

# 필기 숫자의 기계 인식을 위한 인간의 필기 숫자 인식 실험에 대한 고찰

## A Study on Human Recognition Experiments with Handwritten Digit for Machine Recognition of Handwritten Digit

윤성수\*, 정현숙\*\*, 이광오\*\*\*, 이일병\*\*\*\*, 이상호\*

Sungsoo Yoon\*, Hyunsook Chung\*\*, Kwangoh Yi\*\*\*, Yillbyeong Lee\*\*\*\* and Sangho Lee\*

\* 이화여자대학교 공과대학 컴퓨터학과

\*\* 조선대학교 컴퓨터공학부

\*\*\* 영남대학교 문과대학 심리학과

\*\*\*\* 연세대학교 공과대학 컴퓨터과학과

### 요 약

지금까지 기계 기반의 필기 숫자 인식 방법에 대한 많은 연구가 진행되어 왔다. 그러나 여전히 인간이 만족할 만한 인식 성능을 이루지는 못하였다. 이러한 배경에는 단순히 인식률을 나타내는 수치가 낮은 것도 한 부분을 차지 하지만, 인간이 수용할 수 없는 오류 성향도 중요한 부분을 차지한다. 그러므로 본 논문에서는 실제 인간의 숫자 인식이 어떻게 이루어지는지를 확인하는 실험을 먼저 수행하고, 이것에 근거하여 기계 인식을 위하여 필요한 요소들이 무엇인지를 고찰하도록 하였다. 실험결과 한쪽 또는 양쪽 방향으로 혼동하는 숫자 쌍, 전혀 혼동하지 않는 숫자 쌍, 오류 발생의 중복성 등의 결과를 비교 분석하여 인간이 인식과정에서 중요하게 고려하는 특징들을 찾아냈고, 그 결과에 근거하여 기계 인식에 있어서 더 높은 인식 성능을 발휘할 수 있고, 더 나아가 인간적인 측면에서 보다 더 신뢰할 수 있는 인식결과를 이끌어 낼 수 있는 접근 방향에 대하여 제시하였다.

키워드 : 필기 숫자, 기계 인식, 인간 인식.

### Abstract

So far there have been many researches on machine-based recognition of handwritten digit. But we have not yet attained the level of performance that can be satisfactory to men. The dissatisfaction with the performance of machine comes from not only the low accuracy of recognition but also the dissimilarity of the recognition results between man and machine. To reduce the difference of machine from man we first made an experiment with the human recognition of handwritten digits and then inquiry into the way of the human recognition that makes the results of men different from that of machine. We found out the attributes that play an important role in the human recognition process through the analysis of the experimental results like uni- and bi-directional confused pairs of digits, several ones unmixed up with another and the redundancy of mis-recognition, and proposed the approach direction to be able to improve the accuracy of the machine-based recognition, and furthermore the similarity in the recognition results of men and machine on the basis of the found facts above.

Key Words : Handwritten Digit, Machine Recognition, Human Recognition

### 1. 서 론

동일한 숫자라도 필기한 사람에 따라 그 모양이 매우 다르다. 또한 같은 사람이 쓴 숫자라 하더라도 때와 상황에 따라 서로 다른 모양을 보인다. 동일한 숫자의 필기 실험이 이처럼 다양함에도 불구하고 사람들은 필기된 숫자들을 아주 정확하게 인식하는 것처럼 보인다. 그러나 기계에 의한 필기 숫자 인식은 아직까지도 만족할 만한 성과를 얻지 못

하고 있다. 이러한 배경 중에는 단순히 수치적인 인식률이 낮은 것뿐만 아니라 인식 결과가 인간이 수용하기 힘든 경향을 보이는 탓도 크다. 기계에 의한 인식 결과는 인간의 패턴 인식 수행 결과와 비교되고 평가되기 때문에 극단적으로 보면 평균적인 인간의 결과와 같게 나오는 것이 최대 목표라 할 수 있다.

그러나 지금까지의 많은 연구들은 기계의 성능을 향상시키기 위하여 분별력이 좋은 특징을 선택하거나 추출하는 방법에 대한 연구[1,2,3,4,5], 각 클래스를 효과적으로 분류할 수 있는 방법에 대한 연구[6,7] 그리고 인간의 밝혀진 인지 과정을 기계에 직접 적용하려는 좋은 노력들[8,9]도 있었으나, 기계의 수행을 평가하는 데 중요한 근거가 될 수 있는

접수일자 : 2008년 3월 5일

완료일자 : 2008년 5월 13일

\*\*교신저자 : hsch@chosun.ac.kr

인간의 인식 수행 결과를 정밀하게 측정하여 이용되지 않았다. 이에 필기 숫자 인식에 흔히 사용되는 숫자 자극들을 사용하여 인간의 인식 수행 실험을 먼저 수행하고 그 결과를 분석하여 어떠한 특징과 방법을 사용하는 것이 기계 인식에 있어서 더 높은 인식 성능을 발휘할 수 있고, 더 나아가 인간적인 측면에서 보다 더 신뢰할 수 있는 인식결과를 이끌어 낼 수 있는 방법인지를 제시하고자 한다. 2절에서는 인간의 필기 숫자 인식 실험 및 결과에 대해서 서술하고, 3절에서는 인간의 인식 실험 결과에 대한 분석, 4절에서는 기계 인식을 위한 인간의 인식 특성을 분석한 후에 마지막 절에서 결론을 제시한다.

## 2. 인간의 필기 숫자 인식 실험

### 2.1 피험자

교정시력을 포함하여 모두 정상인 남녀 대학생 79명이 실험에 참가하였다.

### 2.2 장치

IBM PC/AT 호환 기종 컴퓨터를 사용하여 자극을 제시하고 반응을 측정 및 기록하였다. 자극제시를 위하여 17인치 모니터와 VGA 그래픽어댑터가 사용되었으며 화면 해상도는 800×600 화소로 고정하였다. 피험자 반응은 키보드의 숫자 키패드에 있는 키를 누르게 하여 기록하였다. 자극의 제시, 반응의 측정, 실험의 통제에는 Forster & Forster가 개발한 실험 생성 소프트웨어 DMASTER를 이용하였다.

### 2.3 자극재료

10,000 개의 필기 숫자 이미지 파일로 구성된 MNIST 데이터베이스[10]에서 각 숫자 당 40개씩 모두 400개의 숫자를 무작위로 선정하였다. 데이터베이스는 다음 표 1과 같이 구성되어 있다. 각 숫자의 이미지는 bmp 형식으로 저장되었으며 가로 28 화소, 세로 28 화소, 256 색상으로 표현되었다.

표 1. MNIST DB에 포함된 숫자

Table 1. Numbers of each digit in MNIST DB

숫자	0	1	2	3	4	5	6	7	8	9
개수	980	1135	1032	1030	982	892	958	1028	974	1009

### 2.4 절차

피험자가 모니터 앞에 착석하면, 우선 자극제시 방식과 반응 방법 및 주의사항이 기술되어 있는 실험 지시문을 읽게 했다. 이어서 지시에 대한 피험자의 이해 여부를 확인하고 바로 실험을 시작했다. 피험자의 눈과 모니터와의 거리는 약 50cm로 설정했다. 화면상에서 각 숫자의 크기는 가로 9.8mm, 세로 9.8mm 이었으며, 이것은 시각으로 가로 1.1°, 세로 1.1°에 해당하였다.

먼저 십자(+) 모양의 응시점을 화면 중앙에 600ms 동안 제시했다. 응시점이 사라지면 그 자리에 자극 숫자가 제시되었다. 피험자가 해야 할 과제는 제시된 숫자에 해당하는 숫자 키를 키보드의 오른쪽 부분에 있는 숫자 키패스에서 찾아 누르는 것이었다. 반응은 가능하면 신속하고 정확하게 하도록 요구했고 피험자의 키 누름 반응과 동시에 제시된

숫자는 사라지고, 그것으로 한 번의 시행이 완료되었다. 다음 시행은 1.5초 후에 시작되었다.

실험은 연습시행 10회와 본시행 400회로 구성되었다. 피험자를 실험 상황에 적응시키기 위해 본시행에서 사용되지 않은 숫자 이미지를 골라서 10회의 연습을 실시했다. 본시행에서는 400개의 자극이 무작위로 제시되었고, 실험 중간에 2회의 휴식시간이 있었다. 실험에 소요된 시간은 피험자 1인당 약 35분이었다.

## 3. 인간의 인식 실험 결과 분석

실험 결과로는 실험에 대한 집중도가 낮았던 5명에 대한 자료를 제외한 74명의 자료가 사용되었다.

인간의 인식 실험 결과에서 보이는 성향, 예를 들면 특정 혼동되는 숫자쌍이 있다는지, 혼동이 일방적으로 일어난다는지, 인식에 시간이 많이 필요하다든지 등을 분석함으로써 기계 인식의 성능과 신뢰성을 높일 수 있는 방법을 모색할 수 있다.

### 3.1 정량적 분석

각 숫자에 대한 평균 정인식률 및 표준편차를 표 2에 제시하였다. 피험자들은 10개의 숫자 모두에 대해 상당히 정확한 반응을 보였다. 전체적인 정인식률 98.88%가 그 것을 뒷받침한다. 숫자 “8”에 대한 반응이 99.76%로 가장 높은 정확률을 보였고, 숫자 “3”에 대한 반응이 98.21%로 가장 낮았다. 그 차이는 불과 1.55%였으나, 정인식률의 패턴을 좀 더 자세히 알아보기 위하여 변량분석을 실시했다. 분석 결과, 자극 종류(숫자)의 주효과가 유의미하였다 [ $F(9,657)=5.01, p<0.0001$ ]. 즉, 숫자에 따른 평균 정인식률 사이에 유의미한 차이가 있었다. 이에 대한 자세한 분석을 위해 숫자 간 평균 정인식률 차이를 사후검증으로 비교하였고, 표 3에 그 결과를 제시하였다. 평균 간 차이 값은 옆에 있는 자극숫자에 대한 평균 정확률에서 행에 있는 자극숫자의 평균 정확률을 뺀 값이다.

표 3의 내용을 보면 숫자 “8”에 대한 정확률이 특이할 정도로 높은 것이었음을 알 수 있다. 모든 숫자에 대해 유의미한 차이를 보이고 있다. 숫자 “0” 또한 다른 여섯 개의 숫자에 비하여 인식률이 높았고 숫자 “3”은 다른 다섯 개의 숫자에 비하여 유의미할 정도로 인식률이 낮았음을 알 수 있다.

정인식률에서의 차이를 이해하기 위해서 숫자 간 반응 혼동행렬을 표 4에서 자극-반응의 혼동행렬(confusion matrix)로 제시하였다. 표 4의 혼동행렬 표를 보면 몇 개의 오 반응 패턴이 드러난다. 구체적으로 “0”을 “6”으로(44회), “1”을 “0”으로(20회), “3”을 “6”으로(27회) 혼동하는 경우가 많았다. 이 외에도 “9”를 “8”로(19회), “6”을 “4”로(16회), “5”를 “6”으로(15회), “4”를 “9”로(11회) 혼동하는 경우도 많았다. 여기서 특기할 사실은 혼동의 방향이 일방적이라는 것이다. 예를 들면, “0”을 “6”으로 혼동하는 경우는 많았으나, 그 반대의 경우 즉, “6”을 “0”으로 혼동하는 것은 매우 드물었다. 이러한 관계를 그림으로 나타낸 것이 그림 1이다.

실험심리학에서는 정인식률 이외에 반응시간도 인지 과정을 반영하는 중요한 측정치로서 간주된다. 본 실험에서도 반응의 정확성 이외에 반응시간을 측정하였으므로 그 결과를 분석하였다. 우선, 각 숫자 별 평균 반응시간을 표 5에 제시하였다.

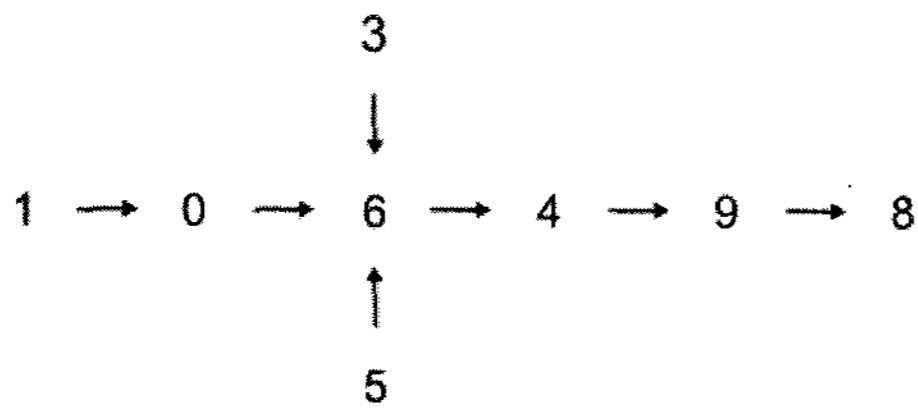


그림 1. 혼동 방향  
Fig 1. Direction of confusion

표 2. 각 숫자에 대한 평균 정인식률(%) 및 표준편차  
Table 2. Recognition rate(%) and standard deviation for each digit

숫자	0	1	2	3	4	5	6	7	8	9	전체
평균	98.24	98.55	99.19	98.21	99.02	98.99	99.05	98.78	99.76	98.95	98.88
표준편차	13.14	11.97	8.97	13.26	9.85	10.02	9.68	10.96	4.86	10.18	10.55
오반응률(%)	1.76	1.45	0.81	1.79	0.98	1.01	0.95	1.22	0.24	1.05	1.12

표 3. 숫자 간 평균 정인식률 차이의 사후 검증  
Table 3. Post-hoc test for the difference of recognition rate

	0	1	2	3	4	5	6	7	8	9
0		-0.31	-0.95**	0.03	-0.78**	-0.75*	-0.81**	-0.54	-1.52**	-0.71*
1			-0.64	0.34	-0.47	-0.44	-0.50	-0.23	-1.21**	-0.40
2				0.98**	0.17	0.20	0.14	0.41	-0.57**	0.24
3					-0.81*	-0.78*	-0.84**	-0.57	-1.55**	-0.74
4						0.03	-0.03	0.24	-0.74**	0.07
5							-0.06	0.21	-0.77**	0.04
6								0.27	-0.71**	0.10
7									-0.98**	-0.17
8										0.81**
9										

p<0.05\*, p<0.01\*\*

숫자 "5"에 대한 평균반응시간이 가장 짧았고(795.44ms), 이어서 "1", "2", "0", "8", "4", "9", "6", "3"의 순으로 길었으며, "7"에 대한 반응시간이 가장 길었다(899.21ms). 그 차이는 무려 104ms에 달하였기에, 각 숫자에 대한 반응시간 사이에 유의미한 차이가 있는지를 알아보기 위하여 변량분석을 실시했다. 변량분석을 위하여, 500ms 이하 반응시간과 3000ms 이상의 반응시간은 피험자의 부주의에 의한 반응으로 간주하여 자료에서 제외하였다. 분석결과, 자극 종류(숫자)의 주효과를 확인할 수 있었다[F(9,658)=18.45, p<0.0001]. 즉 숫자에 따른 평균반응시간 사이에 유의미한 차이가 있었다. 이에 대한 자세한 분석을 위해 숫자 간 평균반응시간 차이를 사후검증으로 비교하였다. 표 6에 각 평균 간 쌍 대 비교의 결과가 제시되어 있다. 표 6에 제시된 각각의 값들은, 열에 있는 자극 숫자에 대한 평균반응시간에서 행에 있는 자극 숫자의 평균반응시간을 뺀 값이다. 예를 들어, 표 5에서 볼 수 있듯이, "0"에 대한 반응시간은 "3", "6", "7"에 비해 유의미하게 길었으며 그 밖의 숫자들과는 유의미한 차이가 없었다.

표 4. 자극-반응의 혼동 행렬

Table 4. Confusion matrix of stimulus-response

자극 숫자	반응 숫자										전체
	0	1	2	3	4	5	6	7	8	9	
0	2908 (98.24)	7 (0.24)	1 (0.03)	0 (0.00)	0 (0.00)	0 (0.00)	44 (1.49)	0 (0.00)	0 (0.00)	0 (0.00)	2960 (100)
1	20 (0.68)	2917 (98.55)	1 (0.03)	0 (0.00)	8 (0.27)	0 (0.00)	10 (0.34)	1 (0.03)	3 (0.10)	0 (0.00)	2960 (100)
2	4 (0.14)	1 (0.03)	2936 (99.19)	1 (0.03)	0 (0.00)	4 (0.14)	5 (0.17)	4 (0.14)	5 (0.17)	0 (0.00)	2960 (100)
3	0 (0.00)	2 (0.07)	4 (0.14)	2907 (98.21)	0 (0.00)	13 (0.44)	27 (0.91)	0 (0.00)	2 (0.07)	5 (0.17)	2960 (100)
4	1 (0.03)	3 (0.10)	1 (0.03)	0 (0.00)	2931 (99.02)	1 (0.03)	6 (0.20)	2 (0.07)	4 (0.14)	11 (0.37)	2960 (100)
5	0 (0.00)	0 (0.00)	1 (0.03)	1 (0.03)	1 (0.03)	2930 (98.99)	15 (0.51)	0 (0.00)	9 (0.30)	3 (0.10)	2960 (100)
6	3 (0.10)	0 (0.00)	1 (0.03)	2 (0.07)	16 (0.54)	2 (0.07)	2932 (99.05)	0 (0.00)	2 (0.07)	3 (0.10)	2960 (100)
7	2 (0.07)	2 (0.07)	3 (0.10)	3 (0.10)	7 (0.24)	2 (0.07)	3 (0.10)	2924 (98.78)	7 (0.24)	7 (0.24)	2960 (100)
8	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	5 (0.17)	0 (0.00)	0 (0.00)	2953 (99.76)	2 (0.07)	2960 (100)
9	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	10 (0.34)	0 (0.00)	2 (0.07)	0 (0.00)	19 (0.64)	2929 (98.95)	2960 (100)

주: ( )속은 백분율

표 5. 각 숫자별 평균 반응시간(ms)

Table 5. Reaction time(ms) for each digit

	자극 숫자									전체	
	0	1	2	3	4	5	6	7	8		9
평균	815.53	811.93	814.44	866.04	843.52	795.44	864.27	899.21	824.65	845.51	838.31
표준편차	257.60	251.90	259.48	264.42	295.17	286.65	280.97	291.13	240.54	261.96	271.07

표 6. 숫자 간 평균 반응간 차이의 사후검증

Table 6. Post-hoc test for the difference of average responses

	0	1	2	3	4	5	6	7	8	9
0		3.60	1.09	-50.50*	-27.99	20.09	-48.74*	-83.67*	-9.12	-29.98
1			-2.51	-54.11*	-31.59	16.49	-52.34*	-87.28*	-12.72	-33.58
2				-51.59	-29.08	19.00	-49.83*	-84.76*	-10.21	-31.07
3					22.51	70.60*	1.77	33.17	41.39	20.53
4						48.09	-20.75	-55.68*	18.87	-1.99
5							-68.83*	-103.77*	-29.21	-50.07*
6								-34.94	39.62	18.76
7									74.56*	53.70*
8										-20.86
9										

p<0.05\*

### 3.2 오류가 발생한 숫자 쌍 분석

인간의 필기 숫자에 대한 인식 오류 결과를 분석해 보면 크게 두 가지 원인에서 기인하는 것으로 나타난다. 첫째는 필기 숫자 형태 자체에 대한 변형이 선행되어 오인식이 일어나는 경우이고 두 번째는 숫자 형태의 변형과는 무관한, 인간의 인식기제에서 기인한 오류이다. 다음 절에서 실험 결과를 두 가지 측면에서 비교 분석하였다. 이때 오인식은

적어도 다섯 번 이상 발생한 경우만 실제 오인식이 발생한 것으로 간주였으며, 분석에 사용된 숫자 쌍을 그림 2에서 보이고 있다. 각 숫자 쌍에 대해서 오인식이 발생한 횟수와 오인식이 발생한 영상의 수를 표시하였다. 숫자 쌍에서 앞의 숫자는 실험에서 제시한 숫자이고, 뒤 숫자는 오인한 숫자를 나타낸다.

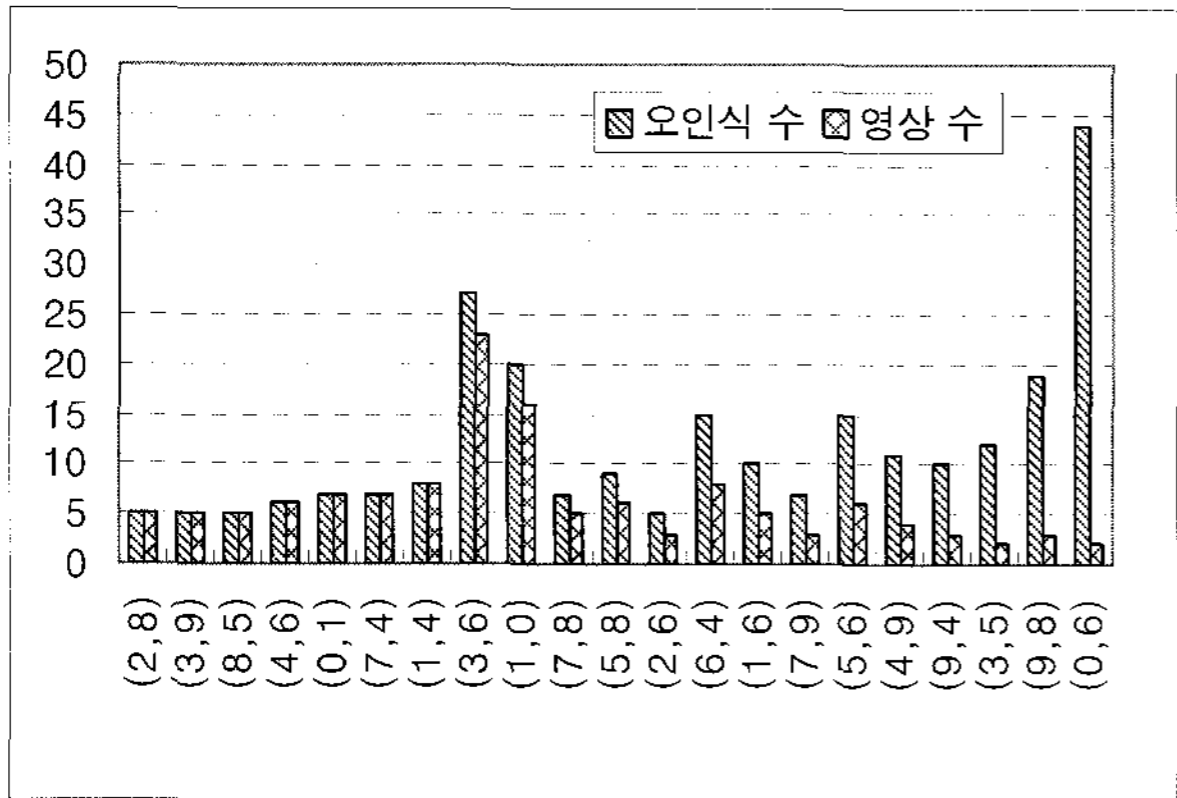


그림 2. 오인식이 발생한 숫자 쌍  
Fig 2. Pairs of mis-recognized digits

3.2.1 필기 숫자의 획 변형에 기인한 오류

오류의 일차 원인이 되는 획의 변형 성향을 분석하는 것이 필기 숫자 오류를 분석하고 이해하는 데 도움이 될 것이다. 이에 오류가 발생한 숫자들 중에 어떤 형태의 획 변형이 존재하고 이로 인한 인식 결과를 분류하여 정리하였다.

- (1) 곡선이 직선으로 또는 직선이 곡선으로 변형되는 경우로, 반원 등이 부등호 기호(<, >)와 같은 직선과 상호 변형되는 경우이다. “4”를 “9”로, “9”를 “4”로 오인한다.
- (2) 곡선 획의 끝 부분은 정확하게 마무리되지 않거나 지나치는 경우로, “0”을 “6”로, “9”를 “4” 등으로 오인한다.
- (3) 숫자가 많이 기울여 쓰이는 경우로, “6”을 “4”로 오인한다.
- (4) 획과 획 사이를 연결하는 획이 축소되거나 변형되는 경우로, “3”을 “5”로, “5”를 “6” 등으로 오인한다.

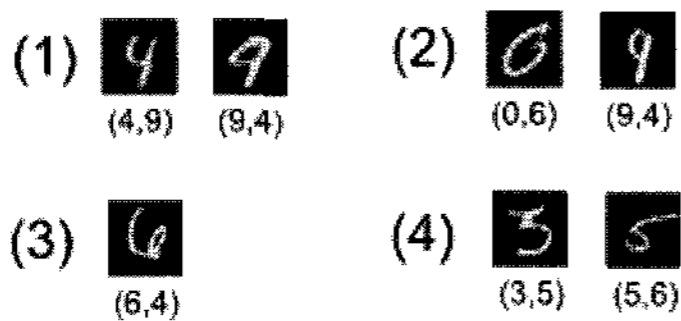


그림 3. 획 변형에 기인한 오인식된 문자의 예  
Fig 3. Examples of mis-recognized digits from the stroke distortion

3.2.2 인간의 인식기체에 기인한 오류

인간의 필기 숫자 인식 실험 결과를 보면 획의 변형은 거의 없어 보이나 혼동을 일으키는 경우가 여전히 있음을 알 수 있다. 더욱이 어떤 경우는 쉽게 납득하기 어려운 숫자들 간의 혼동이 일어나기도 한다. 물론 실험에 참가한 피실험자의 실수가 있을 수도 있으나 전혀 혼동하지 않은 숫자쌍이 존재하는데 반해서 최소 다섯 번 이상 혼동이 일어난다는 것은 분명 유의미한 관련이 있는 것으로 보인다.

다음에서 이러한 인식 오류 결과를 분석하여 오류 종류

와 해당되는 숫자에 대한 내용을 정리하였다.

- (1) 반원을 원으로 인식하는 경우로, “3”을 “9”로, “5”를 “6”로, “5”를 “8”로, “7”을 “8”로, “7”을 “9” 등으로 오인한다.
- (2) 부등호 기호 모양의 획을 원호로 인식하는 경우로, “4”를 “9”로 오인한다.
- (3) 뭉친 획을 원으로 인식하는 경우로, “1”을 “6”로, “2”를 “8”로, “9”를 “8” 등으로 오인한다.
- (4) 전경과 배경을 혼동하여 오류가 일어나는 경우, “0”을 “1”로, “1”을 “0”로, “3”을 “6” 등으로 오인한다.
- (5) 회전 또는 대칭된 숫자로 인식하는 경우로, “2”를 “6”로 오인한다.
- (6) 일부 구조 또는 전체적인 윤곽의 유사성이 존재하여 오인하는 경우로, “1”을 “4”로, “2”를 “8”로, “4”를 “6”로, “7”을 “4”로, “8”을 “5”로 등으로 오인한다.

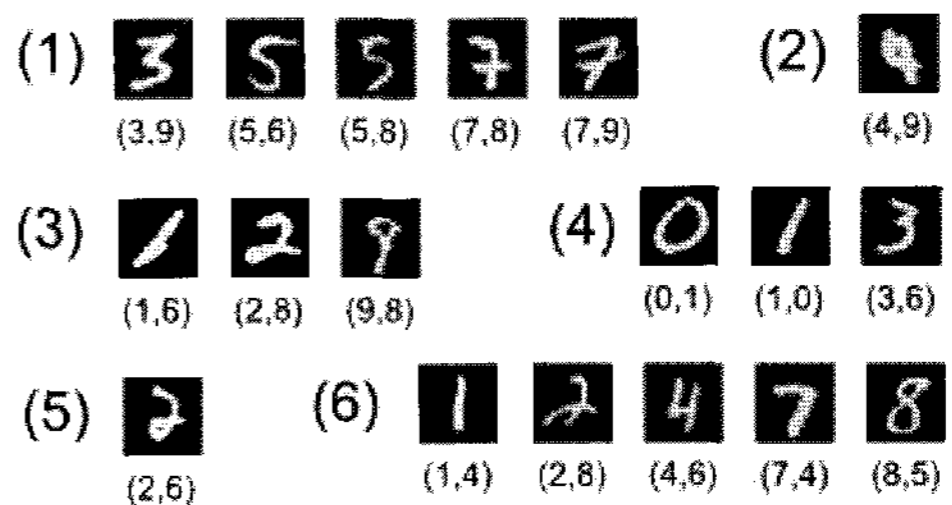


그림 4. 인식기체에서 기인한 오인식한 문자 예  
Fig 4. Examples of mis-recognized digits from recognition mechanism

3.2.3 각 영상의 오인식이 발생 빈도에 대한 분석

각 숫자 집합에 대해 오인식이 일어나는 영상의 분포를 분석해 보면 서로 다른 경향을 보인다. 즉, 어떤 숫자 집합의 경우에는 특정 영상이 계속해서 오인식을 일으키고 또 다른 집합의 경우에는 전체 영상에 대해서 고르게 오인식이 발생함을 알 수 있다. 식 (1)을 이용하여 오류 영상의 빈도를 계산하고 그 결과를 그림 5에서 보였다.

$$(N_m - N_k) \times 100 / N_m + 1 \quad (1)$$

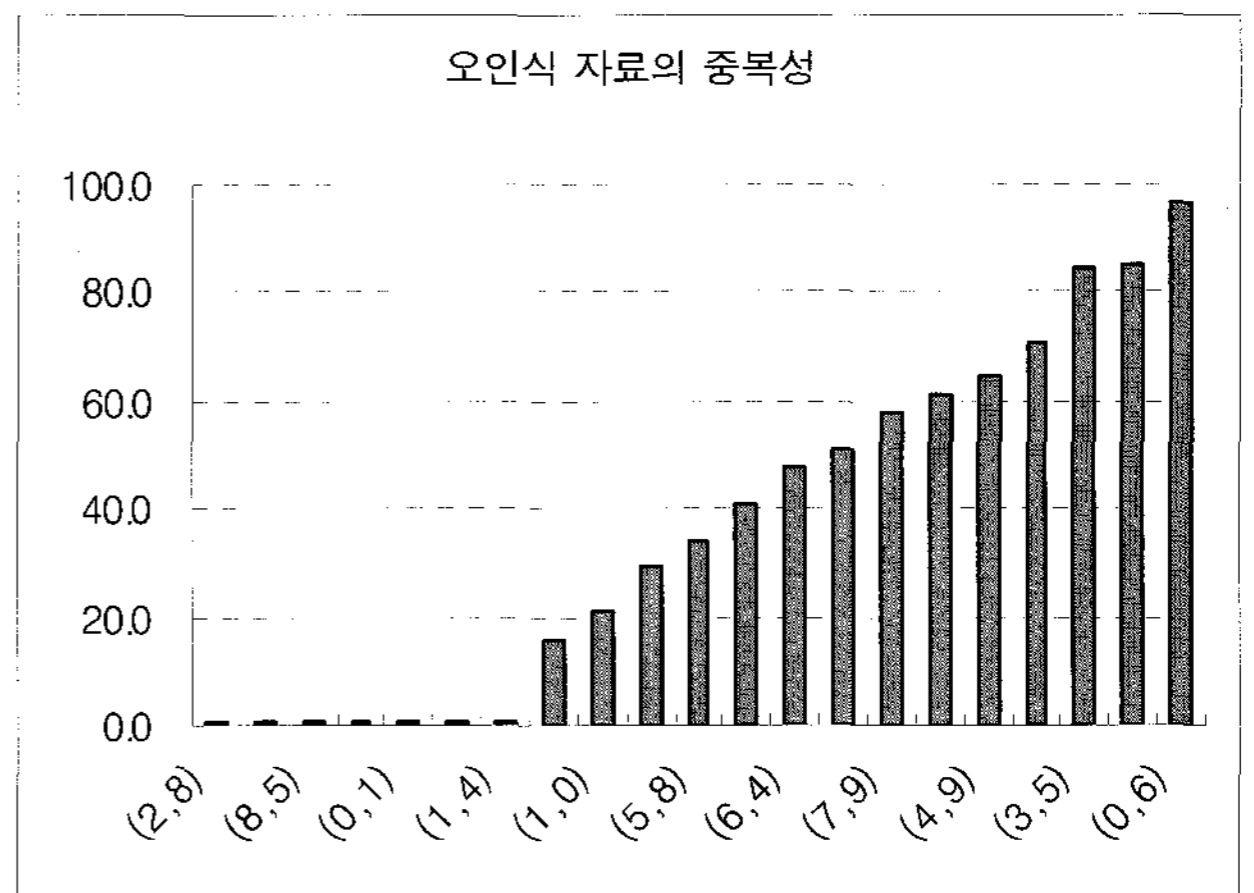


그림 5. 오인식 자료의 중복성  
Fig 5. Overlapping disposition of mis-recognized digits

$N_m$ 은 전체 오인식 수,  $N_k$ 는 오인식이 발생한 자료 수



이다.

오인식의 중복성이 높은 숫자 쌍은 (9,8)의 경우를 제외하면 획 변형에 기인해서 오인식된 숫자 쌍들임을 알 수 있다. 반대로 중복성이 없거나 낮은 숫자 쌍의 경우는 인식기제에서 기인하여 오인식된 숫자 쌍임을 알 수 있다.

중복성이 높은 숫자 쌍의 경우에는 약간의 획 변형으로 매우 유사한 형태를 갖게 됨을 알 수 있다. 그러므로 두 숫자 쌍 간에 중요한 특징이 무엇인지를 확인할 수 있다.

(0,6)의 경우에는 고리 모양을 이루는 위치가 가장 중요한 차이점이고, (3,5)의 경우에는 위쪽 부분의 빠침이 왼쪽이나 오른쪽이냐가 중요한 차이점이다. (9,4)의 경우에는 고리 모양을 이루는 부분이 곡선인지 직선 형태인지가 중요한 차이점이고, (5,6)의 경우에는 고리 모양을 이루었는지 아닌지가 중요한 차이점이다. (7,9)의 경우에도 역시 위쪽 부분이 고리 모양을 이루는지가 중요한 요소가 되고, (6,4)의 경우를 보면 직선과 곡선의 차이보다는 구조적 또는 획의 움직임의 유사성 더 중시되는 것으로 보인다.

중복성이 낮은 숫자 쌍의 경우에는 인식기제와 연관해서 유사성이 존재하는 것으로 봐야할 것이다. 즉, 특별한 획의 변형이 없이 일반적인 자료에 대해서 일어날 수 있는 것으로 보이기 때문이다.

(9,8), (1,6)의 경우를 보면 고리 모양이 나타나지 않아도 단지 획이 뭉친 것 같은 정보만 있어도 고리 모양으로 간주하는 경향이 높아 보이는데, 이 숫자 쌍의 경우에 중복성이 높은 것을 보면 임계값 이상의 특징이 존재해야 하는 것으로 생각된다. (1,0), (3,6)의 경우는 아주 자주(각각, 20, 27회) 많은 영상(각각 16, 23개)에서 혼동을 보이는데 전경뿐만 아니라 배경 정보도 중요하게 처리됨을 짐작할 수 있다.

### 3.3 혼동이 일어나지 않는 숫자 쌍 분석

혼동이 일어나는 경우와 마찬가지로 서로 혼동되지 않는 숫자 쌍에 대한 분석이 역시 인식기제에 대한 많은 정보를 줄 수 있다.

#### 3.3.1 상호간 획 변형이 어려운 경우

숫자의 모양을 분석해 보면 그림 6와 같이 네 개의 기본 모양으로 이루어져 있음을 알 수 있다. 그러나 실제 필기 숫자가 쓰일 때는 획 간의 이동이 중요한 변수로 작용하기 때문에 획의 움직임을 고려해서 획 모형을 분류하면 그림 7(a)와 같다. 이러한 모형과 대부분의 획 변형은 말단에서 발생한다는 사실을 이용하면 획 변형에 의한 혼동되는 숫자 쌍을 설명할 수 있다.



그림 6. 숫자의 획 모형  
Fig 6. Stroke model of digit

- (1) 위에서 밑으로 움직이면서 고리를 구성하는 형태의 특징을 가진 “8”자의 경우에는 획의 변형에 의해서 오인되는 경우는 없다. “5”와 “9”에 대한 혼동은 인식기제에 기인한 것으로 보는 것이 옳다.
- (2) 그림 7(c)에서 보이는 것과 같이 획의 방향이 반대인 원호 또는 원형의 경우는 획의 변형이 발생하더라도 열린 입구가 전혀 달라지기 때문에 오인이 일어나는 경우가 거의 없다. “2”를 “4” 또는 “9”로, “6”을 “2”로 혼동하는 경우가 전혀 없었다. “2”를 “6”으로 혼동하는 경우가 4

회 정도 발생하였지만 이것 역시 인식기제의 원인으로 보는 것이 타당해 보인다. 단순 모양으로는 “2”를 “0”으로 오인할 수 있게 획 변형이 다수 발생할 것으로 보이나 실제 실험에서는 4회 정도만 발생하였다. (0,6) 쌍이 가장 많은 오인식을 보인 것과 대조적으로 (0,9)는 전혀 혼동됨이 없다는 사실을 보면 획 변형은 획의 흐름이 매우 중요함을 확인할 수 있다.

- (3) 곡선으로 이루어진 획이 직선으로 변형되어 잘 못 인식될 경우는 그 반대의 경우보다 낮은 빈도를 보인다. 예를 들면, “8”은 “4”와 전혀 혼동되지 않지만 “4”는 “8”과 혼동되는 경우(4회)가 나타나고, “3”은 “7”과 전혀 혼동되지 않지만 “7”은 “3”으로 오인식이 3회 일어났다.

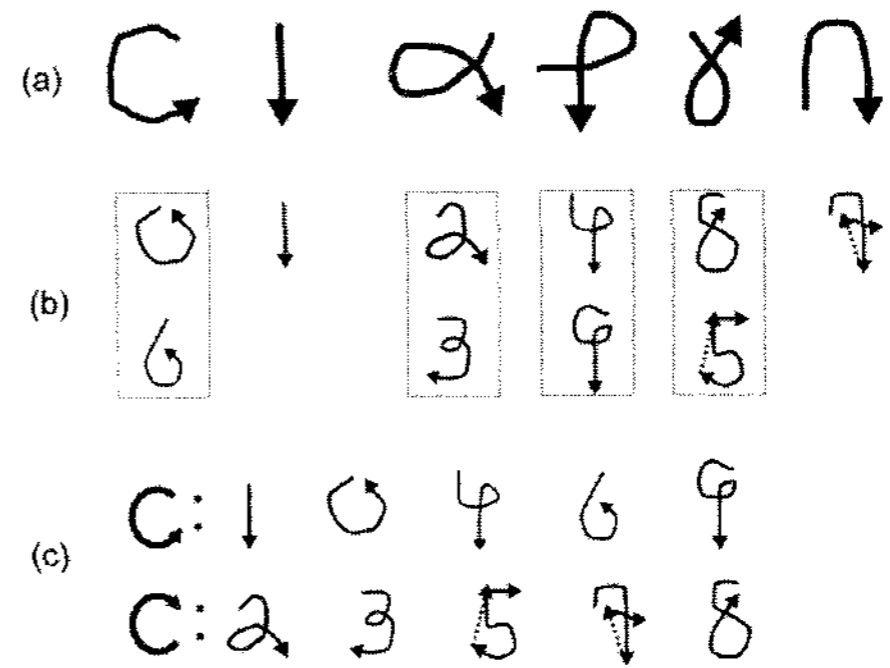


그림 7. (a)획 흐름 모형 (b)각 숫자의 획 흐름  
(c)왼/오른쪽 방향의 숫자 집합

Fig 7. (a) Model of stroke flow (b) stroke flow for each digit (c) two groups of digits with a left/right flow direction

#### 3.3.2 인식기제 측면에서 혼동 발생이 어려운 경우

인식기제에 의해 혼동이 일어나지 않는 경우를 분석해 보면 다음과 같다.

- (1) 곡선으로만 이루어진 숫자들을 선형으로만 이루어진 숫자와 혼동되는 경향이 적다. 예를 들면, “3”의 경우 “4” 또는 “7”과 혼동이 전혀 발생하지 않았고, “6”을 “7”로 혼동하는 경우도 없었다.
- (2) 여러 획으로 이루어진 복잡한 숫자는 간단한 획으로 이루어진 숫자들과 혼동의 여지가 적다. 예를 들면, “3”, “5”, “8” 그리고 “9”는 “0”과는 전혀 혼동을 일으키지 않고, “5”, “6” 그리고 “9”는 “1”과 전혀 혼동이 발생하지 않았다.
- (3) 원호 또는 부등호(<, >) 형태의 특징이 열린 입구의 방향이 차이가 큰 경우에 혼동되는 경향이 적다. 예를 들면, “5”를 “7”로, “6”을 “2”로 혼동하는 경우가 전혀 발생하지 않았다.

### 3.4 혼동의 방향성에 대한 분석

3.1절에서도 밝혔지만 혼동은 일정 방향을 갖고 일어나는 경향이 있다. 숫자 쌍들의 오인식이 어떻게 방향성을 갖게 되는지 쌍방으로 혼동이 일어나거나, 반대로 어느 방향으로 혼동이 일어나지 않는 경우를 비교분석한다.

#### 3.4.1 쌍방 간 혼동이 일어나는 경우

쌍방으로 모두 혼동이 일어나는 경우를 획의 변형에 의해서 두 숫자가 혼동되는 경우와 획 변형 없이 인식기제에

의해 두 숫자가 혼동되는 경우, 그리고 두 이유가 각기 적용되는 경우로 나누어 생각할 수 있다. 다음은 이런 경향을 보이는 숫자 쌍들에 대해 정리한 것이다.

- (1) 쌍방 모두 획 변형이 원인인 경우: {4,9}
- (2) 쌍방 모두 인식기제가 원인인 경우: {0,1}
- (3) (1)과 (2)가 각각 원인인 경우: {4,6}, {5,8}

{4,9}의 경우에는 그림 7의 획 흐름 모형에서도 알 수 있듯이 같은 획 흐름을 갖고 있으므로 획 변형이 발생하기 쉽고, 또한 변형된 숫자는 매우 비슷한 형태가 된다. 이것은 그림 5의 오인식 발생 중복성이 쌍방 모두 높음을 봐도 확인할 수 있다. 이때 “4”와 “9”를 결정짓는 중요한 특징은 얼마나 부드럽게 고리 모양이 이루어지느냐이다. 직접적으로 고리 모양을 이루지 못하고 호 형태를 이루더라도 그 끝이 부드러운 곡선 선상에 있는 것처럼 보인다면 하나의 선으로 보는 경향이 있다.

{0,1}의 경우는 전경과 배경을 혼동하면 상대 숫자로 오인하기 때문에 쌍방 간에 혼동이 일어나고 있다. 배경이 숫자 크기에 비하여 크지 않은 경우에 배경의 경계가 마치 인식된 숫자의 경계 역할을 하게 된다. 만일 숫자 영상의 배경과 이것을 보여주는 화면의 배경이 같은 색을 갖도록 실험하면 다른 결과가 나올 것으로 예상된다. {4,9}와 같이 유사한 획 흐름에 의한 획 변형은 상호 비슷한 횡수의 오인식이 발생하나 {0,1}과 같이 인식기제에 의한 상호 오인식은 각기 발생할 수 있는 상황이 다르다. (0,1)로 오인식하는 경우보다 (1,0)로 오인식하는 경우가 4배 가까이 더 많았다.

{4,6}와 {5,8}은 모두 각각 획 변형과 인식기제에서 기인한 이유로 상호 혼동을 일으키고 있기 때문에 상호 독립적인 환경에서 발생하게 된다. 예를 들어, 만일 인쇄 문자를 이용한 동일한 실험을 한다면 (4,6)으로 혼동하는 경우는 발생할 수 있지만 (6,4)로 혼동하는 경우는 거의 발생하지 않을 것이다. 왜냐하면 “4”를 “6”으로 혼동하는 것은 우리의 인식기제에서 나타나는 오류로 변형 없는 인쇄 문자에서도 발생할 수 있으나, “6”을 “4”로 혼동되는 획 변형은 인쇄 문자에서는 나타날 수 없기 때문이다.

### 3.4.2 쌍방 간 전혀 혼동되지 않는 경우

전혀 한 번도 혼동이 일어나지 않은 숫자 쌍들은 {0,3}, {0,5}, {0,8}, {0,9}, {3,4}, {1,5}, {1,9}, {2,9}이다. 이들은 앞서 ?절에서 설명한 이유로 획 변형이 일어나기 어려울 뿐만 아니라 인식기제에 기인한 혼동도 없는 숫자 쌍이다. 특이한 점은 획의 변형이 일어나지 않는 쌍들 중에 “0”과 혼동을 일으키지 않는 쌍이 많다는 것이다. 또한 “1”과 혼동되지 않는 숫자 쌍도 두 개나 된다. 이 두 숫자가 포함된 경우의 비율이 75%에 이른다.

원칙적으로 고리 모양을 갖고 있는 숫자들(0, 8, 9)이 포함된 경우의 비율을 보면 75%가 된다. 10개의 숫자 문자 중에서 고리 모양을 갖는 숫자는 모두 포함된 것이다.

숫자 모양의 개방된 입구 방향을 보면, 모든 숫자 쌍이 반대 방향을 갖거나 “1” 또는 “0”과 같이 개방된 방향이 없는 경우이다.

### 3.4.3 혼동의 방향성에 대한 원인

결과적으로 한쪽 방향으로만 혼동이 일어나는 이유는 두 가지 원인에 근거한다. 첫 번째는 획을 쓰는 습관과 방향 등에 기인하는 부분이다. 예를 들면, “5”의 위쪽 획이 약간 곡선으로 써지면 “9”로 오인식할 수 있으나 “9”의 위쪽 원형은 “ㄷ”자 모양으로 써지지 않기 때문에 오히려 “4” 또는

“8”로 오인식된다. “5”의 위쪽 부분의 수직 획이 축소됨으로써 “6”과 가장 많은 혼동이 일어나나 상대적으로 “6”은 밑쪽 원형이 개방되는 방향이 왼쪽인 경우가 없기 때문에 오히려 약간 기울어짐으로써 “4”와 가장 많이 혼동된다.

오인식 결과가 한쪽 방향인 두 번째 이유는 인식기제의 특성이 존재하는 특징을 무시하기는 어려우나 반대로 없는 특징은 약간의 단서만 있으면 재구성하기는 쉽기 때문이다. 예를 들면, “6”은 “5”로 잘 혼동되지 않으나 “5”의 호가 조금만 길어져도 “6”자의 아래쪽 원의 왼쪽이 잘린 것으로 생각되어 “6”으로 많은 혼동이 일어난다. “3”은 전경과 배경을 혼동하면서 영상자료 전체에 걸쳐서 많은 수가 “6”과 혼동을 일으키고 있으나 “6”은 전경과 배경을 혼동하더라도 “3”과는 유사한 특징을 갖기 어려워 상대적으로 아주 작은 수(2회)만이 혼동을 일으킨다.

## 4. 기계 기반 인식을 위한 인간의 인식 특성 분석

### 4.1 인간의 인식 특성

지금까지의 분석 내용을 종합하여 인간의 인식 특성을 정리해 보면 다음과 같다.

- (1) 인식에 있어서 직선과 곡선에 대한 경계가 분명하지 않다. 즉, 직선이 약간 변형되어 나타나면 곡선으로 쉽게 받아들일 수 있고 반대로 곡선이 약간 변형되면 직선으로도 쉽게 인식한다.
- (2) 직선 또는 곡선에 대한 특징보다는 전체적인 구조에서 오는 특징이 더 강한 것으로 보인다. 예를 들면, 특정 부분이 직선이나 곡선이나가 중요한 것이 아니라 어떤 방향으로 열린 구조가 있느냐가 더 중요한 정보로 쓰인다.
- (3) 유사성이 높은 숫자 쌍의 경우 특정한 국부적인 특징이 두 수를 구별하는 데 중요하게 쓰일 때도 있다. 예를 들어, {4,9}와 같은 유사성이 높은 숫자 쌍의 경우 위쪽 고리 모양의 꺾인 형태의 분석 결과에 따라 판단한다.
- (4) 획에 대한 길이는 크게 작용하지 않고 전체 획을 구성하는데 약간의 정보 제공만 할 수 있으면 족하다.
- (5) 회전에 의한 변형은 인식에 큰 장애가 되지 못한다. 오히려 회전이 전혀 없는 데도 회전된 형태의 대상과 혼동하는 경우가 있다.
- (6) 뭉친 획이 원 모양으로 받아들여지는 것처럼 획 모양에 대한 정보는 윤곽에서도 추출한다.
- (7) 전경과 배경을 헛갈려 할 정도로 전경과 배경이 서로 상호작용 한다.
- (8) 단순한 획은 쉽게 구별할 수 있다. 즉, “0”과 “1”은 다른 문자와 혼동이 적다.
- (9) 고리 모양은 중요한 특징이 된다. “0”, “8”, “9”의 숫자들은 상대적으로 다른 숫자들과 혼동되는 경향이 낮다.
- (10) 특징들이 중복되면 더 확실한 정보를 제공한다. 예를 들면 “0”과 “1”은 모두 단순한 획이지만 “0”은 고리 모양이라는 추가된 특징이 있어 더 쉽게 구별되는 것으로 나타난다.

### 4.2 기계 기반의 숫자 인식에서 고려해야 할 사항

앞의 결과들을 종합해 보면 다음 사항들을 기계 기반 인식에서 고려할 필요가 있다.

- (1) 지역적인 특징 예를 들면, 직선이나 곡선이나 하는 특징은 보조적인 특징으로 사용하는 것이 타당해 보인다. 즉, 전체 숫자 인식과정에서 문자 분류에 결정적인 역할을 수행하는 것보다는 후반부에 위치하여 판단에 보조적인 정보를 주는 것이 적합하다.
- (2) 전체 구조를 파악할 수 있는 방법이 필요하다. 예를 들면 고리 모양과 같은 특징이 어디에 어떤 크기로 존재하는지 알 필요가 있다. 그렇게 하기 위해서는 화소 단위의 정보를 의미를 파악할 수 있을 만큼의 큰 단위로 통합하는 과정이 필요하며, 이런 과정이 계층적으로 표현되어 지역적인 정보와 전역적인 정보를 필요에 따라 얻을 수 있어야 한다.
- (3) 인식 과정 후반부에는 혼동되는 두 숫자 간 알려진 지식을 이용하여 검증하는 과정이 필요해 보인다. 예를 들면 혼동되는 숫자 간에 특정 위치의 특징들을 비교 분석함으로써 더 가능성이 높은 쪽을 선택할 수 있도록 돕고 결정된 결과를 검증하는 것이다.
- (4) 구조적인 특징을 결정짓는 요소는 획의 길이보다는 획 흐름 안에서의 변화 정도가 더 중요하게 고려되므로 획 흐름의 변화를 추적할 수 있는 방법이 필요하다.
- (5) 회전에 강하고 심지어는 회전된 상태의 것과 혼동을 일으키는 것으로 보아 제시된 숫자에서 추출된 특징들이 절대적 위치 정보와 결합되는 것이 아니라 상대적으로 표현되고 해석되는 것으로 생각된다. 따라서 획의 흐름을 중심으로 하는 좌표 체계를 이용할 필요가 있다.
- (6) 획이 뭉쳐 방향성을 갖지 못하는 상황에 원 요소로 인식하는 것을 보면 획의 흐름에 대한 정보에는 윤곽 정보도 함께 포함을 시킬 필요가 있다.
- (7) 전경 또는 배경 중에서 어떤 정보를 중심으로 구성하는지에 따라서 결과가 달라질 수 있는 것으로 보아, 획 경계에 대한 부분적인 모습이 중요한 정보로도 사용됨을 알 수 있다. 그리고 일부분의 획 또는 획 경계 정보만을 이용하여 전체 숫자를 판단하는 것을 보면 지역적인 특징들이 모여 인식이 일어나는 과정이 있고 또한 이러한 과정의 중간 단계의 결과만으로도 판단을 내릴 수 있는 것으로 생각된다.
- (8) 단순한 획으로 이루어진 숫자가 혼동이 적게 일어난 실험 결과를 보면 인식은 상대적으로 비교해서 결정되는 것으로 생각된다. 즉, 어떤 특징들이 존재해서 결정될 수도 있지만 다른 숫자에 비교해서 특징이 존재하지 않기 때문에도 결정될 수 있는 것으로 보인다.
- (9) 각 숫자 마다 결정에 중요하게 사용되는 특징들을 독립적으로 관리할 필요가 있고 특징들 간의 중요도도 함께 고려되어야 한다. 즉, 고리 모양과 같이 중요한 특징은 다른 특징에 비하여 결정에 많은 영향을 줄 수 있어야 한다.
- (10) 여러 단계와 경로를 통해 중복으로 검증할 필요가 있다. 예를 들면, 고리 모양의 존재 유무에 의해 처리하는 경로뿐만 아니라 숫자 모양의 열린 방향에 대한 처리도 함께 이루어질 필요가 있다.

## 5. 결 론

기계 인식에 있어서 인식률이 가장 중요한 평가 항목이지만 오인식된 결과 또한 중요한 의미를 갖는다. 왜냐하면 사람이 납득할 수 있는 오인식 결과는 사람도 틀릴 수 있기

때문에 큰 문제가 되지 않지만, 받아들이기 어려운 결과는 신뢰성에 중요한 문제를 야기하기 때문이다. 따라서 사람의 필기 숫자 인식 실험을 선행한 후에 그것에 근거하여 기계 인식에서 필요한 사항들을 추론함으로써 기계 인식의 정확성과 신뢰성을 높일 수 있는 방법을 제시하였다. 그러나 인식기제와 관련된 사항 중에는 정확하게 원인을 규명하지 못한 내용들이 있어 좀 더 심도 있는 학제간 연구가 필요할 것으로 생각된다.

본 실험과 관련하여 앞으로의 보강이 필요한 부분들을 열거하면 다음과 같다. 첫째, 제시된 방법론을 적용한 기계 인식의 결과에 대한 추가 실험이 필요하다. 둘째, 실제 인식 오류의 원인이 필기 숫자의 변형에 기인하는 정도와 인식기제에서 기인한 정도를 추정하기 위해 인쇄 숫자와 같이 변형이 없는 문자에 대해서 같은 실험을 해볼 필요가 있다. 셋째, 반응시간은 인간의 인식 수행을 측정하는 유용한 측정치가 될 수 있으므로 이를 적극적으로 활용할 필요가 있다. 특히 정인식률이 매우 높고 숫자 종류 간의 차이가 크지 않은 본 실험과 같은 경우에 반응시간이란 측정치는 정인식률이 말해주지 않는 인식수행의 차이를 보여줄 수 있다. 이를 위해서 키 누름에 의한 반응이 아니라 음독(naming) 반응을 사용할 필요가 있다. 음독 반응은 숫자를 제시하여 그 숫자를 소리 내어 읽기까지의 시간을 측정하므로, 키의 공간적 배치에 따른 영향을 최대한 배제할 수 있는 장점이 있다. 넷째, 자극의 제시 방법을 바꾸어 볼 필요가 있다. 본 실험에서 자극은 분명하게 그리고 충분한 시간만큼 제시되었다. 그러나 인식 수행을 측정하는 실험에서는 인식에 필요한 제시 시간 또는 자극 선명성 등을 통제하여 인식의 역치(threshold)를 측정하는 일이 흔하다. 즉 반응의 단순한 정확성 및 반응 속도가 아니라 인식역을 측정하는 방법을 사용하는 것도 검토하여 볼 만하다. 다섯째, 본 실험에서는 1만개의 숫자 이미지 모집단 중에서 오직 400개의 숫자 이미지 샘플만 사용되었다. 물론 무작위로 선택하였으므로 샘플은 모집단을 잘 반영할 것으로 생각되나, 특정 자극에 특정 오반응이 집중되는 경우가 있었다. 예를 들면, 자극번호 6644번은 "0"의 이미지를 담고 있는데, 피험자의 약 반수가 이를 "6"으로 판단하고 있다. 이러한 오류를 막고, 샘플이 모집단을 더 잘 반영하게 하기 위해서는 피험자마다 자극 샘플을 다르게 하는 방법을 사용할 필요가 있다.

## 참 고 문 헌

- [1] 이찬희, 정순호, "획의 방향 코드 조합에 의한 오프라인 필기체 숫자 인식," *한국정보과학회논문지: 소프트웨어및응용*, 제31권 12호, pp.1581-1590, 2004
- [2] 이진선, "유전 알고리즘을 이용한 특징 결합과 선택," *한국콘텐츠학회논문지*, 제5권 5호, pp.152-158, 2005
- [3] Wierer, J., Boston, N., "A Handwritten Digit Recognition Algorithm using Two-Dimensional Hidden Markov Models for Feature Extraction," *Proc. of ICASSP'07*, vol 2, pp. 505-508, 2007.
- [4] Garg, Naresh Kumar, Jindal, Simpel, "An Efficient Feature Set for Handwritten Digit Recognition," *Proc. of ADCOM'07*, pp. 540-545, 2007.



[5] Javed, I., Ayyaz, M.N., Mehmood, W., "Efficient Training Data Reduction for SVM based Handwritten Digits Recognition," *Proc. of ICEE'07*, pp. 1-4, 2007.

[6] 박중소, 김경민, "SVM 분류기를 이용한 필기체 숫자 인식," *한국신호처리시스템학회논문지*, 제8권 3호, pp.136-142, 2007

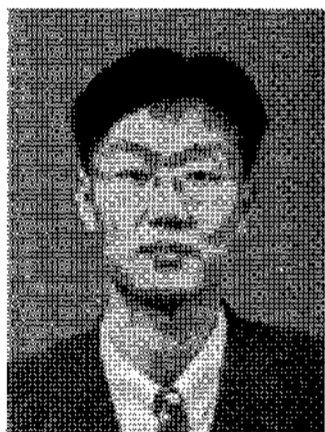
[7] K. Takahashi, D. Nishiwaki, "A class-modular GLVQ ensemble with outlier learning for handwritten digit recognition," *Proc. of ICDAR'03*, vol.1, pp. 268-272, 2003.

[8] 김도현, 김광백, 조재현, 차의영, "인간 시각의 선택적 지각 능력에 기반한 패턴 분류," *한국해양정보통신학회논문지*, 제10권 2호, pp. 398-405, 2006.

[9] 조성목, 조옥래, "인간시각 인식특성을 지닌 효율적 비선형 스케치 특징추출 필터," *한국컴퓨터정보학회논문지*, 제11권 1호, pp. 139-145, 2006.

[10] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition." *Proc. of the IEEE*, vol.86, no.11, pp. 2278-2324, 1998.

저 자 소 개



**윤성수(Sungsoo Yoon)**  
 1997년 : 연세대학교 컴퓨터과학과 공학 석사  
 2002년 : 연세대학교 컴퓨터과학과 공학 박사  
 2004년 : 뉴욕 PACE 대학 방문교수  
 2006년~현재 : 이화여자대학교 연구교수

관심분야 : 패턴인식, 생체인식, 정보보안

E-mail : ssyoon@ewha.ac.kr



**정현숙(Hyunsook Chung)**  
 1995년 : 대구가톨릭대학교 전산통계학과 이학석사  
 2003년 : 연세대학교 컴퓨터과학과 공학 박사  
 2006년~현재 : 조선대학교 컴퓨터공학부 전임강사

관심분야 : 게임공학, 웹컨텐츠제작, 모션그래픽스, 시멘틱웹

E-mail : hsch@chosun.ac.kr



**이광오(Kwangoh Yi)**  
 1979년 : 서울대학교 심리학과 문학사  
 1983년 : 서울대학교 심리학과 문학석사  
 1993년 : Hokkaido University 행동과학 박사  
 1989년~현재 : 영남대학교 심리학과 교수

관심분야 : 단어인지, 문장이해, 음성지각

E-mail : yiko@yu.ac.kr



**이일병(Yillbyeong Lee)**  
 1976년 : 연세대학교 전자공학과  
 1980년 : University of Illinois, Urbana-Champaign Dept. of Computer Science(M.S.)  
 1985년 : University of Massachusetts, Amherst Dept. of Computer & Information Science(Ph.D)

1985년~현재 : 연세대학교 정보산업공학 교수

관심분야 : 패턴인식, 생체인식, 데이터마이닝

E-mail : yblee@cs.yonsei.ac.kr



**이상호(Sangho Lee)**  
 1979년 : 서울대학교 계산통계학과 이학사  
 1981년 : 한국과학기술원 전산학과 이학 석사  
 1987년 : 한국과학기술원 전산학과 공학 박사  
 1990년 : 미국 일리노이대학교 전산학과 방문교수

1984~현재 : 이화여자대학교 컴퓨터학과 교수

관심분야 : 암호 및 보안, 데이터마이닝

E-mail : shlee@ewha.ac.kr