Human Performance in Recognition of Handwritten ZIP Codes from the CEDAR Database

Ibrahim Y. Chaaban

Michael R. Scheessele

Abstract

We established human performance in recognition of handwritten ZIP codes taken from the standard CEDAR database. We expect that the result will serve as a benchmark for machine performance in recognition of handwritten ZIP codes.

INTRODUCTION

Handwritten digit recognition research concentrates on either individual digits or digit strings. In particular, unconstrained handwritten digit recognition has been applied to recognize amounts written on checks for banks, tax returns, data entry for hand-held computers, or ZIP codes on envelopes for postal service. In these domains, handwritten digits rarely appear isolated. Instead they appear as part of a string of digits where some digits may touch and/or overlap. In many of these real world applications, the images are processed by human operators. However, automation may improve production and cut costs. For this to happen, performance of an automated system should compare favorably to human performance. Such comparison is also an essential component in determining whether the problem has been solved or not.

On average a postal worker can sort about 800 letters an hour. On the other hand, an automated sorting machine, reading printed ZIP codes with an optical scanner is estimated to process about 37 times more than the postal worker at a fraction of the cost [3]. Such performance would also be desirable for handwritten ZIP codes. ZIP code recognition has always been a challenging task in pattern recognition. Many systems and classification algorithms have been proposed in recent years, due to the benefits of having an accurate automated system that can recognize handwritten digits at a high rate. Many solutions have been proposed to solve this problem; however it is not known how the

'best' solution compares to human performance because no human recognition rate on handwritten ZIP codes has been established. Therefore, we performed an experiment to establish human performance in recognizing handwritten ZIP codes. The next section gives a description of this human experiment.

EXPERIMENT

Method

Subjects

Four undergraduate Indiana University South Bend (IUSB) students participated in this experiment[1]. Subjects were at least 18 years old and had normal (20/20) or corrected-to-normal vision in both eyes. For completing the experiment each subject was paid $10.00. All subjects gave their informed consent prior to participation, and the study was approved by the IUSB Institutional Review Board beforehand.

Stimuli

For this experiment we used the 436 5-digit ZIP code images available from the CEDAR CDROM-1 in the testing folder under the BINZIPS directory [1]. These ZIP codes have been used by researchers to test various machine classification methods [2]. The ZIP codes were segmented from mail images obtained by the USPS, and they are in binary format. In this database, some ZIP codes have digits that are touching and/or overlapping. Figure 1 shows examples of handwritten ZIP codes used in this experiment.

---

[1] In visual perception studies where large *individual differences* between subjects are expected a larger number of subjects (i.e. 12-30 subjects) should be tested. When individual differences between subjects are not expected in a perceptual study, it is permissible to use fewer subjects, sometimes as few as two subjects. Here, large individual differences between subjects with normal vision were not expected. For example, it would be unlikely that one subject would correctly recognize 95% of the CEDAR ZIP codes, while another subject would recognize just 65%.

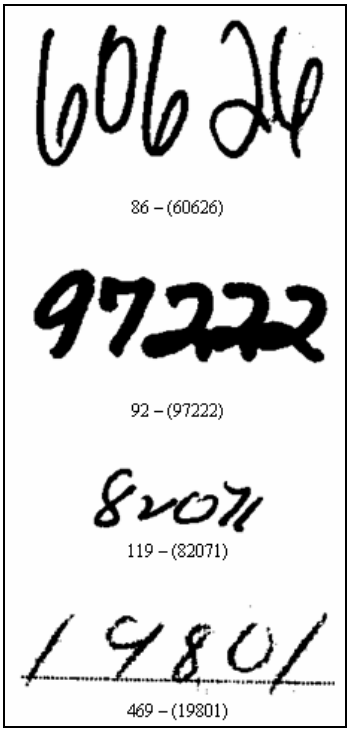

Figure 1. Sample patterns from the CEDAR CDROM-1 database. The number at left represents the pattern number. The number at right is the correct classification of the ZIP code.

Procedure

Each subject attended one session in which he/she sat in front of a computer screen and used software especially created for this experiment (Figure 2).

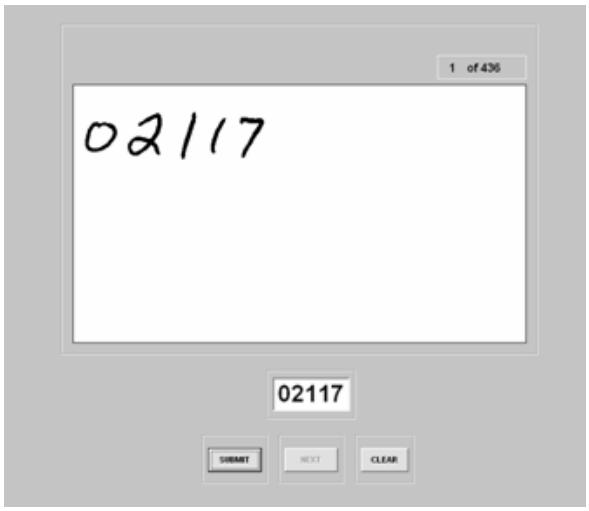

Figure 2. A screenshot of the software used in the Experiment.

A subject's task was to identify a series of handwritten ZIP codes randomly presented on the computer screen, by using a mouse/keyboard. For an ambiguous stimulus, a subject was instructed to do their best, but to still respond.

Results

The average percent error of the four subjects was 3.44% (Table 1).

Subject #	Errors	Percent Error
1	10	2.29%
2	8	1.83%
3	39	8.94%
4	3	0.69%
Average Percent Error		3.44%

Table 1. Results from the Human Experiment. The average percent error was 3.44%.

Due to the ease of the task, the legibility of most ZIP codes, and the performance of subjects 1, 2, and 4, the result of subject 3 appears to be an outlier. Subject number 3 apparently had some lapses of attention because some of his/her errors were on easy ZIP codes. The average percent error with this outlier removed was 1.61% (Table 2).

Subject #	Errors	Percent Error
1	10	2.29%
2	8	1.83%
4	3	0.69%
Average Percent Error		1.61%

Table 2. Results from the Human Experiment, with the result of subject 3 omitted. The average percent error after removing this outlier was 1.61%.

CONCLUSION

Until now, no human recognition rate on handwritten ZIP codes had been established. Therefore, this experiment establishes human performance in recognizing handwritten ZIP codes taken from the standard CEDAR database. The results of 1.61 % average percent error will allow researchers to compare their models to the ultimate ZIP code recognition system – that of the human visual system. However, due to the use of university undergraduate subjects, one must be cautious about over-generalizing the results. For example, if a sample of postal carriers were used as subjects, one may expect even better performance than that reported here.

REFERENCES

[1] CEDAR CDROM-1 Specifications of the Databases. Retrieved September 8, 2005, from http://www.cedar.buffalo.edu/Databases/CDROM1

[2] C.L. Liu, K. Nakashima, H. Sako, H. Fujisawa, Integrated Segmentation and Recognition of Handwritten Numerals: Comparison of Classification Algorithms, International Workshop on Frontiers in Handwritten Recognition (IWFHR), 8 (2002) 303-308.

[3] J. Mead, Speed Reading, Journal of University of Buffalo Research, (1991) vol1.1 page 2 and 3. Retrieved April 8, 2005, from http://www.cedar.buffalo.edu/pub_docs/article41.html.