Piecewise Linear Modulation Model of Handwriting

H. Chen† O. E. Agazzi‡ C. Y. Suen†

Mailing Address

†GM 606, CENPARMI, Concordia University

1455 de Maisonneuve Blvd. West, Montreal, Quebec, Canada H3G 1M8

‡Room 2D-332, Bell Laboratories, Lucent Technologies

700 Mountain Avenue, Murray Hill, NJ 07974-0636, U.S.A.

Phone

†(514)848-7953 ‡(908)582-6568

Fax

†(514)848-4522 ‡(908)582-7308

Email Address

†hchen@cenparmi.concordia.ca

‡oea@research.bell-labs.com

Piecewise Linear Modulation Model of Handwriting

H. Chen† O. E. Agazzi‡ C. Y. Suen†

Abstract

A new piecewise linear modulation model of handwriting is proposed in this paper. In this model, the velocity of handwriting trajectory is modeled as the impulse response of a time varying second order system. For mathematical tractability, the entire trajectory is segmented into several non-overlapping frames, while the natural frequency and the damping factor of the system are assumed to vary linearly with time in each frame and are continuous along the entire trajectory. In other words, handwriting is regarded as an oscillation modulated by a continuous and piecewise linear signal. The parameters of this model are estimated by Powell's optimization algorithm which does not require the computation of the first order derivative. The number and lengths of the frames are decided by a modified binary search algorithm along with the estimation of parameters. This model has achieved very high data compression rate as well as accurate reproduction of real handwriting.

Keywords: model of handwriting, oscillation model, piecewise linear modulation model, parameter optimization, modified binary search algorithm

1 Motivations and Background

Handwriting OCR is regarded as one of the core technologies for the next generation of human computer interface. Although having been applied successfully in some cases, state-of-the-art OCR systems still fail to meet stringent requirement from a wide range of applications. [1]

High performance OCR systems depend vitally on accurate classifiers. Most current OCR classifiers have been designed without the knowledge of the intrinsic properties of handwriting, e.g. How is handwriting generated? What features do different characters written by the same writer have in common? What features do the same characters written by different writers have in common? These OCR classifiers simply regard handwriting as a pixel image (in the case of off-line recognition) or a trajectory (in the case of on-line recognition). As a consequence, they must be trained on very large data set and their performance depends on how well the training data represent the testing data. However, it is often very difficult, if not impossible, to manually collect adequate data to ensure the desired performance of the classifier. To overcome this difficulty, the model based recognition approach was proposed [2] [3]. The underlying philosophy is to incorporate the knowledge of the creation of handwriting into the design of the classifier. To this end, a model of handwriting must be established first. The parameters of a good model can usually be extracted from a relatively small amount of training data. Based upon these parameters, a classifier tightly matched to the model can be built, which is therefore more reliable and robust against the variations of handwriting. This model based approach could lead to major performance improvement of OCR systems.

A model of printed documents has been used very successfully in large-scale applications [4]. Models of handwritten documents have also been proposed in the literature [2]. However, they are far from satisfactory due to the large variation of handwriting. Before presenting our own model, it would be helpful to establish the criteria for a good handwriting model. We argue that a good handwriting model should:

- Be able to reproduce real handwriting precisely.
- Have a small number of parameters, i.e. it should achieve a high data compression rate.
- Have relatively stable parameters, i.e. the parameters for a given character should be largely insensitive to different writing styles.

The following sections will begin with a brief review of related work in this area. Then, a new piecewise linear modulation model of handwriting will be introduced and assessed using the above mentioned criteria. Experimental results and future direction will be given at the end.

2 Related Work

The piecewise linear modulation model of handwriting belongs to a family of models collectively called as *oscillation model.* They were put forward by Eden [5] and further investigated by Hollerbach [6]. Hollerbach assumes that handwriting can be described as a modulated oscillation in both the horizontal and vertical directions with a constant factor added to the horizontal velocity, which is in the form of:

$$\begin{cases} \dot{x}(t) = A_x(t) sin(\omega_x(t)t + \phi_x) + c \\ \dot{y}(t) = A_y(t) sin(\omega_y(t)t + \phi_y) \end{cases} \tag{1}$$

Hollerbach considers the zero-crossings of vertical velocity as the critical points to determine the shape of the character. He shows how different character shapes can be generated by changing the oscillation parameters.

Singer and Tishby [7] investigate how to extract parameters of the oscillation model from real handwritings. They segment the trajectory of handwriting into several frames whose boundaries are decided by the zero crossings of vertical velocity. They further assume that oscillation parameters stay constant inside each frame and they can only be changed at frame boundaries. Thus, zero crossings of vertical velocity become crucial points in their model.

However, as Stettiner and Chazan [8] pointed out, zero crossings of vertical velocity are extremely noise sensitive. According to their experiments, different instances of the same character produced by the same writer can have quite different number of zero crossings of vertical velocity. Thus, contrary to Singer and Tishby's method, they regard the handwriting trajectory as the impulse response of a slowly time-varying second order system. The parameters of the model describe the time evolution of the natural frequency and damping factors. To facilitate mathematical tractability, the trajectory is segmented into several non-overlapping frames of fixed and equal length and the parameters of the model vary linearly inside each frame. However, in their model, the decision on the number of frames for each character is quite heuristic and it does not carry from one writer to the other.

In this paper, we will propose a method which not only extracts the oscillation parameters but also determines the number of frames and their lengths from real handwriting data.

3 Piecewise Linear Modulation Model

In this model, the handwriting trajectory is considered as a modulated oscillation. Its velocity signal is modeled as the impulse response of a second order system whose natural frequency and damping factor vary slowly with time:

$$\hat{V}(t) = \frac{A}{\sqrt{1-\zeta(t)^2}} sin(\omega(t)\sqrt{1-\zeta(t)^2}\, t + \phi) exp\{-\zeta(t)\omega(t)t\} + C \qquad \text{n=1..N} \tag{2}$$

where

$\hat{V}(t)$: The horizontal/vertical velocity of handwriting

A: The amplitude

ϕ: The initial phase

$\omega(t)$: The natural frequency

$\zeta(t)$: The damping factor

C: Constant factor

3.1 Piecewise Linear Parameters

To represent the fact that different segments of the trajectory have different characteristic shapes, the entire handwriting trajectory is segmented into several non-overlapping frames and each frame represents a stroke with a special shape. Within each frame, the natural frequency $\omega(t)$ and damping factor $\zeta(t)$ vary linearly with time. Due to the continuity of handwriting, $\omega(t)$ and $\zeta(t)$ must be continuous along the entire trajectory. In other words, they are continuous and piecewise linear functions, as illustrated in Fig. 1

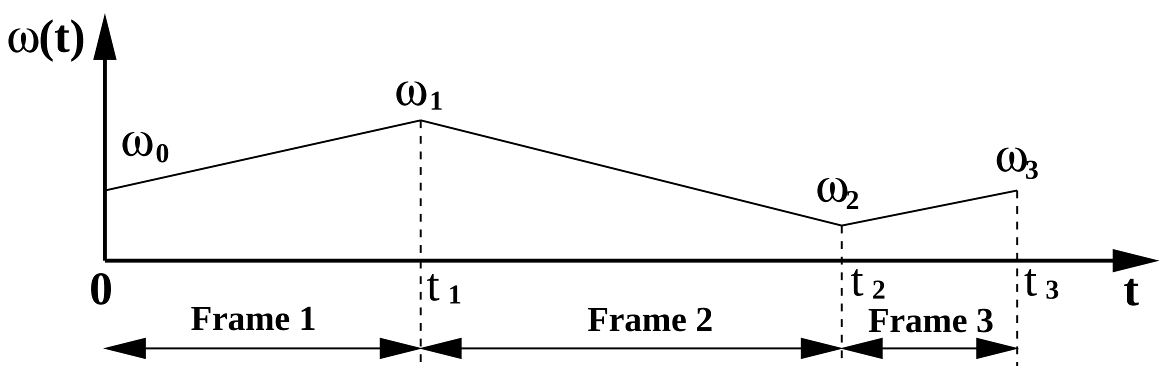

Figure 1: Piecewise Linear Function

3.2 Feature Vector

The feature vector of a character is composed of the amplitude A, the initial phase ϕ, the constant factor C, and a set of natural frequencies $\omega(t)$ and damping factors $\zeta(t)$ for each frame:

$$\theta = \{A, \phi, C, (\zeta_i, \omega_i)\} \qquad \text{i=1..P+1, where P is the number of frames} \tag{3}$$

3.3 Estimation of Parameters

The velocity signal calculated from the handwriting model is an estimation of the real velocity signal, and there exists an error signal which is the difference between them:

$$E(t) = V(t) - \hat{V}(t) \tag{4}$$

where V(t) is the real handwriting velocity and $\hat{V}(t)$ is the velocity calculated from the model. The feature vector can be estimated by minimizing the mean square error:

$$\hat{\theta} = argmin_\theta \sum_t E(t)^2 \tag{5}$$

This is a multi-dimensional non-linear optimization problem, which can be effectively solved by a number of methods. Some of them require the calculation of the derivative of the cost function (such as Conjugate Gradient Method), but some of them do not (such as Powell's Method) [9]. The derivative based methods rely on the derivative values of the cost function for the direction along which they will search in the multi-dimensional space. However, our model function is only meaningful in a fixed range of its parameters ($|\zeta(t)| < 1$) and thus its derivative is undefined outside this range, which will cause troubles for the derivative based methods. However, for the non-derivative based methods, derivative values are not required to guide the search. When the parameters are out of range, we simply evaluate the cost function to be an artificial value (such as a very large number), which efficiently limits the scope of search into meaningful region. For this reason, Powell's method is used in our experiment.

3.4 Estimation of Frames

In Stettiner and Chazan's original model, the entire trajectory is segmented into frames of fixed and equal lengths. However, this approach has several liabilities. First, the decision on how many frames are required for each character is quite empirical and lack of objective criteria. Second, the real handwriting trajectory is generally not composed of strokes of equal length. It may have one long stroke followed by another short one. Thus, the model which segments the trajectory into equal length strokes does not represent real handwriting precisely.

To overcome the above mentioned problem, in our model the trajectory is divided into a variable number of unequal frames. The number and lengths of the frames are decided by a modified binary search algorithm which is carried out together with the estimation of model parameters.

Let E(t) denote the minimum mean square error between the real and synthetic velocity signal inside the frame of length `t` starting from a fixed point on the trajectory. Obviously, E(t) must be monotonically increasing. The optimization process proceeds as follows:

At the beginning, a short starting segment of the trajectory is used as the initial frame and its E(t) is computed. If E(t) is less than a threshold, the frame length is extended and E(t) is computed again. However, if E(t) is larger than the threshold, the frame length is decreased until it reaches a value t_0, where $E(t_0) \leq threshold$ and $E(t_0+1) > threshold$. Then, the first segment of length `t` of the trajectory is treated as the first frame. The optimization process continues on the rest of the trajectory, until the entire trajectory is covered. Since E(t) is monotonically increasing, a modified binary search algorithm is applied to extend or shrink the length of the frame to accelerate computation. This process is illustrated in the following pseudo code:

```
using a short starting segment of the trajectory as the initial frame
while (end of trajectory has not been reached)
{
  compute MSE inside the frame;
  if (MSE < threshold)
    increase the length of the frame;
  else
    if (boundary position right before has been tested)
      /* and was successful */
      create a new frame;
    else
      decrease the length of the frame;
}
```

4 Experimental Results

4.1 Experiment

Handwriting trajectory is collected from a graphical tablet (Wacom PL100V) connected to a PC running Linux 1.2. The horizontal and vertical resolutions of the tablet are both 20 pixels/mm. The handwriting trajectory is sampled at a constant sampling rate. The data is then transmitted to a SGI workstation running IRIX 5.2 and all subsequent computation of the model is performed on SGI workstation. The data pass through a low pass filter before the subsequent experiments.

4.2 Results

Fig. 2 shows respectively the trajectory, horizontal and vertical velocity signals of the real and synthetic character `d` . Only two frames are necessary to model this complicated shape. Fig. 3 shows the same for character `m`. In this case, only one frame is necessary to represent the shape. Fig. 4 shows character `8`. Fig. 5 and Fig. 6 show the real versus synthetic shapes of many lowercase characters. These figures clearly suggest that our model can precisely represent the real handwriting of characters.

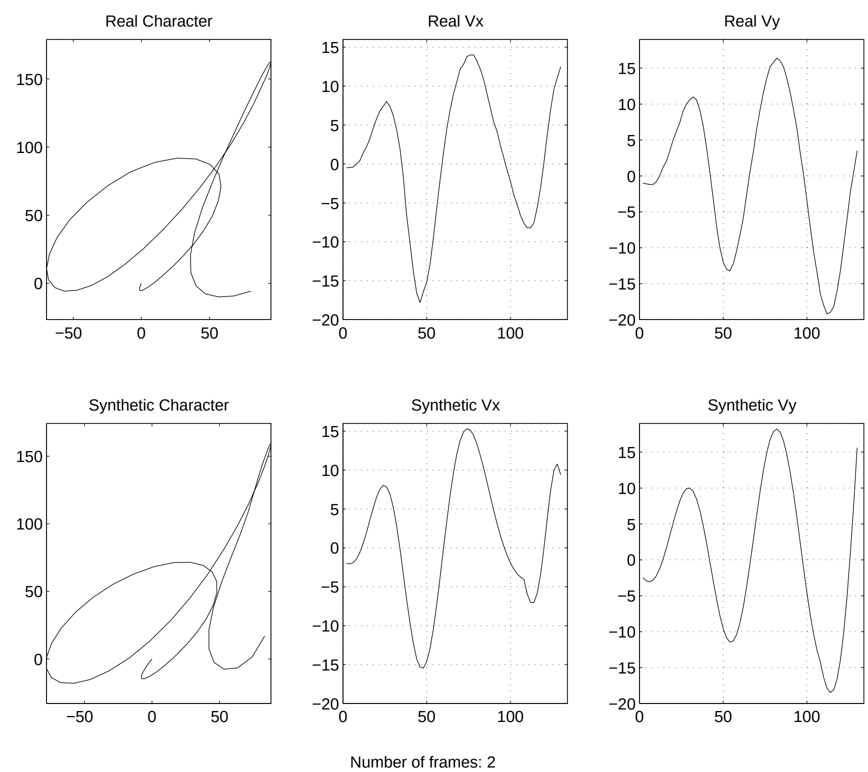

Figure 2: Real and Synthetic Character d

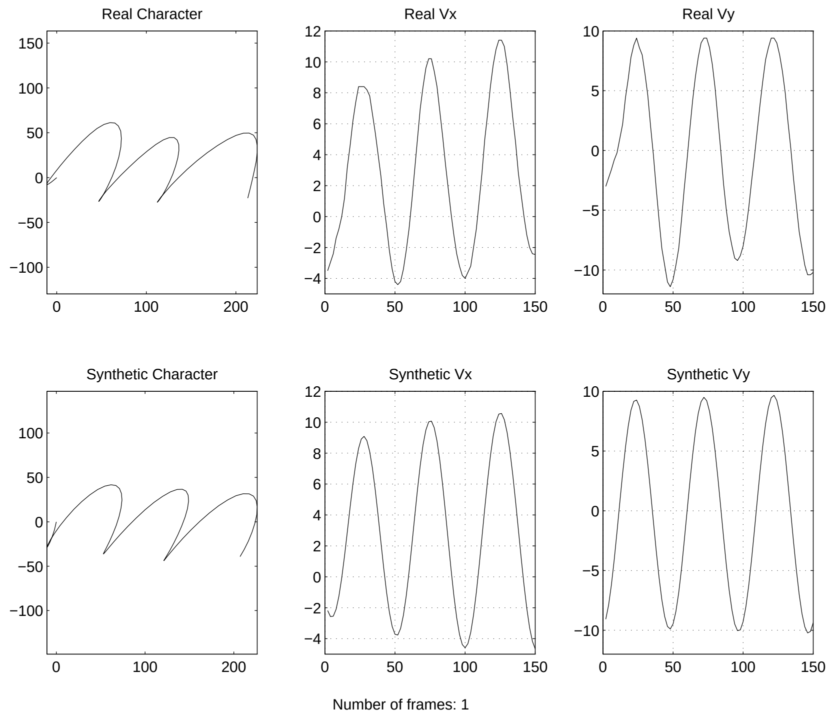

Figure 3: Real and Synthetic Character m

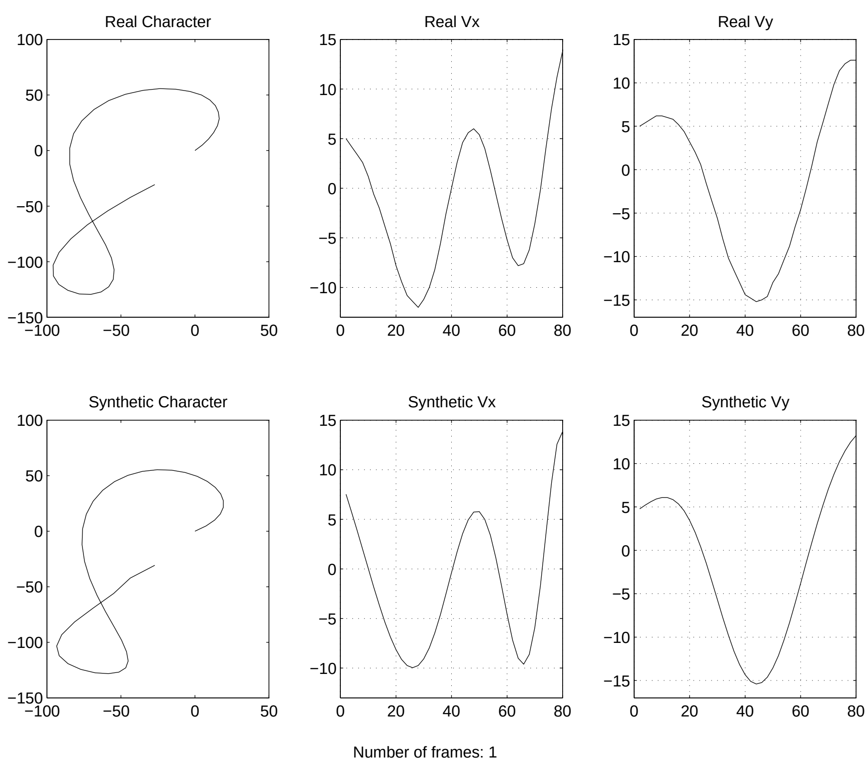

Figure 4: Real and Synthetic Character 8

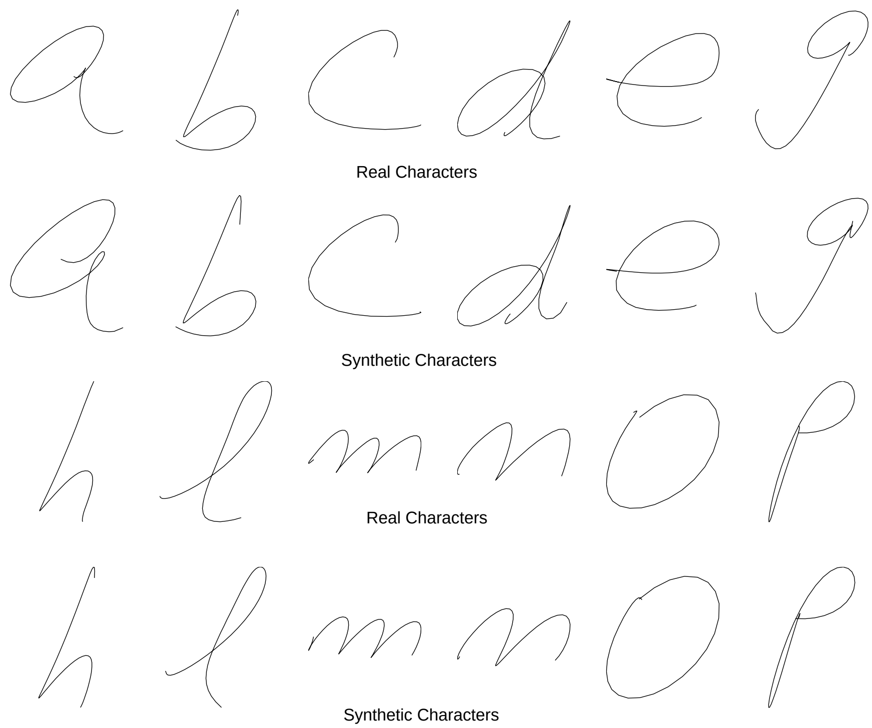

Figure 5: Some Real and Synthetic Lowercase Characters, Part 1

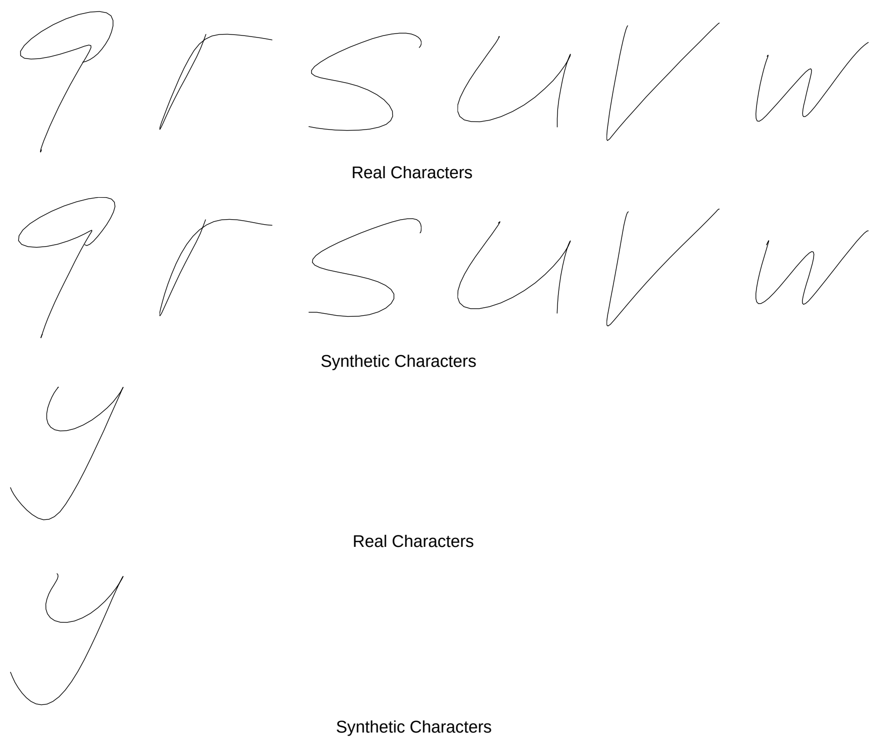

Figure 6: Some Real and Synthetic Lowercase Characters, Part 2

5 Conclusion & Discussion

At the beginning of this paper, we proposed three criteria for a good handwriting model, i.e. (1) precise reproduction of real handwriting, (2) high data compression rate, (3) stability of parameters. Examine our model according to these criteria and it can be found that:

- Our model is capable of reproducing synthetic handwriting characters that are very close to the real ones.
- Our model has achieved very high data compression rate while keeping most of the information to restore the original character. In our experiment, most of the characters can be modeled with only one frame, which requires only 14 parameters. Considering the fact that the trajectory of one character has about 100 points (thus 200 data entries) in our experiment, our model has achieved a data compression rate at the order of 14.
- Linear Predictive Coding (LPC) and short time Fourier analysis on the velocity signal suggest that the natural frequency ω is quite stable for a given writer[10].

Compared to the original piecewise linear modulation model by Stettiner and Chazan, our model segments the entire handwriting trajectory into a variable number of frames with unequal lengths and the segmentation is carried out dynamically together with the estimation of model parameters. Two obvious advantages result from this approach:

Higher data compression rate Our model tries to minimize the number of frames required for any given character. It does not create a new frame unless deemed necessary in the estimation process. Compared to the approach of fixed and equal length frame, our model reduces the number of parameters for most characters.

Better Optimization Our model starts the optimization from a short initial frame of the trajectory, where the cost function is relatively simple and it helps to reach a global minimum. When the length of the frame increases, the complexity of the cost function increases gradually, and the new global minimum drifts away from its previous position with a small displacement. This helps to reach the global minimum each time a new frame length is tried.

Further work on this project involves validating the stability of parameters of the model by experiment on large scale database, although we have got some preliminary evidence. Also our experiments are carried out on handwritten characters only. It would be more valuable but straight forward to extend them to cursive handwritings, where conceivably more frames will be generated by the algorithm of the model.

6 Acknowledgement

Special thanks go to Dr. Henry S. Baird at Bell Laboratories, Lucent Technologies for many exciting discussions and comments.

References

[1] S. Mori, C. Y. Suen, and K. Yamamoto. Historical review of ocr research and development. *Proc. of the IEEE*, 80:1029–1058, 1992.

[2] R. Plamondon and F. J. Maarse. An evaluation of motor models of handwriting. *IEEE Trans. on System, Man, and Cybernetics*, 19(5):1060–1072, 1989.

[3] A. M. Alimi and R. Plamondon. Performance analysis of handwritten strokes generation models. In *Third International Workshop on Frontiers in Handwriting Recognition*, Buffalo, N.Y., 1993.

[4] H. S. Baird. Document image defect models. In *Structured Document Image Analysis*, pages 546–556. Springer-Verlag, 1992.

[5] M. Eden. Handwriting and pattern recognition. *IRE Trans. on Information Theory*, pages 160–166, 1961.

[6] J. M. Hollerbach. An oscillation theory of handwriting. *Biol. Cybern.*, 39:139–156, 1981.

[7] Y. Singer and N. Tishby. Dynamical encoding of cursive handwriting. *Biol. Cybern.*, 71:227–237, 1994.

[8] O. Stettiner and D. Chazan. A statical parametric model for recognition and synthesis of handwriting. In *Proc. of 12th Int. Conf. on Pattern Recognition*, pages 34–38, 1994.

[9] W. H. Press et al. *Numerical Recipes in C.* Cambridge University Press, 2 edition, 1992.

[10] L. R. Rabiner and R. W. Schafer. *Digital Processing of Speech Signals.* Prentice-Hall Inc., 1978.