

# Dynamical encoding of cursive handwriting

Y. Singer, N. Tishby

Institute of Computer Science and Center for Neural Computation, The Hebrew University, Jerusalem 91904, Israel

Received: 13 April 1993 / Accepted in revised form: 18 February 1994

**Abstract.** A model-based approach to on-line cursive handwriting analysis and recognition is presented and evaluated. In this model, on-line handwriting is considered as a modulation of a simple cycloidal pen motion, described by two coupled oscillations with a constant linear drift along the line of the writing. By slow modulations of the amplitudes and phase lags of the two oscillators, a general pen trajectory can be efficiently encoded. These parameters are then quantized into a small number of values without altering the writing intelligibility. A general procedure for the estimation and quantization of these cycloidal motion parameters for arbitrary handwriting is presented. The result is a discrete *motor control representation* of the continuous pen motion, via the quantized levels of the model parameters. This motor control representation enables successful word spotting and matching of cursive scripts. Our experiments clearly indicate the potential of this dynamic representation for complete cursive handwriting recognition.

---

## 1 Introduction

Cursive handwriting is a complex graphic realization of natural human communication. Its production and recognition involve a large number of highly cognitive functions including vision, motor control, and natural language understanding. Yet the traditional approach to handwriting recognition has focused so far mostly on computer vision and computational geometric techniques. The recent emergence of pen computers with high resolution tablets has made available dynamic (temporal) information as well and created the need for robust on-line handwriting recognition algorithms. Considerable effort has been spent in the past years on on-line cursive handwriting recognition (for general reviews see Plamondon et al. 1989; Plamondon and Leedham 1990; Tappert et al. 1990), but there are no robust, low error rate recognition schemes available yet.

Research of the motor aspects of handwriting has suggested that the pen movements produced during cursive handwriting are the result of 'motor programs' controlling the writing apparatus. This view was used for natural synthesis of cursive handwriting (see e.g., E. Doojies, pp. 119–130 in Plamondon et al. 1989). There have been several attempts to construct dynamical models of handwriting for recognition. Some of these works are based on a similar approach to ours (e.g., Rumelhart 1992). None of the previous works, however, have actually solved the inverse dynamics problem of 'revealing' the 'motor code' used for the production of cursive handwriting.

Motivated by the oscillatory motion model of handwriting, as introduced by, e.g., Hollerbach (1981), we develop a robust parameter estimation and regularization scheme which serves for the analysis, synthesis, and coding of cursive handwriting. In Hollerbach's model, cursive handwriting is described by two independent oscillatory motions superimposed on a constant linear drift along the line of writing. When the parameters are fixed, the result of these dynamics is a cycloidal motion along the line of the drift (see Fig. 1). By modulations of the cycloidal motion parameters, arbitrary handwriting can be generated. The difficulty, however, is to generate writing by a *low rate* modulation, much lower than the original rate of the oscillatory signals. In this work, we propose an efficient low rate encoding of the cycloidal motion modulation and demonstrate its utility for robust synthesis and analysis of the process.

The pen trajectory is discretized in time by considering only the zero vertical velocity points. In between these points, the handwriting is approximated by an unconstrained cycloidal motion using the values of the parameters estimated at the zero vertical velocity points. Further, we show that the amplitude modulation can be quantized to a small number of levels (five for the vertical amplitude modulation and three for the horizontal amplitude modulation), and the results are robust. The vertical oscillation is described as an almost synchronous process, i.e. the angular velocity is transformed to be

constant. The horizontal oscillation is then described in terms of its phase lag to the vertical oscillation and thus becomes synchronous as well. The modeling and estimation processes can be viewed as a many-to-one mapping, from the continuous *pen motion* to a discrete set of *motor control symbols*. While this dramatically reduces the coding bit rate, we show that the relevant recognition information is regularized and preserved.

In Sect. 2, we discuss Hollerbach’s model and demonstrate its advantages in representing handwriting over standard geometric techniques. In Sect. 3, we describe our analysis-by-synthesis methodology and define the goal to be an efficient motor encoding of the process. In Sect. 4 we introduce two global transformations: correction of the writing orientation and slant equalization. We show that such preprocessing further assists in regularizing the process, which simplifies the parameter estimation phase. In Sect. 5 we discuss the model’s parameter estimation. Sects. 6-8 introduce a series of quantizations and discretizations of the dynamic parameters, which both lower the encoding bit rate and improve the readability of the writing. Section 9 summarizes the discrete representation of the cursive handwriting process and shows that this representation is stable in the sense that similar words result in similar *motor control symbols*. Finally, we describe word spotting and matching experiments which demonstrate the power of our approach for recognition.

## 2 The cycloidal model

Handwriting is generated by the human motor system, which can be described by a spring muscle model near equilibrium. This model assumes that the muscle operates in the linear small deviation regions. Movements are excited by selecting a pair of agonist-antagonist muscles, modeled by a spring pair. If we further assume that the friction is balanced by an equal muscular force, then the process of handwriting can be approximated by a system of two orthogonal opposing pairs of ideal springs. In a general form, the spring muscle system can be described by the following differential equation

$$M \begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} = -K \begin{bmatrix} x \\ y \end{bmatrix} \quad (1)$$

where  $M$  and  $K$  are  $2 \times 2$  matrices that can be diagonalized simultaneously. This system can be transformed to a diagonalized system described by the following decoupled equations set

$$\begin{cases} M_x \ddot{x} = K_{1,x}(x_1 - x) - K_{2,x}(x - x_2) \\ M_y \ddot{y} = K_{1,y}(y_1 - y) - K_{2,y}(y - y_2) \end{cases} \quad (2)$$

where  $K_{1,x}, K_{2,x}, K_{1,y}, K_{2,y}$  are the spring constants, and  $x_1, x_2, y_1, y_2$  are the spring equilibrium positions. Solving these equations with the initial condition that the system has a constant velocity (drift) in the horizontal direction yields the following parametric form

$$\begin{cases} x(t) = A \cos(\omega_x(t - t_0) + \phi_x) + C(t - t_0) \\ y(t) = B \cos(\omega_y(t - t_0) + \phi_y) \end{cases} \quad (3)$$

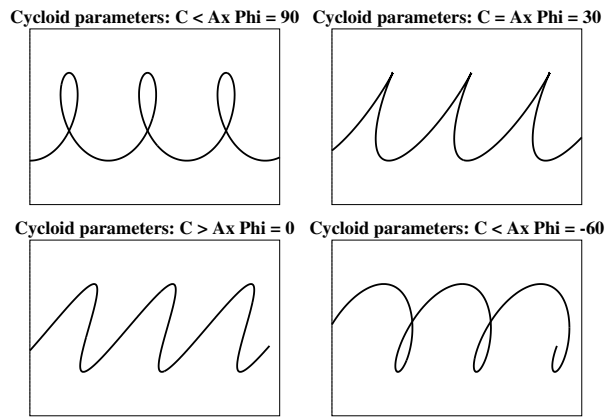


Fig. 1. Various cycloidal writing curves

The angular velocities  $\omega_x$  and  $\omega_y$  are determined by the ratios between the spring constants and masses.  $A, B, C, \phi_x, \phi_y, t_0$  are the integration parameters determined by the initial conditions. This set describes two independent oscillatory motions, superimposed on a linear constant drift along the line of writing, generating cycloids. Different cycloidal trajectories can be achieved by changing the spring constants and zero settings at the appropriate time. The relationship between the horizontal amplitude modulation  $A_x(t)$ , the horizontal drift  $C$ , and the phase lag,  $\phi(t) = \phi_x(t) - \phi_y(t)$ , controls the letter corner shape (cusp), as demonstrated in Fig. 1.

We further restrict the model by assuming that the angular velocities are tied, i.e.  $\omega_x(t) \approx \omega_y(t) \triangleq \omega(t)$ , and that  $\phi_y(t) = 0$ . These assumptions are not too restrictive, as will be shown later. With these assumptions, the equations governing the oscillations in the velocity domain can be written as

$$\begin{cases} V_x(t) = A_x(t) \sin(\omega(t)(t - t_0) + \phi(t)) + C \\ V_y(t) = A_y(t) \sin(\omega(t)(t - t_0)) \end{cases} \quad (4)$$

where  $t_0$  is the writing onset time,  $A_x(t)$  and  $A_y(t)$  are the horizontal and the vertical instantaneous amplitude modulations,  $\omega(t)$  is the instantaneous angular velocity,  $\phi(t)$  is the horizontal phase lag, and  $C$  is the horizontal drift velocity. By definition, the oscillation phase  $\theta(t) = \int_0^t \omega(t) dt$  is monotonic in time. Hence, the time parameterization of the velocity equations can be changed, using the chain rule,  $\frac{dX}{dt} = \frac{dX}{d\theta} \frac{d\theta}{dt}$ , to phase parameterization of the following form

$$\begin{cases} V_x(\theta) = A_x(\theta) \sin(\theta + \phi(\theta)) + C \left[ \frac{dt}{d\theta} \right] \\ V_y(\theta) = A_y(\theta) \sin(\theta) \end{cases} \quad (5)$$

As already demonstrated, different cycloid parameters yield different letter forms. The transition from one letter to another can be achieved by a gradual change in the parameter space. A smooth pen trajectory can be obtained in this way. Standard differential geometry parameterizations (e.g., curvature versus arc-length), however, have difficulties expressing infinite curvature (corners), which are handled naturally in our model.

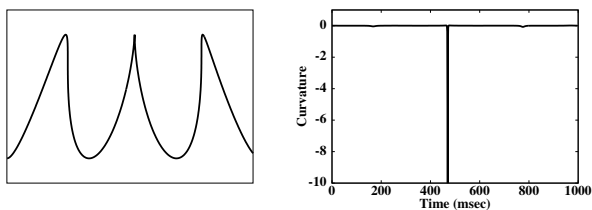


Fig. 2. A synthetic cycloid and its curvature

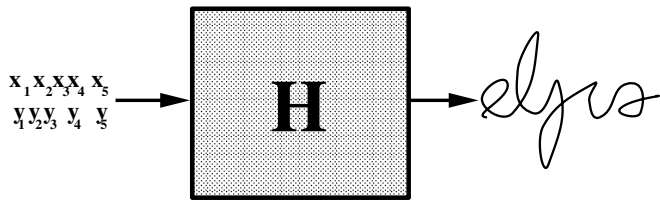


Fig. 3. A discrete controlled system that maps *motor control symbols* to pen trajectories

This problem is demonstrated in Fig. 2. In this simple example, a cycloid trajectory was produced by setting the parameters  $A_x(t) = A_y(t) = C = 1$  and gradually changing  $\phi(t)$  from  $0^\circ$  to  $+180^\circ$ . The resulting trajectory after integration of the velocities is a smooth curve which has the form of the letter **w**. However, the curvature diverges at the middle cusp.

### 3 Methodology

Using the velocity equation presented in the previous section, handwriting can be represented as a slowly varying dynamical system whose control parameters are the cycloidal parameters  $A_x(t)$ ,  $A_y(t)$ , and  $\phi(t)$ . In this work, it is shown that these dynamical parameters have an efficient discrete coding that can be represented by a discretely controlled, dynamical system. The inputs to this system are *motor control symbols* which define the instantaneous cycloidal parameters. These parameters change only at restricted times. Our motor system ‘translates’ these motor control symbols to continuous arm movements. An illustration of this system is given in Fig. 3 where the system is denoted by  $H$  and the control symbols by  $(x_i, y_i)$ .

Decoding and recognition, as implied by this model, are done by solving an *inverse dynamics* problem. The following sections describe our solution to this inverse problem. A series of parameter estimation schemes that reveal the discrete control symbols is presented. Each stage in the process is verified via an *analysis-by-synthesis* technique. This technique uses the estimated parameters and the underlying model to reconstruct the trajectory. At every stage the synthesized curve is examined to determine whether the relevant recognition information is preserved. The result is a mechanism that maps the continuous pen trajectories to the discrete motor control symbols. A more systematic approach which uses control theoretical schemes is being developed.

## 4 Global transformations

On-line handwriting need not be oriented horizontally, and usually the handwriting is slanted. In this section normalization processes that eliminate different writing orientations and writing slants are described. These transformations are performed prior to any modeling to make the input scheme more robust. In this process, we do not estimate any of the dynamic parameters but use the general form of the dynamic equations.

### 4.1 Correction of the writing orientation

On-line handwriting is sampled in a general unconstrained position. This results in a nonhorizontal direction of writing. Even when the writing direction is horizontal, there are position variations due to the oscillations; thus, the general orientation is defined as the average slope of the trajectory. Robust statistic estimation (Wald 1940) is used to estimate the general orientation, rather than a simple linear regression, since there are measurement errors both in the vertical and the horizontal pen positions. The sampled points  $(X(i), Y(i))$  are randomly divided into pairs  $\{(X(2i_k), Y(2i_k)), (X(2i_k + 1), Y(2i_k + 1))\}$ , such that  $X(2i_k + 1) \geq X(2i_k)$ . The estimated writing orientation is  $\hat{W} = \frac{\sum_k \{Y(2i_k + 1) - Y(2i_k)\}}{\sum_k \{X(2i_k + 1) - X(2i_k)\}}$ . The angle of the writing direction  $\alpha$  is  $\tan^{-1} \hat{W}$ , and the velocity vectors are rotated as follows

$$\begin{cases} V'_x(t) = V_x(t) \cos(\alpha) + V_y(t) \sin(\alpha) \\ V'_y(t) = -V_x(t) \sin(\alpha) + V_y(t) \cos(\alpha) \end{cases} \quad (6)$$

### 4.2 Slant equalization

Handwriting is normally slanted. In the spring muscle model, this implies that the spring pairs are not orthogonal and only the general (1) is valid. The amount of coupling can be estimated by measuring the correlation between the horizontal and vertical velocities. Removing the slant is equivalent to decoupling the oscillation equations. This decoupling is desired since it is a writer-dependent property which does not contain any context information. The decoupling enables an independent estimation of the oscillation parameters for the *phase-lag regularization* stage, described in Sect. 7, and simplifies the estimation scheme.

The decoupling can be viewed as a transformation from a nonorthogonal to an orthogonal coordinate system in which one of the axes is the direction of writing. The horizontal velocity after slant equalization (denoted by  $\tilde{V}_x$ ) is statistically uncorrelated with the vertical velocity  $V_y$  ( $\tilde{V}_x - V_y$ ). Therefore, the original velocity can be written as  $V_x(t) = \tilde{V}_x + A(t)V_y(t)$ . Assuming stationarity, this requirement means that  $E(\tilde{V}_x V_y) = 0$  and  $A(t) = A$ . If we assume that the slant is almost constant, then the stationarity assumption holds. The

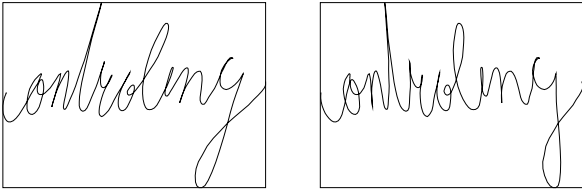


Fig. 4. The result of the slant equalization process

*maximum likelihood* estimator for  $A$ , assuming that the measurement noise is gaussian, is

$$\hat{A} = \frac{E(V_x V_y)}{E(V_y V_y)} = \frac{\sum_{t=1}^N V_x(t) V_y(t)}{\sum_{t=1}^N V_y(t) V_y(t)} \quad (7)$$

There are writers whose writing slant changes significantly even within a single word. For those writers, the *projection coefficient*  $A(t)$  is estimated locally. We assume though that along a short interval the slant is constant (local stationarity assumption). In order to estimate  $A(t_0)$  we compute the short time correlation between  $V_x(t)$  and  $V_y(t)$  after multiplying them by a window centered at  $t_0$

$$\hat{A}(t_0) = \frac{\sum_{t=1}^N V_x(t) V_y(t) W(t_0 - t)}{\sum_{t=1}^N V_y(t) V_y(t) W(t_0 - t)} \quad (8)$$

where  $W$  is a *Hanning window*<sup>1</sup>, frequently used in *short time Fourier analysis* applications (Oppenheim and Schaffer 1975). We empirically set the width of the window to contain about 5 cycles of  $V_y$ . After finding  $\hat{A}(t)$  (or  $\hat{A}$  if we assume a constant slant) the horizontal velocity after slant equalization is  $\hat{V}_x(t) = V_x(t) - \hat{A}(t) V_y(t)$ . The slant equalization process is depicted in Fig. 4, where the original handwriting is shown with the handwriting after slant equalization with a stationary slant assumption.

## 5 Estimating the model parameters

The cycloidal (4) is too general. The problem of estimating its continuous parameters is ill-defined since there are more parameters than observations. Therefore, we would like to constrain the values of the parameters while preserving the intelligibility of the handwriting. It is shown in this section that by restricting the values of the parameters, a compact coding of the dynamics is achieved while preserving intelligibility.

Assuming that the model is a good approximation of the true dynamics, then the horizontal drift,  $C$ , can be estimated as follows,  $\hat{C} = \frac{1}{N} \sum_{i=1}^N V_x(n)$ , where  $N$  is the number of digitized points. Under the model assumptions,  $\hat{C}$  converges to  $C$  and is an unbiased estimator. In order to check the assumption that  $C$  is really constant we calculated it for different words and locally within a

<sup>1</sup> Hanning window of length  $N$  is defined as  $W_{Hanning}(n) = \frac{1}{2} (1 - \cos(\frac{2\pi n}{N-1}))$ .

word using a sliding window. The small variations in the estimator  $\hat{C}$  indicate that our assumption is correct. At this point we perform one more normalization by dividing the velocities  $V_x(t)$  and  $V_y(t)$  by  $\hat{C}$ . The result is a set of normalized equations with  $C = 1$ . Henceforth, the constant drift is subtracted from the horizontal velocity and it is added whenever the spatial signal is reconstructed. Integration of the normalized set results in a fixed height handwriting, independent of its original size. The normalizations and transformations presented so far are supported by physiological experiments (Hogan and Flash 1987; Lacquiniti 1989) that show evidence of spatial and temporal invariance of the motor system.

We assume that the cycloidal trajectory describes the natural pen motion between the velocity zero-crossings and changes in the dynamical parameters occur at the zero-crossings only, to keep the continuity. This assumption implies that the angular velocities  $\omega_x(t)$ ,  $\omega_y(t)$  and the amplitude modulation  $A_x(t)$ ,  $A_y(t)$  are almost constant between consecutive zero-crossings. A good approximation can be achieved by identifying the velocity zero-crossings, setting the local angular velocities to match the time between two consecutive zero-crossings, and setting the velocities to values such that the total pen displacement between two zero-crossings is preserved. Denote by  $t_i^x$  and  $t_i^y$  the  $i$ th zero-crossing of the horizontal and vertical locations, and by  $L_i^x$  and  $L_i^y$  the horizontal and vertical progression during the  $i$ th interval (after subtracting the horizontal drift), respectively. The estimated amplitudes are

$$\int_{t_i^x}^{t_{i+1}^x} \hat{A}_i^x \sin(\frac{\pi}{t_{i+1}^x - t_i^x} (t - t_i^x)) dt = L_i^x \Rightarrow \hat{A}_i^x = \frac{L_i^x \pi}{2(t_{i+1}^x - t_i^x)}$$

$$\int_{t_i^y}^{t_{i+1}^y} \hat{A}_i^y \sin(\frac{\pi}{t_{i+1}^y - t_i^y} (t - t_i^y)) dt = L_i^y \Rightarrow \hat{A}_i^y = \frac{L_i^y \pi}{2(t_{i+1}^y - t_i^y)}$$

The angular velocities are set independently and the phase lag,  $\phi(t)$ , is currently set to 0. The result of this process is a compact representation of the writing process, demonstrated by the resynthesized curve which is similar to the original, as shown in Fig. 5.

At this stage we can represent the writing process as two statistically independent, single-dimensional, oscillatory movements. Free oscillatory movement is assumed between consecutive zero-crossings, while switching of the dynamic parameters occurs only at these points.

Each of the original sampled points, denoted by  $(x, y)$ , is quantized to 8 bits. Quantizing the amplitudes and the zero-crossings indices to 8 bits reduces the number of bits needed to represent the curve, as shown in Fig. 13. The original code length is indexed as stage 1. Stage 2 is the velocity approximation described in this section. The total description length of the trajectory at this point is reduced by a factor of 7.

## 6 Amplitude modulation discretization

The amplitudes  $A_x(t)$ ,  $A_y(t)$  define the vertical and horizontal scale of the letters. From measurements of written words, the possible values of these amplitudes appear to be limited to a few typical values with small variations.

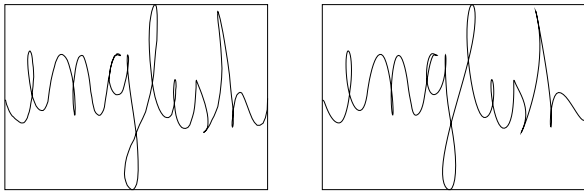


Fig. 5. The original and the reconstructed handwriting after amplitudes coding

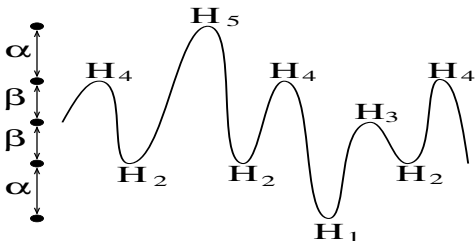


Fig. 6. Illustration of the vertical positions as a function of time

We assume statistical independence of the amplitude values and perform discretization separately for the horizontal and vertical velocities. Nevertheless strong correlations remain between the velocities, which can be reduced in later stages.

### 6.1 Vertical amplitude discretization

Examination of the vertical velocity dynamics reveals the following:

- There is a virtual center of the vertical movements. The pen trajectory is approximately symmetric around this center.
- The vertical velocity zero-crossings occur while the pen is at almost fixed vertical levels, which correspond to high, normal and small modulation values.

These observations are presented in Fig. 6, where the vertical position is plotted as a function of time. Using this apparent quantization we allow five possible pen positions, denoted by  $H_1, \dots, H_5$ , which satisfy the symmetry constraints,  $\frac{1}{2}(H_1 + H_5) = \frac{1}{2}(H_2 + H_4) = H_3$ . Let,  $\alpha = H_2 - H_1 = H_5 - H_4$  and  $\beta = H_3 - H_2 = H_4 - H_3$  (Fig. 6). Then, the possible curve lengths are,  $0, \alpha, \beta, \alpha + \beta, \alpha + 2\beta, 2\beta, 2(\alpha + \beta)$ .

The five-level description is a qualitative view. The levels achieved at the vertical velocity zero-crossings vary around  $H_1, \dots, H_5$ . The variation around each level is approximated by a normal distribution with an unknown common variance. The distributions around the levels are assumed to be fixed and characteristic for each writer. Let  $I_t$  ( $I_t \in 1, \dots, 5$ ) be the level indicator, i.e., the index of the level obtained at the  $t$ th zero-crossing. We need to estimate concurrently the five mean levels  $H_1, \dots, H_5$ , their common variance  $\sigma$ , and the indicators  $I_t$ . Yet the observed data are just the actual levels,  $L(t)$ , which are composed of the ‘true’ levels,  $H_{I_t}$ , and an observation gaussian noise  $\xi$ ,  $L(t) = H_{I_t} + \xi$  ( $\xi \sim N(0, \sigma)$ ).

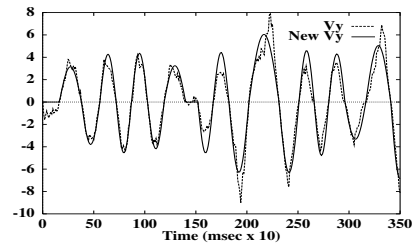


Fig. 7. The original and the quantized vertical velocity (*top*), the original handwriting (*bottom left*), and the reconstructed handwriting after quantization of the horizontal and vertical amplitudes (*bottom right*)

Therefore, the *complete data* consist of the sequence of levels and indicators  $\{I_t, L(t)\}$ , while the observed data (also termed *incomplete data*) are just the sequence of levels,  $L(t)$ . The task of estimating the parameter  $\{H_i, \sigma\}$  is a classical situation of *maximum likelihood* parameters estimation from *incomplete data*, commonly solved by the *EM* algorithm (Dempster et al. 1977). A full description of the use of *EM* in our case is given in the Appendix. The handwriting synthesized from the quantized amplitudes is depicted in Fig. 7.

### 6.2 Horizontal amplitude discretization

The quantization of the horizontal progression between two consecutive velocity zero-crossings is simpler. In general, there are three types of letters, thin (like *i*), normal (*n*), and fat (*o*). These typical levels can be found using a standard scalar quantization technique.

## 7 Horizontal phase lag regularization

After performing slant equalization, the velocities  $V_x(t)$  and  $V_y(t)$  are approximately statistically uncorrelated. Since  $\omega_x \approx \omega_y$ , the two velocities can be statistically uncorrelated if the phase lag between  $V_x$  and  $V_y$  is  $\pm 90^\circ$  on the average. Thus, the horizontal velocity,  $V_x$ , is close to its local extrema, while  $V_y$  is near zero, and vice versa. Since the phase lag changes continuously, a change from a positive phase lag to a negative one (or vice versa) must pass through  $0^\circ$ . There are places of local halt in both velocities, so a zero phase lag is also common. When the phase lag is  $0^\circ$ , the vertical and horizontal oscillations become coherent, and their zero-crossings occur at about the same time. These observations are supported by empirical evidence, as shown in Fig. 8, where the horizontal and the vertical velocities of the word shown in Fig. 6 are plotted. Note that the phase lag is likely to be  $\pm 90^\circ$  or  $0^\circ$ . This phenomenon supports our

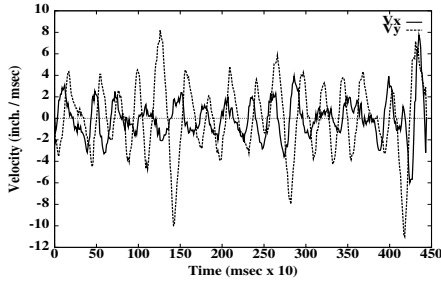


Fig. 8. The horizontal and the vertical velocities of the word shown in Fig. 4 (after removing the slant)

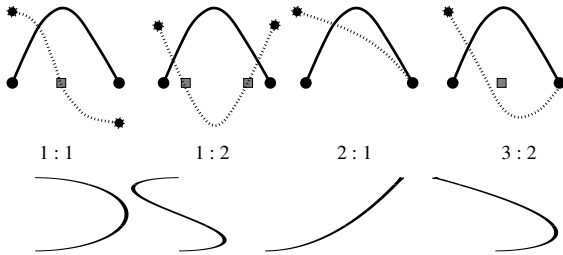


Fig. 9. The possible phase-lag relations and the corresponding spatial curves

discrete dynamical approach, and the phase lag between the oscillations is discretized to  $\pm 90^\circ$  or  $0^\circ$ . We now describe how the best discrete phase-lag trajectory is found.

Examining the cycloidal model for each Roman curvise letter reveals that the horizontal to vertical angular velocity ratio is at most 2, i.e.,  $\max \left\{ \frac{\omega_x}{\omega_y}, \frac{\omega_y}{\omega_x} \right\} \leq 2$ . Thus, for English cursive handwriting the ratio  $\frac{\omega_x}{\omega_y}$  is restricted to the range  $[\frac{1}{2}, 2]$ . Combining the angular velocity ratio limitations with the discrete set of possible phase-lags implies that the possible angular velocity ratios are: 1:1, 1:2, 2:1, 2:3, and 3:2. Four of these cases are plotted in Fig. 9 with the corresponding spatial curves, assuming that the horizontal drift is zero. The vertical velocity  $V_y$  is plotted with a solid line and the horizontal velocity  $V_x$  with a dotted line.

We view the vertical velocity  $V_y$  as a ‘master clock’, where the zero-crossings are the clock onset times.  $V_x$  is viewed as a ‘slave clock’ whose pace varies around the ‘master clock’. The rate ratio between the clocks is limited to at most 2. Thus,  $V_y$  induces a grid for  $V_x$  zero-crossings. The grid is composed of  $V_y$  zero-crossings and multiples of quarters of the zero-crossings (the bold circles and the grey rectangles in Fig. 10).  $V_x$  zero-crossings occur on a subset of the grid. The phase trajectory is defined over a subset of this grid, which is consistent with the discrete phase constraints and the angular velocities ratio limit. The allowed transitions for one grid point are plotted by dashed lines in Fig. 10. For each two allowed grid points the phase trajectory is calculated. For example, if  $t_i$  and  $t_j$  are two grid points and there is a  $V_y$  zero-crossing at  $t_k$  where  $t_i < t_k < t_j$ , then the horizontal velocity phase along the time interval  $[t_i, t_j]$  should meet the following conditions:

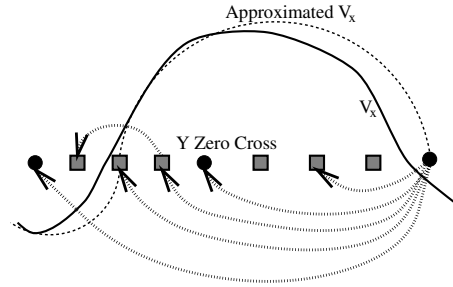


Fig. 10. Phase lag trajectory optimization by *dynamic programming*.  $V_x$  is approximated by limiting its zero-crossings to a grid which is denoted in the figure by bold circles ( $V_y$  zero-crossings) and grey rectangles

$\theta_x(t_i) = 2\pi n$ ,  $\theta_x(t_k) = 2\pi(n + \frac{1}{4})$ ,  $\theta_x(t_j) = 2\pi(n + \frac{1}{2})$ . The phase trajectory is linearly interpolated between the induced grid points. Hence, the phase along the time interval  $[t_i, t_j]$  is

$$\theta_x(t) = \begin{cases} 2\pi n + \frac{\pi}{2} \frac{t-t_i}{t_k-t_i} & t_i \leq t < t_k \\ 2\pi n + \frac{\pi}{2} + \frac{\pi}{2} \frac{t-t_k}{t_j-t_k} & t_k \leq t < t_j \end{cases}$$

If there is no  $V_y$  zero-crossing between the grid points or there are two  $V_y$  zero-crossings, the  $V_x$  phase lag changes linearly between the zero-crossings. In those cases, the phase trajectory along the grid points is

$$\theta_x(t) = 2\pi n + \pi \frac{t-t_i}{t_j-t_i}$$

Given the horizontal phase lag and assuming that the amplitude modulation is constant along one grid interval, the amplitudes that will preserve the horizontal progression are calculated. Denoting by  $L$  the horizontal progression, the approximated horizontal amplitude modulation along the time interval  $[t_i, t_j]$  is

$$A'_{i,j} = \frac{L}{\int_{t_i}^{t_j} \sin(\theta_x(t)) dt}$$

and the approximation error along this interval is

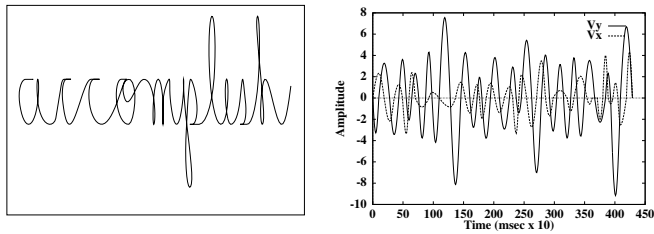
$$Error_{Approx}([t_i, t_j]) = \int_{t_i}^{t_j} (V(t) - A'_{i,j} \sin(\theta_x(t)))^2 dt$$

Formally, let the set of possible grid points be  $\bar{T} = \{t_1, t_2, \dots, t_N\}$ . We are looking for a subset  $\tilde{T} = \{t_{i_1}, t_{i_2}, \dots, t_{i_K}\} \subset \bar{T}$  such that all the pairs  $t_{i_j}, t_{i_{j+1}}$  are allowed, with the minimal induced approximation error

$$\tilde{T} = \arg \min_{T' \subset T} \sum_{i_j \in T'} Error_{Approx}([t_{i_j}, t_{i_{j+1}}])$$

For each grid point  $t_i$  a set of allowed previous grid points  $S_{t_i}$  is defined. The accumulated error at the grid point  $t_i$  can be calculated by *dynamic programming* using the following local minimization,

$$Error(t_j) = \min_{t_i \in S_{t_j}} \{Error(t_i) + Error_{Approx}([t_i, t_j])\}$$



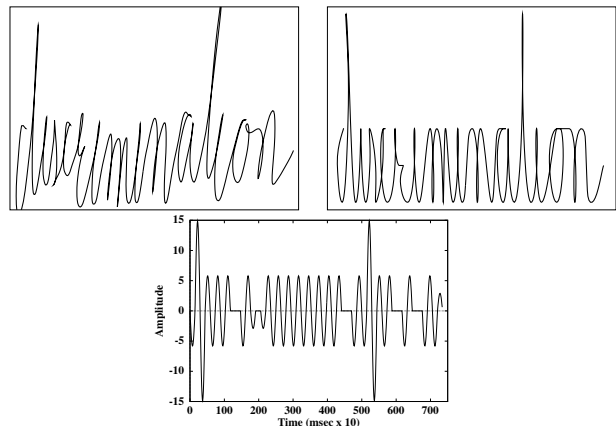
**Fig. 11.** The horizontal and vertical velocities and the reconstructed handwriting after phase-lag regularization

An illustration of the optimization process is depicted in Fig. 10. The best phase trajectory is found by backtracking from the best grid point of the last  $V_x$  zero-crossing. The result of this process is plotted in Fig. 11. This process ‘ties’ the two oscillations and represents the horizontal oscillations in terms of the vertical oscillations. Therefore, only the vertical velocity zero-crossings have to be located in the estimation process. This further reduces the number of bits needed to code the handwriting trajectory as indicated by stage 5 in Fig. 13. Since the horizontal oscillations are less stable and more noisy, this scheme avoids many of the problems encountered when estimating the horizontal parameters directly.

## 8 Angular velocity regularization

Until now the original angular velocities of the vertical oscillations were preserved. Hence, in order to reconstruct the velocities, the exact timing of the zero-crossings is kept. Our experiments reveal that all writers have their own typical angular velocity for the oscillations. These findings seem to be in contradiction to previous experiments, where it has been shown that a tendency exists for spatial characteristics to be more invariant than the temporal characteristics (Teulings et al. 1986, Thomassen and 1986) and to Hollerbach’s claim that both the amplitudes and the angular velocity are scaled during the writing of tall letters like **l**. For the purpose of representing handwriting as the output of a discrete, controlled, oscillatory system, fixing the angular velocity does not incur difficulties, and the approximated velocities preserve the context as shown in Fig. 12. Fixing the angular velocity can also be seen as a basic writing rhythm which may actually be supported by neurobiological evidence (Bergman et al. 1990). Since the horizontal oscillations are derived from the vertical oscillations by changing the phase lag, fixing the vertical angular velocity implies that the angular velocities of both the vertical and the horizontal oscillations are fixed. The angular velocity variations for each writer are small except in short intervals, where the writer hesitates or stops. The total halt intervals can be omitted or used for natural segmentation. The angular velocity is fixed to its typical value, and the time between two consecutive zero-crossings becomes constant.

The amplitudes are modified so that the total vertical and horizontal progressions are preserved. Since the



**Fig. 12.** The original and the reconstructed handwriting after angular velocity regularization (*top figures*) and the final vertical velocity (*bottom figure*)

horizontal and vertical progressions are quantized discrete values, the time scaling implies that the possible amplitudes are discrete as well. The time scaling can be viewed as a change in parameterization of the oscillation equations from time to phase, as denoted by (5). Assuming that the angular velocity,  $\omega$ , is almost constant implies that  $\frac{dt}{d\theta}$  is almost constant as well. The normalized dynamic equations which describe the handwriting become

$$\begin{cases} V_x(\theta) = A_x(\theta) \sin(\theta + \phi(\theta)) + 1 \\ V_y(\theta) = A_y(\theta) \sin(\theta) \end{cases} \quad (9)$$

where  $\phi(\theta) \in \{-90^\circ, 0^\circ, 90^\circ\}$ ,  $A_x(\theta) \in \{A_x^1, A_x^2, A_x^3\}$ , and  $A_y(\theta) \in \{A_y^1, A_y^2, A_y^3, A_y^4, A_y^5\}$ . The result of this process is shown in Fig. 12 where the original script and reconstructed script (after all stages including angular regularization) are plotted together with the synchronized vertical velocity. Note that the vertical velocity attains only a few discrete values at the maximal points of the oscillations. The number of bits needed to encode the writing curves is reduced after this final stage by a factor of about *100* compared with the original encoding of the writing curves (stage 6 in Fig. 13).

The synthesized velocities are not ‘natural’ due to the switching scheme of the velocity parameters, which results in very large accelerations at the zero-crossings. Our simple synthesis scheme was used in order to verify our assumption that cursive handwriting can be represented as the output of a discrete, controlled system. Other synthesis schemes can be applied to yield more ‘natural’ velocities. For example the principle of minimal jerk by Hogan and Flash (1987) can be used for synthesis.

## 9 The discrete control representation

So far, we have introduced a dynamic model which describes the velocities of a cursive writing process as a

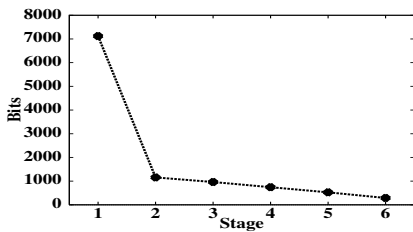


Fig. 13. The number of bits needed to encode cursive handwriting along the various stages

constrained modulation of underlying oscillatory processes. The imposed limitations on the dynamical control parameters result in a good approximation which is similar to the original. We then introduced a series of transformations which led to synchronous oscillations. As a result, a many-to-one mapping from the continuous velocities  $V_x(t), V_y(t)$  to a discrete symbol set was generated. This set is composed of a cartesian product of the discrete vertical and horizontal amplitude modulation values and the phase-lag orientation between the horizontal and vertical velocities.

Tracking the number of bits that are needed to encode the velocities (Fig. 13) reveals that the discretization and regularization processes gradually reduce the bit rate. This indicates that our discrete controlled system representation is well suited for compression and recognition applications. The transformation closes part of the gap between different writing styles and different writers. Keeping track of the transformations themselves can be used for writer identification. Here we introduce one possible discrete representation of the resulting discrete control. Our representation does not correspond directly to the original dynamic parameters but rather involves a one-to-one transformation of them.

Further, we describe the two discrete control processes as the output of two synchronized stochastic automata. The output can be written in two rows. The first row describes the appropriate vertical level (which can be one of 5 values) each time  $V_y(t) = 0$ . Whenever there is a vertical velocity zero-crossing, the corresponding automaton outputs a symbol which is the index of the level obtained at the zero-crossing. Similarly, the second automaton outputs a symbol when a horizontal velocity zero-crossing occurs. This symbol corresponds to the horizontal amplitude modulation for the next interval. Special care is taken when tracking the discrete control of the horizontal oscillations, since the phase is not explicit but changes its state implicitly. Yet if the initial horizontal oscillation phase is known, then the total phase trajectory can be reconstructed from this information. The first output symbol of the horizontal automaton is the initial phase denoted by  $\pm$ . Since the oscillation processes are synchronized by the angular velocity regularization, we only need to record the order of the automata output. When the two automata output symbols at the same time, it means that the oscillation phases became coherent; otherwise, there is a  $\pm 90^\circ$  phase lag. The angular velocity ratio limitation implies that each of the automata can output at most two con-

secutive symbols, while the other automaton is silent. The following is an example for such a representation for the same word ('toccata') written twice.

```

└25└2└4└44└2└4└3└4└└442└44└42└44└42└4└2└5└2└44└42└4└2
+22└1└1└23└2└3└3└1└1└23└3└1└133└3└2└2└2└2└2└3└2└2└2└3└

```

```

└25└2└4└442└4└3└44└4└2└4└442└4└442└4└2└5└2└4└442└4└2
+122└1└132└3└3└1└13└3└1└1233└1└132└2└3└2└3└1└232└2└3└

```

Note that the discrete representation is almost the same, and that simple rules may be found to match the two sequences. In fact, in this example, if we omit the horizontal (lower) output and squeeze the gaps for the vertical (upper) one, then the upper sequences for the two words are identical. This implies that much of the information is embedded in the vertical oscillations. We use this observation in a rudimentary spotting and matching experiment to conclude this paper.

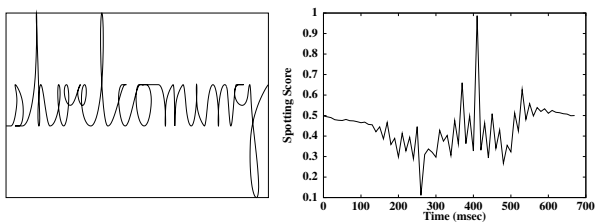
## 10 Word spotting and matching experiments

In order to test our representation, some rudimentary experiments were performed. We describe two of them in this section. Only a small amount of data was collected, so the experiments are restricted to nonparametric schemes. The first scheme is based on *learning vector quantization* (LVQ) which is essentially a static template matching. The second scheme is based on dynamic programming and called *dynamic time warping* (DTW). The second scheme was compared with direct DTW on the original handwriting velocities to check the robustness of our discrete representation.

### 10.1 Word spotting by learning vector quantization

A set of 10 words, all beginning with the letter sequence **com** was collected. Another set of 10 words with other letters without the sequence **com** but with the sequences **co** and **om** was also collected. A third set which included the sequence **com** and other sequences was used as a test set. A nonparametric method based on LVQ (Kohonen 1989) was developed for word spotting.

During the estimation process, there were false detections and misses of zero-crossings. This affects directly the output of the automata. Deletion and insertion of symbols may occur. In order to overcome this difficulty, the spotting algorithm should be made locally *shift tolerant*. Shift tolerance is achieved by training the classifier on 'windowed' parts of the sequences (McDermott and Katagiri 1991). Denote a sequence by  $\{S_i\}_{i=1}^N$  where  $N$  is the length of the sequence. Each sequence is divided into  $L$  overlapping sequences,  $S^l = \{S_i\}_{i=l}^{N-L+l}$ . A two-class LVQ was trained for each chopped sequence, by building a vector set (also termed codebook) for each position  $i \in \{1, \dots, N - L + 1\}$ . Denote by  $M$  the number of



**Fig. 14.** A reconstructed handwriting and the corresponding spotting score of the word-part `com`

code vectors for each position, then the total number of code vectors is  $ML$ .

A compound distance was developed in order to compare different sequences. First, the outputs of the vertical automaton are compared by calculating the Euclidean distance between the chopped sequences. Then the vertical automaton outputs are used to divide the horizontal automaton output into blocks. The horizontal output symbols between two consecutive vertical automaton outputs are gathered into one block. Denote the discrete output times of the vertical automaton by  $i_1, i_2, \dots, i_K$  and the horizontal output symbols by  $X_j$ . The  $k$ th horizontal block is  $B_k = \{X_j \mid i_k \leq j < i_{k+1}\}$ . The compound distance (denoted by  $D_c$  for each class  $c$ ) is the sum of the vertical symbol distance and the horizontal block distance

$$\begin{aligned} D_c(\{S_i^1\}, \{S_i^2\}) &= D(\{Y_i^1\}, \{Y_i^2\}) + D(\{X_i^1\}, \{X_i^2\}) \\ &= \sum_{i=1}^{N-L+1} \|Y_i^1 - Y_i^2\|_2 + \sum_{i=1}^K \|B_i^1 - B_i^2\|_2 \end{aligned}$$

Spotting was tested on more than 30 words. The closest code vector for each class within each position was found for all the chopped sequences. Let  $D_c(l)$  be the distance for class  $c$  ( $c = 1, 2$ ) at the  $l$ th position. An activation is defined for each window position using these distances as follows

$$A_c(I) = \frac{1}{L} \sum_{l=1}^L \frac{D_{1-c}(l)}{\sum_{c=1}^2 D_c(l)} = \frac{1}{L} \sum_{l=1}^L \frac{D_{1-c}(l)}{D_0(l) + D_1(l)}$$

For words that contain the word-part `com`, there is a window position with close vectors from the code; hence,  $D_0(l)$  is small compared with  $D_1(l)$ , and the activation is high. For random strings, the distances are about the same, hence the score fluctuates about 0.5.

The result of spotting in the word `shortcoming` is shown in Fig. 14, where the reconstructed word is plotted together with the corresponding spotting score. The sharp peak occurs in the last part of the learned word-part `com` since the accumulated activation scores are calculated backward.

## 10.2 Word matching by dynamic time warping

DTW (Sankoff and Kruskal 1983) is a well-known technique for string matching, when the strings are corrupted by a noisy channel that can omit, insert, or substitute

symbols. If the strings are over a finite alphabet, then the distance between two strings is defined as the minimal number of string modifications (deletions, insertions, or substitutions) on one string that make it equal to the other. This is a symmetric distance since a deletion of a symbol in the first string is equivalent to an insertion of the symbol in the proper place of the second string. The optimal sequence of operations can be obtained via dynamic programming. Substitutions can be performed by successive insertions and deletions, hence the distance  $D$  between strings  $S_1, S_2$  up to place  $i_1$  in string  $S_1$  and place  $i_2$  in string  $S_2$  is,

$$D_{S_1, S_2}(i_1, i_2) = \min \begin{cases} D_{S_1, S_2}(i_1 - 1, i_2) + C_{\text{del}} \\ D_{S_1, S_2}(i_1, i_2 - 1) + C_{\text{ins}} \end{cases} \quad (10)$$

where  $C_{\text{del}}$  and  $C_{\text{ins}}$  are the costs of a deletion and an insertion. The total distance between two strings is the accumulated distance at the last two symbols in the strings,  $D(S_1, S_2) \triangleq D_{S_1, S_2}(|S_1|, |S_2|)$ .

When the strings' alphabet is an infinite set a slightly different scheme is applied. Usually the square distance, denoted by  $\|\cdot\|_2$ , is used to measure the distance between symbols. The distance between two strings is

$$\begin{aligned} D_{S_1, S_2}(i_1, i_2) &= \\ \min \begin{cases} D_{S_1, S_2}(i_1 - 1, i_2) + \|S_1(i_1), S_2(i_2)\|_2 \\ D_{S_1, S_2}(i_1, i_2 - 1) + \|S_1(i_1), S_2(i_2)\|_2 \\ D_{S_1, S_2}(i_1 - 1, i_2 - 1) + \|S_1(i_1), S_2(i_2)\|_2 \end{cases} \end{aligned} \quad (11)$$

Based on the distance  $D_{S_1, S_2}$  a *forward distance matrix*  $M^f$  is defined, where  $M_{i,j}^f = D_{S_1, S_2}(i, j)$ . Therefore, the  $(i, j)$  element in the matrix represents the distance between the two strings from the start up to place  $i$  in  $S_1$  and  $j$  in  $S_2$ . Let  $S'_1 = \epsilon \circ S_1, S'_2 = \epsilon \circ S_2$  be two pseudo-strings, where  $\epsilon$  represents the null symbol and  $\circ$  represents concatenation. We define the matrix  $M^f$  as the distance matrix between  $S'_1$  and  $S'_2$ . Denote the length of the sequences by  $N_1 = |S'_1|$  and  $N_2 = |S'_2|$  and assume that  $N_1 > N_2$ . We want to spot occurrences of the string  $S_2$  in  $S_1$ . The following scheme builds the distance matrix: Initialize  $M_{i,0}^f = 0 : 0 \leq i \leq N_1$ , i.e. the string  $S_2$  can start anywhere in  $S_1$ . The rest of the matrix is computed by the recursion (11). The values  $M_{i,N_2}^f$  are the distances (number of string operations needed) when the string  $S_2$  ends in location  $i$  in  $S_1$ . The smaller  $M_{i,N_2}^f$  is, the more likely that the string  $S_2$  ends in location  $i$  in  $S_1$ . We can define a backward distance similarly

$$D_{S_1, S_2}(i_1, i_2) = \min \begin{cases} D_{S_1, S_2}(i_1 + 1, i_2) + C_{\text{del}} \\ D_{S_1, S_2}(i_1, i_2 + 1) + C_{\text{ins}} \end{cases} \quad (12)$$

The *backward distance matrix*  $M^b$  is defined based on this distance. The symmetric distance between strings  $S_1$  and  $S_2$  for a position centered around the  $i$ th symbol in  $S_1$  is  $SD_{S_1, S_2}(i) = M_{i - \frac{N_2}{2}, 0}^b + M_{i + \frac{N_2}{2}, N_2}^f$ . This distance measures the number of insertion and deletion operations

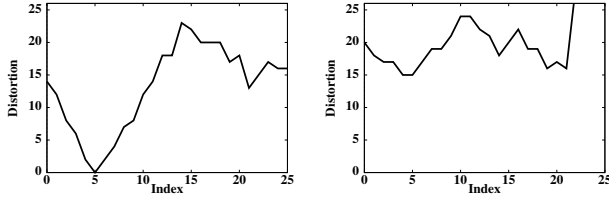


Fig. 15. The DTW spotting distance for the word-part `com` in the words `locomotion` (left) and `sybiotic` (right)

needed to transform part of string  $S_1$ , centered around  $i$ , to  $S_2$ .

Based on the distances  $SD(i)$ , spotting occurrences of the sequence representing the word-part `com` is performed. The results of this method are similar to the results of the *LVQ*-based method. Since the *DTW* process can accommodate for local displacements by extra insertions or deletions at the ends of the string, the match is also good (small distance) around the best spotted place. A typical result is shown in Fig. 15 for the words `locomotion` and `sybiotic`. Note that the best distance for the word `locomotion` was 0, i.e. a perfect match was found !

Another experiment using *DTW* was performed on complete words. It tested the discrimination ability of the discrete representation by performing word matching based on discrete *DTW* versus a continuous domain *DTW*. A continuous domain *DTW* was performed by applying the distance defined in (11) on the original horizontal and vertical velocities. Five different words were collected, each written twice (by a single author): `bifocal dignify horizon quantum wolfish`. The estimation process was applied to each of the words. The horizontal and vertical velocities were kept after finding the best affine transformation between each pair of words. The distance was calculated by initializing  $M_{0,0}^f = 0$  and  $M_{i,j}^f = \infty$  ( $1 \leq i \leq N_1, 1 \leq j \leq N_2$ ) and defining the distance for complete strings as  $D_{S_1, S_2}(|S_1|, |S_2|)$ . The results of the discrete and continuous *DTW* are shown in Fig. 16. There, the distance between two words  $S_i, S_j$  is denoted by a rectangle. Larger rectangles indicate larger distances. The ‘ideal’ distance matrix should be all black except for a  $2 \times 2$  block diagonal. The distances for both experiments were normalized to have the same average. Clearly, the matching based on the discrete representation outperforms the continuous, indicating the robustness of our representation.

## 11 Conclusions and future research

Although the idea that the pen movements in the production of cursive script are the result of a simple ‘motor program’ is quite old, revealing this ‘motor code’ is a difficult inverse-dynamic problem. In this paper, we present a robust scheme which transforms the continuous pen movements into discrete *motor control symbols*. These symbols can be interpreted as a possible high

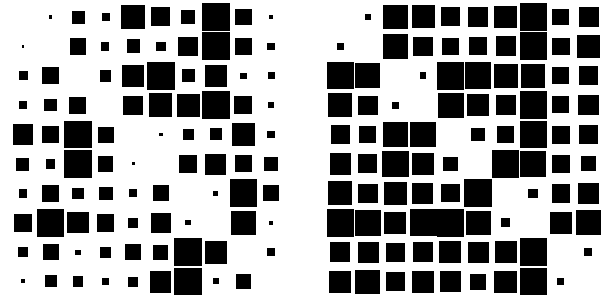


Fig. 16. Comparison of *DTW* on the discrete representation (right) and on the original writing velocities (left)

level coding of the motor system. The relationship between this representation and the actual cognitive representation of handwriting remains open, though there is some psychophysical experimental evidence linking the recognition time to the writing time for handwriting (Frederiksen and Kroll 1976). The *discrete motor control* representation largely reduces the variability in different writing styles and writer specific effects. The rudimentary recognition experiments that we performed indicate the potential of this representation for cursive recognition tasks, which is our primary goal. Since different writing styles are transformed to the same representation, the transformation itself can be used for *text independent* writer identification and verification tasks.

## 12 Appendix

We assume that there is a virtual center for the vertical movements and that the amplitudes are symmetric about this center. The problem becomes similar to a mixture density estimation, but it is more involved since the parameters are tied via the symmetry constraints. The five levels correspond to five normal distributions with unknown means and a common variance. Initially, each level is chosen by the a priori probability  $P_i$ . We need to estimate the parameters  $H_i$  and find the most probable level indices  $I_t$ , when the available observations are the noisy vertical positions at the zero-crossings.

Let  $\mu_i \triangleq H_i$  and denote the stochastic levels by  $Y_i \sim N(\mu_i, \sigma)$  ( $i \in \{1, \dots, 5\}$ ). At each of the zero-crossings one of the levels is chosen with probability  $P_i$  ( $\sum_{i=1}^5 P_i = 1$ ). The observed information is a noisy sample of the chosen level. We would like to estimate concurrently the vertical amplitude parameters and the levels obtained at the zero-crossings. Denote the parameter set by  $\Theta = \{\{\theta_i\}, \sigma\} = \{\{P_i\}, \{\mu_i\}, \sigma\}$ . The joint distribution of the levels  $Y$  is  $Z \sim \sum_{i=1}^5 P_i N(\mu_i; \sigma)$ . The symmetry constraints imply that  $\mu_5 = 2\mu_3 - \mu_1$  and  $\mu_4 = 2\mu_3 - \mu_2$ . The complete data are denoted by  $(\bar{Y}, \bar{I}) = (\{Y_t\}, \{I_t\})$  where  $I_t$  is the index of the chosen level at time  $t$ , and  $Y_t$  is the observed level value at that time. Let  $I_t(i)$  be the levels indicator vector due to the index  $I_t$ , i.e.,  $I_t(i) = 1$  if  $I_t = i$  and  $I_t(i) = 0$  otherwise. The likelihood of an observation sequence  $\{Y_t\}_{t=1}^T$  is

$$\begin{aligned} \log L_{\Theta}(\bar{Y}) &= \log \sum_{t=1}^T P_{I_t} N(Y_t; \mu_{I_t}, \sigma) \\ &= \sum_{t=1}^T \sum_{i=1}^5 I_t(i) \log P_i N(Y_t; \mu_i, \sigma) \end{aligned} \quad (13)$$

The first step in each *EM* iteration is to find the expectation of (13) using the current estimation of the parameter set denoted by  $\Theta_1 = \{ \{P_i^1\}, \{\mu_i^1\}, \sigma^1 \}$ . The following weights are calculated using the current parameters

$$W_t(i) \triangleq E(I_t(i) | Y_t, \Theta_1) = \frac{P_i^1 e^{-\frac{1}{2} \left( \frac{Y_t - \mu_i^1}{\sigma^1} \right)^2}}{\sum_{i=1}^5 P_i^1 e^{-\frac{1}{2} \left( \frac{Y_t - \mu_i^1}{\sigma^1} \right)^2}} \quad (14)$$

The second stage of each *EM* iteration maximizes the current set of parameters, denoted by  $Q(\Theta; \Theta_1)$ , using the expectation of (13)

$$\max_{\Theta} Q(\Theta; \Theta_1) = \max_{P_i, \mu_i, \sigma} \sum_t \sum_i W_t(i) \cdot \left( \log P_i - \log \sigma - \frac{1}{2} \left( \frac{Y_t - \mu_i}{\sigma} \right)^2 \right) + \text{Const} \quad (15)$$

Taking the partial derivative of (15) with respect to  $P_i$  under the constraint that  $\sum_{i=1}^5 P_i = 1$  and equating it to zero results in the following estimator,  $P_i = \frac{\sum_t W_t(i)}{\sum_i \sum_t W_t(i)}$ . The estimation of the current optimal level averages  $\mu_i$  is more complicated due to the symmetry constraints. We rewrite Equ. (15) by substituting the symmetry constraints. Therefore, the explicit form for  $Q$  is

$$Q(\Theta; \Theta_1) = \text{Const} + \sum_t \sum_{i=1}^5 W_t(i) (\log P_i - \log \sigma) - \sum_t \sum_{i=1}^3 \frac{1}{2} W_t(i) \left( \frac{Y_t - \mu_i}{\sigma} \right)^2 - \sum_t \sum_{i=1}^2 \frac{1}{2} W_t(6-i) \left( \frac{Y_t - (2\mu_2 - \mu_i)}{\sigma} \right)^2 \quad (16)$$

Define  $\omega_i \triangleq \sum_t W_t(i)$  and  $\chi_i \triangleq \sum_t W_t(i) Y_t$ . Minimizing (16) with respect to  $\mu_0, \mu_1, \mu_2$  yields the following set of linear equations

$$\begin{cases} \mu_0 \omega_0 - \chi_0 - (2\mu_2 - \mu_0) \omega_4 + \chi_4 = 0 \\ \mu_1 \omega_1 - \chi_1 - (2\mu_2 - \mu_1) \omega_3 + \chi_3 = 0 \\ \mu_2 \omega_2 - \chi_2 + 2(2\mu_2 - \mu_0) \omega_4 - 2\chi_4 + 2(2\mu_2 - \mu_1) \omega_3 - 2\chi_3 = 0 \end{cases}$$

These equations are explicitly solved using the symmetry constraints, to obtain the new values for  $\mu_i$  as follows

$$\begin{aligned} D &\triangleq 4\omega_4 \omega_0 \omega_1 + 4\omega_4 \omega_3 \omega_1 + 4\omega_3 \omega_0 \omega_1 + 4\omega_4 \omega_0 \omega_3 + \\ &\quad \omega_2 \omega_0 \omega_1 + \omega_4 \omega_2 \omega_1 + \omega_4 \omega_2 \omega_3 + \omega_2 \omega_0 \omega_3 \\ \mu_0 &= D^{-1} (4\omega_4 \omega_3 \chi_1 + 2\omega_4 \omega_3 \chi_2 + 2\omega_4 \omega_1 \chi_2 - 4\omega_1 \omega_3 \chi_4 - \\ &\quad \chi_4 \omega_2 \omega_1 - \omega_3 \omega_2 \chi_4 + 4\omega_1 \omega_4 \chi_3 + 4\chi_0 \omega_4 \omega_1 + \\ &\quad 4\omega_3 \chi_0 \omega_1 + 4\omega_4 \omega_3 \chi_0 + \chi_0 \omega_2 \omega_1 + \omega_3 \omega_2 \chi_0) \\ \mu_1 &= D^{-1} (2\omega_4 \omega_3 \chi_2 + 4\omega_4 \omega_3 \chi_1 + 4\omega_4 \omega_0 \chi_1 - \omega_4 \omega_2 \chi_3 + \\ &\quad 4\omega_0 \omega_3 \chi_4 - \omega_2 \omega_0 \chi_3 + 4\omega_4 \omega_3 \chi_0 + \omega_4 \omega_2 \chi_1 + \\ &\quad \omega_2 \omega_0 \chi_1 + 4\omega_3 \omega_0 \chi_1 - 4\omega_4 \omega_0 \chi_3 + 2\omega_3 \omega_0 \chi_2) \\ \mu_2 &= D^{-1} (\omega_4 \omega_3 \chi_2 + 2\omega_4 \omega_3 \chi_0 + 2\omega_4 \omega_3 \chi_1 + \omega_3 \omega_0 \chi_2 + \\ &\quad 2\omega_3 \omega_0 \chi_1 + 2\omega_0 \omega_3 \chi_4 + 2\omega_1 \omega_0 \chi_3 + \omega_4 \omega_1 \chi_2 + \\ &\quad 2\chi_0 \omega_4 \omega_1 + \omega_1 \omega_0 \chi_2 + 2\omega_1 \omega_4 \chi_3 + 2\omega_1 \omega_0 \chi_4) \\ \mu_3 &= 2\mu_2 - \mu_1 ; \mu_4 = 2\mu_2 - \mu_0 \end{aligned}$$

Finally, the new variance is estimated using the new means,  $\sigma^2 = \frac{\sum_{t,i} W_t(i) (Y_t - \mu_i)^2}{\sum_{t,i} W_t(i)}$ . This process is iterated until convergence, which normally occurs within a few iterations. The final weights  $W_t(i)$  correspond to the posterior probability that at time  $t$  the pen was at the vertical position  $H_i$ . Choosing the maximal value as the indicator of the level is the *maximum a posteriori* decision. This process can be performed on-line on a word basis or

off-line for several words. In the latter case, the estimated a priori probabilities  $P_i$  reflect the stationary probability to be at position  $H_i$ . These probabilities are influenced by the motor characteristics of the handwriting as well as by the linguistic characteristics.

*Acknowledgements.* We would like to thank Scott Kirkpatrick and Andrew Senior for reading the manuscript and for their valuable comments. This work was partly supported by a grant from ART Ltd. and by a grant from the Israeli Ministry of Science and Arts. Y.S. would like to thank the Clore Foundation for its support.

## References

- Bergman T, Wichmann T, DeLong MR (1990) Reversal of experimental parkinsonism by lesions of the subthalamic nucleus. *Science* 249:1436-1438
- Dempster A, Laird N, Rubin D (1977) Maximum likelihood estimation from incomplete data via the EM algorithm. *J R Stat Soc* 39B:1-38
- Edelman S, Flash T, Ullman S (1991) Reading cursive handwriting by alignment of letter prototypes. *Int J Comput Vision* 5(3):303-331
- Frederiksen JR, Kroll JF (1976) Spelling and sound: approaches to the internal lexicon. *J Exp Psych Hum Percept Perform* 2(3):361-379
- Hogan N, Flash T (1987) Moving gracefully: quantitative theories of motor coordination. *Trends Neurosci* 10(4):170-174.
- Hollerbach JM (1981) An oscillation theory of handwriting. *Biol Cybern* 39:139-156
- Kohonen T (1989) *Self-organization and associative memory*. Springer, Berlin Heidelberg New-York
- Lacquit F (1989) Central representations of human limb movement as revealed by studies of drawing and handwriting. *Trends Neurosci* 12(8):287-291
- McDermott E, Katagiri S (1991) LVQ based shift-tolerant phoneme recognition. *IEEE Trans ASSP* 39(6):1398-1411
- Oppenheim AV, Schaffer RW (1975) *Digital signal processing*. Prentice-Hall, Englewood Cliffs, NJ
- Plamondon R, Leedham CG (1990) *Computer processing of handwriting*. World Scientific, Singapore
- Plamondon R, Suen CY, Simner ML (1989) *Computer recognition and human production of handwriting*. World Scientific, Singapore
- Rumelhart DE (1992) *Theory to practice: a case study - recognizing cursive handwriting*. Proc. of 1992 NEC Conf. on Computation and Cognition
- Sankoff D, Kruskal JB (1983) *Time warps, string edits and macro-molecules: the theory and practice of sequence comparison*. Addison-Wesley, Reading Mass.
- Tappert CC, Suen CY, Wakahara T (1990) The state of art in on-line handwriting recognition. *IEEE Trans Pattern Anal Mac Intell* 12(8):787-808.
- Teulings HL, Thomassen AJWM, Galen GP van (1986) *Invariants in handwriting: the information contained in a motor program*. In: Kao HSR, van Galen GP, Hoosain R (eds) *Graphonomics: contemporary research in handwriting*
- Thomassen AJWM, Teulings HL (1986) *Time size and shape in handwriting: exploring spatio-temporal relationships at different levels*. In: Michon JA, Jackson JL (eds) *Time, mind, and behavior*. pp 252-263
- Wald A (1940) Fitting of straight lines if both variables are subject to error. *Ann Math Stat* 11:284-300