Digit Detection Using Adaptive Spline Models

Ansuya Ahluwalia[1], Eric Kim[1], and Nicholas Brett Marcott[1]

University of California, Los Angeles
`ansuya@cs.ucla.edu, eric@ucla.edu, brett@ucla.edu`

Abstract. Automated handwriting detection remains an interesting yet challenging problem in the Vision field. Due to the curve-like nature of handwriting, it seems natural to consider approaches that directly model these curves. This project will investigate a particular approach from Hinton et. al [5] that uses an elastic model to recognize digits. Each digit class is represented by a cubic spline in an "ideal" configuration. To classify a test image, an iterative algorithm performs an elastic match between the test image and each digit model - the digit class with the best score wins. In addition, this project will investigate several extensions to the original model. Validation will be performed against the publically-available handwritten digit dataset, MNIST [1].

Keywords: Digit Recognition, Cubic Spline, Elastic Match, Expectation- Maximization

1 Introduction

Automated handwriting detection is an interesting yet challenging problem in the Vision field. Traditional approaches to detection problems, such as template-based matching, are ineffective in this domain due to nature of writing style variability: different people will write the same character in drastically different ways. Due to the curve-like nature of handwriting, however, it seems natural to consider approaches that attempt to model these curves.

On the one hand, pure template matching involves finding digit instances close to data and involves having a large repository of instances of each digit. This entails a large number of matching steps if the images are not accurately normalized. One solution is to use affine transformations, which significantly reduces this number. On the other hand, using elastic models increases the complexity of matching, while the number of matches decreases. In this project, we use elastic models based on splines, containing a few parameters which captures many of the variations of a given digit.

We investigate a particular approach from Hinton et. al [5] that uses an elastic model to recognize digits. Each digit class is represented by a cubic spline with a particular "home" configuration. To classify a test image, an iterative

algorithm is used that performs an elastic match between the test image and each digit class - the digit class with the best score wins, as shown in Fig. 1. To remain invariant against factors such as translation, rotation, and scaling, all matching is performed in a canonical "object frame". In addition, we investigate a few extensions to the original model. Validation is performed against the publically-available handwritten digit dataset, MNIST [1].
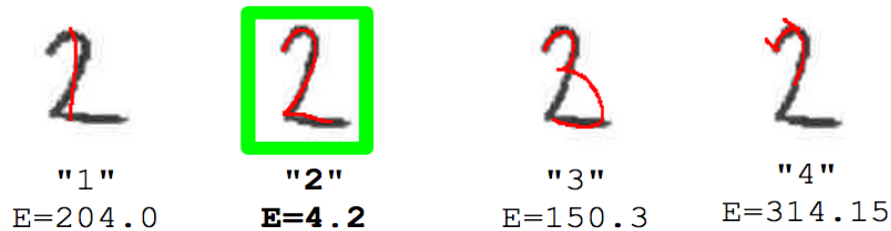

"1"   "2"   "3"   "4"
E=204.0   **E=4.2**   E=150.3   E=314.15

Fig. 1. Figure shows best fit model is the model with the lowest total energy (sum of deformation energy and fitting energy). Each digit model is matched to the test image, with energy scores to determine best fit.

2 Related Work

Prominent amount of work in the field of handwriting recognition revolves around neural networks (recurrent, hierarchical, deep forward etc.). Neural networks learn from an image training set for character recognition. Each neural network uniquely learns features that differentiate training images. The target image is classified based on the similarity in properties with the training images. Neural networks are quick to set up but can be inaccurate if unimportant properties are learnt in context to the target data.

For discrete character recognition, neural networks are trained on weights of template neural nets that generate the character (digit/ alphabet) with highest likelihood. Normalization is used to label the data. However, this system can only be applied to recognition of cursive and unconstrained words if templates of each word are made. This is due to the problem of character segmentation that exists in cursive and unconstrained handwriting. To handle unconstrained words, most researchers use Time-Delay Neural Networks (TDNN's) to generate labels for segments of data [8]. In [10], authors use a time-delay neural network to estimate a posteriori probabilities for characters in a word. A hidden markov model segments the word in a way that optimizes the global word score, using a dictionary in the process.

A lot of research has been in developing advanced algorithms for handwriting

recognition using neural networks and hidden markov models. In [7], authors use backpropagation to recognize handwritten zip codes. Researchers have also developed novel, biologically motivated approaches to classify handwritten characters. Dystal (Dynamically Stable Associative Learning) is one such neural network algorithm [3]. In [4], researchers use neocognitron, a neural network model for deformation-invariant visual pattern recognition, for handwritten alphanumeric character recognition. [6] incorporate HMM's into a complex stochastic language model for handwriting recognition. The pattern elements of the handwriting model are subcharacter stroke types modeled by HMMs. These HMMs are concatenated to form letter models, which are further embedded in a stochastic language model. [2] use a neural network/ HMM hybrid model for on-line handwriting recognition. They also employ EM algorithm in their approach. They perform a word-level normalization by fitting a model of the word structure using the EM algorithm.

Deformable models are efficient for characterizing handwritten digits since they have relatively few parameters and are able to capture many variations in digit instances. We investigate the system described in [5] that uses learned digit models consisting of splines whose shape is governed by a small number of control points. This method uses a single model for each digit class. As as extension to their work in [5], authors developed multi-models for each digit class, with little computational overhead, to better characterize the variations in the instances of each digit [9].

3   Methodology

In this section we discuss the details for performing digit recognition. Cubic B-spline models are defined for each digit class, with Gaussian distributions for their control points, in an object-based frame. Their affine transformations project the models into an image-based frame. The initial configuration consists of eight beads equally spaced along the spline. To fit the models to the data we use EM method. In each iterative step of the EM method, mixtures of Gaussians are fit to the data, dealing with the affine transformations from the object-based frame to the image-based frame and constraints on the Gaussians due to the spline. At each stage of fitting a model to data, the variance is decreased and the number of beads are progressively increased. The number of beads and the number of iterations at different variances are predefined in the annealing schedule. The goal is to achieve the best fitting model, determined by lowest total energy. We use an elastic matching algorithm to minimize an energy function that includes both the deformation energy of the digit model and the log probability that the model would generate the inked pixels in the image.

4 Results and Discussion

5 Conclusion

References

[1] The mnist database of handwritten digits. `http://yann.lecun.com/exdb/mnist`. (1998).

[2] Yoshua Bengio, Yann Lecun, Craig Nohl, and Chris Burges. Lerec: A nn/hmm hybrid for on-line handwriting recognition. *Neural Computation*, 7:1289–1303, 1995.

[3] K.T. Blackwell, T.P. Vogl, S.D. Hyman, G.S. Barbour, and D.L. Alkon. A new approach to hand-written character recognition. *Pattern Recognition*, 25(6):655 – 666, 1992.

[4] K. Fukushima and N. Wake. Handwritten alphanumeric character recognition by the neocognitron. *Neural Networks, IEEE Transactions on*, 2(3):355–365, May 1991.

[5] Geoffrey E. Hinton, Christopher K. I. Williams, and Michael Revow. Adaptive elastic models for hand-printed character recognition. In *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS*, pages 512–519. Morgan Kauffmann, 1992.

[6] Jianying Hu, M.K. Brown, and William Turin. Hmm based online handwriting recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(10):1039–1045, Oct 1996.

[7] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural Comput.*, 1(4):541–551, December 1989.

[8] J Mantas. An overview of character recognition methodologies. *Pattern recognition*, 19(6):425–430, 1986.

[9] Michael Revow, Christopher KI Williams, and Geoffrey E Hinton. Using mixtures of deformable models to capture variations in hand printed digits. 1993.

[10] M. Schenkel, I. Guyon, and D. Henderson. On-line cursive script recognition using time-delay neural networks and hidden markov models. *Machine Vision and Applications*, 8(4):215–223, 1995.