b_j 라 하면 j 번째 인식기의 가중치 w_j 는 식 (3) 같이 계산한다.

$$w_j = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(b_i - b_j)^2}{2\sigma^2}\right] \quad (3)$$

식 (3)에서 가우시안 분포의 표준편차 σ 는 실험적으로 결정할 수 있다. 본 연구에서는 사전 실험을 통해 높은 인식율을 나타내는 값을 선택하여 사용하였다.

각 인식기의 가중치가 결정되면 다음 식을 이용하여 인식 결과 c^* 를 결정한다.

$$c^* = \arg \max_c \sum_{j=1}^N w_j s_j(c) \quad (4)$$

$s_j(c)$ 가 거리 등 부정적인 척도일 경우에는 $\arg \max$ 대신 $\arg \min$ 을 사용한다. Liu의 방법이 여러 인식기 중 한 개만을 선택하여 사용하는 것에 반해 제안하는 방법은 모든 인식기의 결과를 활용하기 때문에 더 우수한 결과를 얻을 수 있다.

입력 영상의 화질을 추정하는 방법은 여러 가지가 있다. Chiang는 입력 영상의 지역 에지가 이상적 에지와 가우시안 PSF의 컨볼루션에 의해 생성되었다고 가정하고 입력 영상으로부터 가우시안 PSF를 역으로 추정하였다[8]. 이 방법은 가우시안 PSF를 사용한 화질 저하 모델에 적합하지만, 복잡도가 높아 실제 인식 시스템에 사용할 경우 속도가 많이 저하된다. Liu는 학습 데이터를 상중하 3종류의 화질로 구분하고 각 화질의 데이터로부터 명도 히스토그램을 추출한 후, 이를 이용하여 참조 특징을 생성하였다. 인식 단계에서는 입력 영상의 명도 히스토그램으로부터 추출한 특징을 참조 특징과 비교함으로써 입력 영상의 화질을 추정하였다[5].

본 연구에서는 특정 화질에 특화된 인식기가 학습 영상과 유사한 화질의 입력 영상에 대하여 높은 점수를 출력하는 점

을 이용하여 각 인식기의 클래스 별 점수로부터 화질을 추정하였다. 결합에 참여하는 인식기들이 각각 특정 화질에 특화되었기 때문에, 입력 영상에 대하여 전체적으로 높은 점수를 출력하는 인식기는 입력 영상의 화질과 가장 가까운 화질에 특화된 인식기일 가능성이 높다. 따라서, 입력 영상에 대하여 높은 점수를 출력하는 인식기를 선택한 후, 그 인식기가 학습된 화질을 입력 영상의 추정 화질로 선택하였다.

모든 조합의 인식기 j 및 클래스 c 에 대한 $s_j(c)$ 는 <표 1>과 같이 $N \times C$ 크기의 행렬로 나타낼 수 있다. 입력 영상에 대하여 가장 높은 점수를 출력하는 인식기는 $s_j(c)$ 중 큰 값을 갖는 K 개의 원소들에 대하여 순위를 반영한 가중 투표에 의해 결정한다. <표 1>의 행렬 중 큰 값을 갖는 원소 K 개를 $s_{j(k)}(c(k))$ ($1 \leq k \leq K$)라고 하자. 여기에서 $j(k)$ 와 $c(k)$ 는 각각 k 번째 원소의 행(인식기)과 열(클래스)를 의미한다. 입력 영상의 화질에 가장 적합한 인식기, 즉 화질 추정치 b_i 는 다음 식에 의해 결정된다.

$$\begin{aligned} score(b_j) &= \sum_{s_{j(k)}(c(k)), s.t. j(k)=j} (K+1-k) \\ b_i &= \arg \max_{b_j} score(b_j) \end{aligned} \quad (5)$$

식 (5)에 의해 결정된 b_i 는 식(3)의 가중치 계산에 사용된다. 제안하는 방법은 별도의 영상 분석이나 별도의 특징 추출을 필요로 하지 않고, 각 인식기들의 출력만을 이용해 화질을 추정하기 때문에 적은 계산량으로도 입력 영상의 화질을 효과적으로 추정할 수 있다.

<표 1> 각 인식기의 클래스별 점수 $s_j(c)$

	클래스 1	클래스 2	클래스 3	...	클래스 C
인식기 1	$S_1(2)$	$S_1(3)$	$S_1(4)$...	$S_1(C)$
인식기 2	$S_2(2)$	$S_2(3)$	$S_2(4)$...	$S_2(C)$
인식기 3	$S_3(2)$	$S_3(3)$	$S_3(4)$...	$S_3(C)$
...
인식기 N	$S_N(2)$	$S_N(3)$	$S_N(4)$...	$S_N(C)$

4. 실험

제안하는 방법의 실제적 효과를 평가하기 위하여 한글 인식에 적용하여 성능을 측정하였다. 실험 데이터로는 ETRI 문자 영상 데이터의 인쇄 한글 영상을 사용하였다[11]. ETRI 데이터는 한글 2350클래스에 대하여 클래스당 300자의 샘플로 구성되어 있다. 이 중 사용 빈도가 높은 문자 클래스 520개를 실험에 사용하였으며 학습 데이터와 테스트 데이터는 2:1의 비율로 나누어 사용하였다. 실험은 Q6600 2.4MHz CPU와 4GB 메모리를 갖는 PC에서 Visual Studio 6.0을 이용해 실시하였다.







다양한 화질의 영상을 생성하기 위하여 식(2)의 b_j 를 0.01에서 0.1까지 0.01의 간격으로 변화시키며 각 원본 영상에 블러링을 가하였다. 원본 영상을 포함하여 모두 11단계의 화질

이 사용되었다. <표 2>는 이와 같이 생성한 영상 데이터의 예이다.

그러나, 학습 데이터의 화질을 지나치게 세분화할 경우 복잡도가 증가하고 화질 측정의 정확도가 저하될 뿐 아니라, 각 인식기가 지나치게 화질에 민감해 진다. 따라서 본 실험에서는 11단계의 화질을 소수의 그룹으로 결합하여 다중 인식기 구성에 사용하였다. 사전 실험을 통해 그룹의 수를 변화시키며 성능을 측정한 결과 3개의 그룹으로 나누었을 때 가장 높은 성능을 보였다. 이 때, 각 그룹은 b_{γ} 가 각각 0~0.02, 0.03~0.06, 0.07~0.10인 데이터들로 구성된다.

제안하는 방법을 다양한 인식 방법에 적용하였을 때의 효과를 측정하기 위하여 문자 인식에 널리 사용되는 LDF (Linear Discriminant Function)와 MLP(Multi-Layer Perceptron) 등 두 가지 인식 방법을 이용하여 실험하였다[12,13]. LDF의 경우 인식 성능의 향상을 위하여 LVQ (Learning Vector Quantization)를 적용해 유사 문자에 대한 구분력을 강화하였다[14]. LDF는 부정적 척도인 클래스별 거리를 출력하기 때문에, 식(4)에서 argmax 대신 argmin 을 사용하였다. 특징 추출 단계에는 문자 인식에 성능이 우수하다고 알려진 NCGF (normalization-cooperated gradient feature)을 이용하였다[15]. 특징 벡터는 8가지 방향에 대하여 8 x 8 그물망(mesh)을 이용해 추출된 512 차원의 벡터를 사용하였다.

<표 2> 실험에 사용한 영상 데이터의 화질

					
원본	$b_{\gamma}=.01$	$b_{\gamma}=.03$	$b_{\gamma}=.05$	$b_{\gamma}=.07$	$b_{\gamma}=.10$

4.1 다양한 화질의 입력 영상에 대한 각 인식기 성능

먼저 다양한 화질의 테스트 영상에 대한 각 인식기의 성능을 파악하기 위하여 여러 화질의 테스트 데이터들을 각 인식기에 입력하여 성능을 측정하였다. 실험 결과는 <표 3>과 같다. 각 행은 인식기가 적용된 화질을, 각 열은 인식률 측정에 사용된 테스트 영상의 화질을 나타낸다.

각 화질의 테스트 영상에 대하여 동일한 화질에 적용된 인식기는 매우 높은 인식률을 나타내었다. 그러나, 다른 화질에 적용된 인식기에서는 인식률이 크게 저하되었다. 이러한 인식률의 저하는 실험자의 예상보다도 훨씬 급격하였다. 이는 인식 성능에 있어서 화질 변화의 영향이 매우 크다는 것을 보여준다. 그러므로, 다양한 화질의 입력 영상에 대하여 좋은 인식 성능을 얻기 위해서는 입력 영상의 화질과 인식기가 적

<표 3> 입력 영상의 화질 저하 정도에 따른 인식기 별 성능

		LDF			MLP		
		테스트 영상 화질			테스트 영상 화질		
		1	2	3	1	2	3
인식기 적용 화질	1	99.71	77.27	15.59	95.19	67.9	12.89
	2	46.49	99.72	76.2	88.06	94.63	45.25
	3	1.9	70.63	99.57	18.53	60.81	85.13

용된 화질 간의 상관 관계가 반드시 고려되어야 함을 확인할 수 있었다.

4.2 다중 인식기 결합 방법에 따른 인식 성능

제안하는 결합 방법의 성능을 평가하기 위하여 다양한 인식기 결합 방법들과 비교하였다. 각 인식기에 적용한 인식 방법, 특징 추출 방법, 화질의 단계 수 등은 모두 동일하다. 실험에 사용된 인식기 결합 방법은 다음과 같다.

- 입력 영상의 화질을 반영하지 않은 방법
 - A. 각 인식기의 1위 후보만을 이용한 투표[16]
 - B. 각 인식기의 복수 후보를 이용한 다중 투표[16]
 - C. 모든 인식기가 출력한 클래스 별 측정치 중 최대값
- 입력 영상의 화질에 따른 가중치를 적용한 결합 방법
 - D. 각 인식기의 1위 후보에 가중치를 적용한 투표
 - E. 각 인식기의 복수 후보에 가중치를 적용한 다중 투표
 - F. 제안하는 방법

화질 추정을 위한 상수 K 와 가중치 결정을 위한 식 (3)의 상수 σ 는 사전 실험에 의해 좋은 성능을 보인 값을 선택하였다. 실제 사용된 값은 LDF에 대하여는 $K=3, \sigma=2.0$, MLP에 대하여는 $K=3, \sigma=0.5$ 이다. 또한, 제안하는 방법을 화질 적용을 이용한 단일 인식 방법들과 비교하기 위하여 다음과 같은 방법과 비교하였다. 이 중 G는 방법론적으로 Ishida의 방법과 동일하며 H는 화질 추정 방법을 제외하면 Liu의 방법과 동일하다[3,5]. 다만, 다양한 결합 방법을 비교하는 과정에서 실험 환경을 통일하기 위하여 학습 데이터의 조합을 다소 수정하였다.

- 화질 적용을 적용한 단일 인식 방법
 - G. 모든 화질의 데이터로 학습된 단일 인식기[3]
 - H. 입력 영상의 화질에 가장 가까운 화질에 적용된 단일 인식기[5]

각 결합 방법 별 인식 성능의 측정 결과는 <표 4> 및 <표 5>와 같다. LDF가 MLP에 비하여 우수한 성능을 보였다. 두 가지 인식 방법 모두에서 6가지 다중 인식기 결합 방법 중 제안하는 방법이 가장 우수한 성능을 나타내었다. 제안하는 방법은 또한, 화질을 고려한 단일 인식 방법들보다도 우수한 성능을 나타내었다. 두 방법 중 높은 성능을 보였던 LDF의 경우 제안하는 방법의 에러율을 단일 인식 기반 기존 방법들인 G, H와 비교하면 각각 81.55%, 59.57%의 오류 감소율을 보였다.

<표 4> 결합 방법에 따른 인식 성능 (LDF)

	1	2	3	평균	오류율	각 방법론 대비 오류감소율
A	80.9	95.4	82.1	86.11	13.89	95.90%
B	82.1	96.3	85.5	87.95	12.05	95.27%
C	98.9	97.3	99.6	98.59	1.41	59.57%
D	96.3	97.6	99.5	97.76	2.24	74.55%
E	92.2	97.8	93.8	94.58	5.42	89.48%
F	99.2	99.6	99.5	99.43	0.57	-
G	99.0	98.9	92.9	96.91	3.09	81.55%
H	98.9	97.3	99.6	98.59	1.41	59.57%

〈표 5〉 결합 방법에 따른 인식 성능 (MLP)

	1	2	3	평균	오류율	각 방법론 대비 오류감소율
A	88.8	88.0	54.6	77.13	22.87	60.79%
B	87.5	87.9	60.1	78.50	21.50	58.29%
C	94.4	92.7	85.0	90.70	9.30	3.58%
D	94.4	92.7	85.0	90.70	9.30	3.58%
E	94.4	92.6	85.0	90.67	9.33	3.93%
F	94.8	93.3	85.0	91.03	8.97	-
G	94.2	93.3	76.7	88.07	11.93	24.86%
H	95.7	93.6	81.5	90.27	9.73	7.88%

5. 결 론

본 논문에서는 화질 저하 모델에 기반한 다중 인식기 결합을 이용하여 저화질 영상을 효과적으로 인식할 수 있는 방법을 제안하였다. 화질 저하 모델에 기반하여 다양한 화질의 학습 데이터를 생성한 후 이를 이용하여 인식기들을 학습함으로써 특정 화질에 각각 특화된 다수의 인식기를 생성하였다. 결합 과정에서는 입력 영상의 화질을 반영하여 각 인식기의 가중치를 결정하였다. 제안하는 다중 인식기 생성 및 결합 방법은 인식기 상호간에 상보성을 강화할 뿐 아니라, 다양한 화질 변이에도 잘 적응할 수 있어 저화질 영상을 효과적으로 인식할 수 있다.

제안하는 방법의 효과를 검증하기 위하여 ETRI 데이터에 포함된 한글 520 클래스의 문자 영상에 대하여 실험을 실시하였다. 다양한 화질의 영상이 입력될 경우 각 인식기의 성능은 그 인식기가 학습된 화질에 매우 크게 영향을 받는 것으로 나타나 인식 과정에 입력 영상의 화질이 반드시 고려되어야 함을 정량적으로 확인하였다. 또한, 다양한 결합 방법을 비교 평가한 결과 본 논문에서 제안하는 다중 인식기 결합 방법이 기존의 다중 인식기 결합 방법이나 화질 적응을 적용한 단일 인식 방법 보다 우수한 성능을 나타내었다.

참 고 문 헌

[1] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *Signal Processing Magazine, IEEE*, Vol.20, No.3, pp.21-36, 2003.

[2] T. Boulton and G. Wolberg. "Local Image Reconstruction And Sub-Pixel Restoration Algorithms," *CVGIP: Graphical Models and Image Processing*, Vol.55, pp.63-77, Jan 1993.

[3] H. Ishida, S. Yanadume, T. Takahashi, I. Ide, Y. Mekada and H. Murase. "Recognition of low-resolution characters by a generative learning method," *Pattern Recognition* Vol.41, Issue 7, pp.2253-2262, 2008.

[4] S. Omachi, F. Sun, and H. Aso. "Precise Recognition of Blurred Chinese Characters by Considering Change in Distribution," *IEEE TPAMI*, Vol.22, No.3, pp.314-319, 2000.

[5] C. Liu, C. Wang, and R. Dai. "Low Resolution Character Recognition by Image Quality Evaluation," *Proc. of ICPR2006*,

pp.864-867, Hongkong, 2006.

[6] L. Xu, A. Krzyzak, and C. Y. Suen, *Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition*. *IEEE Transaction on Systems, Man, and Cybernetics*, Vol.22, No.3, pp.418-435, 1992.

[7] C. L. Liu, "Normalization-Cooperated Gradient Feature Extraction for Handwritten Character Recognition," *IEEE TPAMI*, Vol.29, No.8, pp.1465-1469, 2007.

[8] M. Chiang, T. Boulton. "Local Blur Estimation and Super-Resolution," *proc. of CVPR1997*, pp.821-826, 1997.

[9] <http://people.csail.mit.edu/hasinoff/320/sliding-notes.pdf>

[10] T.G. Ditterich, "Ensemble methods in machine learning, Multiple Classifier Systems," *LNCS*, Vol.1857, pp.1-15, Springer, 2000.

[11] KAIST 인공지능 연구실 홈페이지(<http://ai.kaist.ac.kr/Resource/dbase/Image%20Database.htm#HangulCharacter>)

[12] R.O. Duda, P.E. Hart and D.G. Stork, "Pattern Classification 2nd ed.," pp.36-45, Wiley-Interscience, 2001.

[13] http://en.wikipedia.org/wiki/Multilayer_perceptron

[14] C. L. Liu, I. J. Kim, and J. H. Kim, "High Accuracy Handwritten Chinese Character Recognition by Improved Feature Matching Method," *Proc. 4th ICDAR, Ulm, Germany*, pp.1033-1037, 1997.

[15] C. L. Liu, "Normalization-Cooperated Gradient Feature Extraction for Handwritten Character Recognition," *IEEE TPAMI*, Vol.29, No.8, pp.1465-1469, 2007.

[16] A. F. R. Rahman, H. Alam and M. C. Fairhurst, *Multiple Classifier Combination for Character Recognition: Revisiting the Majority Voting System and Its Variations*, *LNCS* Vol. 2423, pp.167-178, 2002.



류 상 진

e-mail : pardon6@naver.com
 2009년 한동대학교 전산전자공학부(학사)
 2010년~현재 한동대학교 정보통신공학과 석사과정



김 인 중

e-mail : ijkim@handong.edu
 1994년 KAIST 전산학과(학사)
 1995년 KAIST 전산학과(공학석사)
 2001년 KAIST 전산학과(공학박사)
 2001년~2005년 (주)인지소프트 책임연구원
 2006년~현재 한동대학교 교수

관심분야 : 패턴인식, 영상처리, 인공지능, 모바일 프로그래밍, 증강현실 등