Invariant Handwriting Features Useful in Cursive-Script Recognition

Hans-Leo Teulings

Nijmegen Institute for Cognition and Information (NICI), P.O. Box 9104, 6500 HE Nijmegen, The Netherlands (*).

Abstract. The within-writer variability of handwriting forms one of the problems in the automatic recognition of cursive script. Variability can be handled by choosing handwriting features based upon the process of handwriting generation or upon computational models. Handwriting patterns are represented by a sequence of motor actions, i.e., "strokes", which can be identified by invariant segmentation. Each stroke is characterized by features related to motor memory parameters which can be identified by their high signal-to-noise ratios.

Keywords. Cursive-script recognition, on-line handwriting, movement production models, computational models, movement variability, segmentation, movement features, signal-to-noise ratio, off-line reconstruction from on-line.

1. Introduction

Automatic recognition of cursive handwriting is difficult because handwriting comes in great variety of graphical shapes. Not only do different writing styles exist, but, also within a writer, identical letters vary between instances. In order to make automatic recognition systems robust with respect to these variations it is not only instructive to study human perception and reading, but also the production process of handwriting. This paper provides a guide for finding and verifying handwriting features which are relatively insensitive to spontaneous variations of cursive-script. These features form the corner stone for non-chaotic handwriting recognition systems. A macroscopic handwriting production model will be discussed, and various computational handwriting models will be presented.

---

*) Presently at: Exercise Science and Physical Education, Arizona State University, Tempe, AZ 85287-0404, USA. E-mail: hlt@espe1.la.asu.edu. Fax: +1-602-965-8108.

The macroscopic model of handwriting production proposed, consists of a sequence of modules. The higher-level modules have been hypothesized on the basis of slips of the pen [1], and neurologic disturbances [2]. The lower-level modules have been hypothesized on the basis of delays in movement initiation and movement execution [3, 4]. The handwriting motor system contains in its most simplified form (1) a motor memory, where all letter shapes have been stored [5], (2) movement-unit retrieval, where the appropriate allographs are extracted from that motor memory, and (3) a motor-pattern buffer, the contents of which can be defined as the motor program. Subsequently, this motor program needs to be adapted to each concrete execution condition. The adaptation is done at two levels: (4) At the parameter-setting level, where muscle-independent global scale parameters (e.g., speed, size) are specified, and (5) at the muscular-initiation level, where muscle-dependent parameters (e.g., orientation, slant, limb) are specified [3].

A serial-module model is applicable if feedback loops between modules are not essential. In cursive script, several letters are prepared in advance during writing. Subsequent letters of a writing pattern are prepared while executing the first letters and processing feedback from previous letters [4]. In the multi-module model, feedback loops are included, but, play only a marginal role because handwriting is performed as fast as 200 - 400 ms per letter, which is too fast for processing feedback to monitor letter formation. For example, a sudden increase or decrease of pen-to-paper friction results in an immediate reduction of the letter size and several letters are produced before letter size is restored on the basis of visual or tactile feedback [6]. As biomechanical properties of the wrist and finger joints do not seem to be corrected [7], visual feedback is apparently not used to monitor the performance of straight lines. Only orientation and slant are controlled by visual feedback [8].

A cursive-script recognition system, should operate in the reverse way as the handwriting production system. Firstly, the cursive-script recognition system segments the continuous handwriting movement, into units of movement. Then, lower-level, global features, such as orientation, size, slant, and speed are normalized. Subsequently, features are estimated which are closely related to the movement parameters in motor memory. In order to check the completeness of the features extracted, a legible handwriting pattern should be reconstructible, but not necessarily an exact reproduction.

2. Segmentation of Cursive Script into Strokes

Movement parameters in motor memory are not global parameters but rather local parameters generating small segments of the handwriting pattern. Handwriting patterns need to be separated into discrete segments, each representing a single movement command corresponding to a ballistic "stroke". It consists of an acceleration phase, a velocity peak, and a deceleration phase. Before discussing the features, the invariant segmentation of handwriting will be addressed. Although Guiard [9] questioned whether continuous movement sequences can at all be segmented into discrete aiming movements, segmentation offers a great advantage in reducing the complexity of handwriting patterns. The ideal segmentation procedure should segment similar writing patterns in an invariant way, regardless of the writing speed. In Teulings et al. [10] a highly invariant segmentation procedure has been proposed. Other cursive-script segments have been suggested, e.g., segments of constant parameter settings [11]; partly overlapping segments of constant curvature [12]; pen-down segments [13]; segments of constant curvature versus speed relation [14]; overlapping speed and curvature profiles [15]; segments between absolute velocity minima that are robust under lowpass filtering [16]; segments between top, bottom, or stroke crossing [17]; segments at top and bottom [18-20]; and segments between two tops yielding a pair of ballistic strokes and between two bottoms yielding the overlapping stroke pairs [21-23].

Segmenting at speed minima does not always yield a desirable result [27]. Firstly, at low handwriting rates, a simple, straight stroke is not always ballistic so that several absolute velocity minima appear, suggesting several ballistic "substrokes". Secondly, perfect circles at constant velocities do not yield any segmentation points at all because of the perfectly out-of-phase movement components. In order to make the segmentation more invariant and more intuitively correct (i.e., corresponding to integer up or down strokes), four refinements of the algorithm are necessary [27]. (1) In the estimate of the absolute velocity the horizontal component squared is weighted by 0.1 relative to the vertical component, i.e., v = (0.1 * vx**2 + vy**2) **1/2, which shows low minima between vertical segments, and moderate minima between horizontal segments. (2) Only velocity minima are accepted, which do not have a lower minimum within a time window of -60 and +60 ms. This period corresponds to the time constants in the muscles and the nerves. (3) Segments, shorter than the minimum stroke size of about 0.5 mm, are considered non-significant and are joined with the previous segment. (4) Finally, a segmentation point is removed when the direction changes less then 15 degrees. The change of direction is defined as the angle between the vectors from the current segmentation point to the previous and the following points of speed maxima. This refined segmentation appears to be invariant when handwriting rate is changed across a wide range.

3. Frequency Spectra

When handwriting rate is increased, one may suggest that individual motor commands increasingly overlap [24], or increasingly become smoothed by the hypothetical low-pass filtering of the pen-hand system [25]. In order to check whether this may cause segmentation problems, the frequency spectrum of the handwriting movements was inspected. The frequency amplitude spectrum of the horizontal and the vertical velocity components forms a symmetrical peak around 5 Hz [26]. When the rate of handwriting is increased, it appears that all component frequencies of the spectrum increase proportionally [27]. There is no evidence that the pen-hand system acts as a fixed low-pass filter, nor that the strokes become increasingly overlapping with writing speed. Only topological distortions seem to occur which cause atypical letter shapes.

3.1. Theoretical Minimum Number of Features

A time function of limited frequency bandwidth W and duration T can be reconstructed by 2WT (isochronous) samples (sampling theorem; See [28] for an overview). Handwriting consists of two time functions, one for the horizontal and the other for the vertical component. If both functions are independent, 4WT samples are required to reconstruct a handwriting segment of duration T. The frequency spectrum of handwriting shows a gradual descent to noise level at about W = 10 Hz [26]. Durations of the fastest ballistic strokes are about T = 0.1 s. At least 4 samples, or 2 coordinate pairs, per ballistic stroke are required. Depending upon the speed of writing, the samples will be at different phases of the strokes, and therefore these 4 samples are highly variable and not appropriate as features. Nevertheless, the notion can be upheld that 4 features are sufficient to disambiguate a stroke. In order to compare the invariance of features, a measure for the invariance will be proposed.

4. Defining Invariant Stroke Features by the Signal-to-Noise Ratio (SNR)

When a handwriting pattern is segmented into a sequence of strokes and when each stroke yields a feature, the "pattern" of a stroke feature can be defined. If several replications of a handwriting pattern are available, the random variations of these sequences can be considered as the motor noise of a feature. The signal-to-noise ratio (SNR) is a dimensionless measure, which is useful comparing the extent of invariance of different stroke features. The SNR is the ratio of the standard deviations of the movement amplitude, and that of the movement noise:

$$SNR = (sd(Signal) / sd(Noise))$$

The standard deviations can be derived as follows. Assume Xij describes the pattern of a feature for all strokes i (I strokes) and replications j (J replications). Averages across replications j are denoted as Xi., etc. An analysis of variance schema for the main effect Strokes (I levels) based on J replications yields:

$$sd(Noise)^{**}2 = \Sigma[j=1,J]\ \Sigma[i=1,I]\ (Xij - Xi. - X.j + X..)^{**}2 / (I-1)(J-1)$$

and

$$sd(Signal)^{**}2 = \Sigma[i=1,I]\ (Xi. - X..)^{**}2 / (I-1) - sd(Noise)^{**}2 / J$$

According to Fitts' law, the expected magnitude of the signal-to-noise ratios can be estimated from the speed-accuracy tradeoff. Fitts' law states that the duration T of reciprocal aiming movements of amplitude A (or within a circle of width 2A) towards a target of width W increases linearly with the accuracy measure log2 (2A/W) in "bits" of information. The duration increase with accuracy is less then 50 ms/bit (26 ms/bit and 43 ms/bit for the finger and wrist movements, respectively [29]). This suggests that during short handwriting strokes of 100 ms, the movement amplitude, or stroke sawe, contains about 2 bits of information. Two bits of information implies an "effective quantization" into 4 levels, which corresponds to typical SNR values of 4 [10, 30]. It is possible to estimate the number of discriminable classes of letters under the assumptions that 4 features exist with SNR values of 4 and that the average number of strokes per letter is 4. Then, only 64 different letter shapes can be disambiguated. More than 64 different letter shapes may exist (e.g., 26 lower-case letters and their allographic variations). Some of the features may be dependent, and the SNR values of most features are less than 4. For these reasons it is imperative to employ the features with the highest SNRs and a high extent of independence.

5. Independence of Features

The abstract movement parameters stored in motor memory are the ideal features for the recognizer's database of cursive-script prototypes. One may suppose that these parameters are highly invariant under many different conditions, as they are used to generate all other movement parameters, required to perform the writing pattern in many different conditions. The reverse seems also reasonable. Invariant handwriting features are closely related to the memory representation of the motor program. Therefore, invariant features observed in replications of a handwriting pattern are useful in cursive-script recognition.

In order to evaluate the appropriateness of a set of stroke features, various hierarchical relations can be discerned. A top-down hierarchy can be discerned with high-level features, related to the movement information in memory, and lower-level features, derived from these higher-level features. Furthermore, a random variation occurring in one stroke may, to some extent, occur in neighboring strokes. This constitutes the sequence hierarchy.

5.1. Top-Down Hierarchy

The top-down hierarchy assumes that movement memory contains a parsimonious set of source features which generate the lower-order features. The lower-order features show more and more noise and therefore less and less invariance. Source features can be identified by their high invariance [10, 30]. The following is an example of hierarchical features, which are related through a mechanical rule. The stroke size S is proportional with peak acceleration A and duration T squared, i.e., S = Eff * A * T**2, where Eff refers to an efficiency constant, characterizing the effect of a particular shape of the acceleration-versus-time curve [30]. It appeared that the pattern of normalized stroke sizes S is more invariant than the pattern of stroke durations T, peak accelerations A, or efficiencies Eff [10, 30] (See Table 1). This suggests that the patterns of duration, peak acceleration, and efficiency are derived from the stroke size according to some rules. In handwriting recognition, normalized stroke size is a more useful feature than the dynamic features such as stroke duration and peak acceleration.

The "between-feature correlation" provides additional evidence in favor of a top-down hierarchy. One may assume that random variations of a higher-order feature are not compensated by variations of lower-order features. If features A, B, and C satisfy a mechanical relation such as A = B + C or A = B * C, and if B and C are independent, then their absolute or relative variances are additive, i.e.: var(A) = var(B) + var(C) or var(A) / mean(A)**2 = var(B) / mean(B)**2 + var(C) / mean(C)**2, respectively. If B and C are negatively correlated, then the (relative) variance of A will be significantly smaller then expected in the uncorrelated case, supporting the conclusion that feature A is a higher-order invariant feature, controlling both B and C. Note that the correlation between B and C can only be caused by noise in the motor system, as opposed to any movement law describing how B changes with varying C, because B and C are based on "identical" strokes in a series of "identical" writing patterns.

Under the assumption that the same memory representation is used to perform a writing pattern under various execution conditions, such as, different sizes, speeds, limbs, orientations, or slants there should be a high similarity between the higher-order patterns of features generated by the same memory representation. If the pattern of a feature changes proportionally or, at least, linearly from one condition to the other, then the "between-condition correlation" will be high. A feature needs to be rescaled or normalized if the transformation of a feature between conditions is proportional.

5.2. Sequence Hierarchy

Sequence hierarchy deals with the variations affecting neighboring strokes. The extreme case of sequence hierarchy is "rescalability", or the proportional change of a feature of all strokes. Depending upon the extent of the sequence hierarchy, two or more consecutive strokes have to be taken into account when representing prototypes in a cursive-script recognizer’s database. If the sequence hierarchy would be nonsignificant, only separate strokes would be sufficient to store all prototype letters. This may seem reasonable if one realizes that, within a writer’s handwriting, ascenders and descenders of different letters are similar. However, when motor-noise affects each stroke separately, atypical letters would result [22]. In fact, there is evidence that whole letters, counting several strokes, form units in the handwriting motor system [5], so that it is likely that sequential strokes are affected by similar noise sources.

Significantly negative correlations between successive stroke sizes both in x and in y directions were found [10] (See Table 1), indicating that when motor noise causes one stroke-size component to be bigger or smaller than normal, the subsequent stroke will be bigger or smaller in the opposite direction. Noteworthy is that stroke durations do not satisfy any sequence hierarchy, as the correlations between successive stroke durations appear not to be robust. The latter finding confirms that a writing pattern can be split into a strict sequence of movement units. Recognition systems dealing with stroke pairs seem to take the sequence hierarchy of stroke sizes into account to some extent. However, recognition rates of complete letters, counting up to 6 strokes, is not higher than recognition rates of letters on the basis of separate strokes [31] which may indicate that the sequence hierarchy is limited to a few strokes.

Table 1

SNRs and correlations among various stroke features, forming a mechanical relation: vertical strokes size (Sy), vertical peak acceleration (Ay), stroke duration (T), and efficiency (Effy). Results are based on 16 replications of the writing pattern "elementary" in the normal condition and compared to the fast condition in a subject.

	Features			
	Sy =	Effy *	Ay *	T**2
Sequence Hierarchy				
Correlation between Strokes	-0.45	-	-	-0.15
Top-Down Hierarchy				
SNR	5.2	2.5	2.0	3.6
Correlation between Features	+0.24	-0.46	-0.15	
		+0.59	-0.41	
		+0.41		
Correlation between Conditions	0.99	0.95	0.95	0.95

6. Context

Context is a specific case of mutual influencing of letters in handwriting. Yet, the context effects between cursive "e" and "l", are only marginal, especially for spatial features [32]. The major context effects due to size concern the duration of the neighboring strokes. In on-line cursive-script recognition, duration plays a role when characterizing strokes, unless resampling to constant velocity is done [33]. For that reason, it is useful to examine the relation between time and size. Three context levels have been discerned [34]. In each context, specific duration-versus-size relations exist:

o Macro context (i.e., a word in the context of other words of different sizes).

o Meso context (i.e., a single stroke in the context of other strokes of different sizes).

o Micro context (i.e., the local curve radius in the context of a single stroke with varying curvature).

In macro context, complete writing patterns are considered in isolation. As mentioned earlier, in the frictionless case, stroke size S is proportional to the peak acceleration A and the stroke duration T squared: S = Eff * A * T**2, where Eff is an efficiency factor characterizing the effect of a particular shape of the acceleration-versus-time curve. The efficiency Eff plays a minor role in controlling stroke sizes in handwriting [30], which is consistent with the finding that the shapes of the acceleration-versus-time curves are rather constant [35]. The time to produce a writing pattern is virtually independent of size in writing sizes between 0.25 and 1 cm [6, 36, 37]. Apparently, duration is limited by the frequency bandwidth of the pen-hand system so that size is controlled by acceleration level, which is proportional to force level, under the assumption of low friction. Consequently, duration T is constant in the range of writing sizes, so that peak acceleration A is proportional with stroke size S. Force levels will reach a ceiling when the writing sizes are much larger than 1 cm. The height of the ceiling depends on the instructed pace or time pressure [37]. Therefore, when producing large writing sizes or arm movements, force levels remain constant, and size variations are programmed entirely by variation of duration. In macro context, a power relation between T and S, can be proposed:

T = k * S**b (b = 0 if S < 1 cm; b = 1/2 if S >> 1 cm)

where k is constant per context and instruction.

Meso context refers to the adjacency of strokes of different sizes (such as in "el"). Most observations are compatible with an increase of both time and force. The following power function to describe the relation between duration T and size S in meso context is proposed [34]:

T = k * S**b (b = 0.33)

where k is a constant depending upon macro context. The data of "el" or "le" pairs in the literature [32, 34, 11, 38, 39, respectively] yield values for b = 0.22, 0.34, 0.34, 0.38, and 0.41, respectively.

Finally, micro-context describes the time needed per infinitesimal part of the writing trajectory. An estimator of "local size" is curve radius r of the circle, fitting the curve at that point. An estimator of "local duration" is the inverse angular velocity, i.e., r/v = 1/omega. In a large variety of drawing tasks the "2/3-power law" appears to hold [40]:

omega = r/v = k r**b   (b = 2/3)

where k is a constant gain factor which depends upon meso and macro context. The 2/3-power law can be derived for sinusoidal movements, i.e., two equal-frequency movement components without left-to-right translation, which produces only arbitrary ellipses. In normal handwriting the 2/3-power law holds only marginally [34]. A normal handwriting pattern yields b = 0.41 (instead of 2/3) and correlation 0.83 (instead of approximately 1.0) which are virtually the same values as those of a random-walk. Only continuous ellipses and "llll" patterns yield values which are closer to the ideal value, than to those of a random walk: b = 0.59 and correlations as high as 0.95. The 2/3-power law does not hold for handwriting in general due to the width of the frequency band of handwriting [26], and the left-to-right, which mainly takes place during the upstrokes [41, 42].

7. The choice of features

On-line handwriting can be recorded at a high accuracy [43] without the thinning and sequencing problem known in optical scanning. An appropriate set of features to characterize on-line handwriting can be derived from the movement parameters used in computational models of handwriting production. The set of features contains enough information to reconstruct writing patterns as a legible word. Presenting reconstructed words to subjects, yields indications on the importance of certain features. It is suggested that the downstrokes are more important for recognizing words than the upstrokes, which has been exploited in recognition systems [18].

Stroke features used in cursive-script recognition which show high SNRs [31, 44] are the horizontal and vertical stroke sizes (SNRs = 5.3 and 5.2, respectively), path length (SNR = 4.9), direction of the straight line from begin to end (SNR = 7.8), stroke duration (SNR = 3.6), and the surface of loops (SNR = 5.4). Characterization of stroke shapes can be done by the directions of 5 equal-duration subsegments (SNRs = 6.8, 7.5, 6.5, 4.2, and 2.9, respectively). The SNR values are typical values based on 16 replications of a pattern ’elementary’ by one subject. It is interesting to note that the directions of the subsegments at the end of a stroke tend to be unreliable; this is probably due to the programming of the endpoint of the stroke without programming the exact path towards the endpoint [31]. The axial pen pressure is not relevant in on-line handwriting recognition as pen pressure is not related to particular letters or shapes [45].

The following models will be discussed: Orientation-free models, which do not require normalization of orientation of the writing pattern, mass-spring or oscillatory models, movement optimization models, which allow the reconstruction of dynamics from off-line cursive script, and symbolical models.

7.1 Orientation-Free Models

There is a class of models which does not assume any specific x or y axes, nor axes related to biomechanical joints. Movement patterns are unlikely to be specified in the body-oriented horizontal and vertical coordinates [12]. Instead, discrete "circular strokes" are proposed while the inertia of the finger-hand-arm system smooths the movement by acting as a low-pass filter [25]. Circular strokes are described by the following [12]: (1) Circle segments of specific curve radius, arc length, and orientation are fitted to the trace at the point of peak velocity, (2) a half Gaussian velocity function is fitted from the previous absolute-velocity peak to the current peak and the other half phase from the current peak to the following peak, (3) the overlapping functions are averaged sample-by-sample. Therefore, 6 features per ballistic stroke are needed: curve radius, arc length, orientation, peak velocity, duration between previous and following velocity peak, and the asymmetric position of the peak velocity between the previous and following ones. The accuracy of this model does not seem as high as the simpler models, which were based on component movements per axis [35].

A model which allows a more accurate reconstruction of writing patterns has been proposed, but it requires more features per stroke [15]. As in the previous model, absolute velocity is approximated by a piecewise Gaussian function, but, instead of the constant-radius approximation, the angular velocity is also approximated by a piecewise Gaussian function. The piecewise Gaussian functions are generated by a hypothetical neural velocity signal consisting of a sequence of rectangular time functions. The height of each block represents the gain of the muscular speed-generator. In order to simulate sharp movement reversals, discontinuities had to be inserted in the angular velocity. The cursive-script word "bug", containing about 12 ballistic strokes, could be reconstructed accurately using 60 features for the tangential velocity, and, 52 features for the angular velocity, or 9 features per ballistic stroke.

7.2. Mass-Spring Model

The mass-spring model was introduced to model fluent movement trajectories in a parsimonious way [11]. This model inspired several cursive-script recognizers [21, 46]. A horizontal x axis, parallel to the left-to-right translation and a vertical y axis, and fitted piecewise sinusoids as if they resulted from a frictionless mass-spring system. In the vertical movement component, the mass consists of the weights of the pen and the fingers, whereas for the horizontal movement component, the mass consists of the weights of the whole hand and the pen. Various sinusoids can be concatenated to a fluent handwriting trajectory, while the equilibrium points, the stiffness, and some initial conditions of the mass-spring system are changed at specific moments. The major advantage of the model is that it allows the generation of repetitive loops such as "eeee" very parsimoniously as no features need to be changed during this pattern. The model is most easily expressed in terms of the velocity-time functions Vx(t) and Vy(t), respectively:

$$Vx(t) = Vx\_peak * \sin (omega\_x * t + phi\_x) + c$$

$$Vy(t) = Vy\_peak * \sin (omega\_y * t + phi\_y)$$

where Vx_peak and Vy_peak are the velocity amplitudes, omega_x and omega_y the angular frequencies, or frequencies multiplied by 2*pi, and phi_x and phi_y the initial phases, respectively. The constant-velocity, horizontal left-to-right movement c, is added to the horizontal component. The position-time functions can be found by integration using zero initial velocity.

Three simplifications can be made. (1) Only the phase difference is relevant, i.e., phi_x - phi_y = phi. (2) Peak velocities can be chosen equal to twice the horizontal left-to-right velocity, i.e., Vx_peak = Vy_peak = 2c. (3) The horizontal and vertical frequencies are equal, i.e., omega_x = omega_y = omega, because unequal frequencies are only needed in rare patterns like "8". Varying the phase difference phi yields various basic patterns: upright loops (phi = 90 degrees), slanted loops (phi = 60 degrees), garlands (phi = 30 degrees), waves (phi = 0 degrees), and arcades (phi = -30 degrees). Another important parameter to influence stroke shape is the horizontal velocity Vx, at the top of a stroke, where Vy changes from upward to downward, i.e., Vx = c -Vx_peak * sin (phi). If a sharp movement reversal has to be programmed, requiring Vx to change sign, the parameters phi, Vx_peak and c need to be adjusted simultaneously.

Ascenders and descenders can be generated by both increasing Vy_peak and decreasing omega. In order to maintain a stable baseline, Vy_peak can only be changed at the top or at the bottom of a stroke, i.e., Vy = 0, yielding a descender or ascender, respectively. If Vy_peak is changed, then the slant would change as well. In terms of model parameters, the slant beta equals:

$$beta = \arctan (Vy\_peak / (Vx\_peak*\cos(phi))),$$

which again depends upon several parameters. When increasing Vy_peak in order to generate an ascender, either phi or Vx_peak have to be changed in order to maintain slant. Changing two parameters simultaneously does not seem very natural. Ascenders and descenders can be generated, without changing slant, by decreasing angular frequency omega, but this does not seem to be natural either [34].

In summary, it seems not easy to simulate handwriting patterns using these parameters. It is difficult to generate more complex repetitive patterns such as "ellellell...". Furthermore, normal handwriting does not have a predominant sinusoidal movement [26], nor a constant left-to-right trend [41]. An example of a simulation of the letter sequence "elye", counting 11 ballistic strokes, requires 30 parameter changes and time moments apart from the 10 initial parameter settings. This yields only 3.6 parameters per stroke, which is parsimonious indeed, as it is less than 4.

7.3. Movement Optimization Models in Off-line Cursive Script

A general trajectory-formation model minimizes the mean squared rate of force change (i.e., jerk) [24]. In the frictionless case, force can be approximated by the second time derivative of the horizontal and the vertical position time functions. The rate of force change can be approximated by the third time derivative. This model has been used in a cursive-script recognizer [22]. This model allows generation of the "kinematics from shape", so that the recognizer would accept thinned, static images. Scanned images of cursive script can be handled in less flexible stroke models [19]. Pixel maps are analyzed in terms of common strokes having the shapes of half ellipses, thus preventing awkward pixel-quantization effects.

In the "minimum-jerk model" cursive script consists of a sequence of curved segments, e.g., up-down or down-up stroke pairs [24]. In order to simulate curved segments, the minimum-jerk model needed to be extended by a "via point" near the point of maximum curvature [47]. Such a curved segment will be referred to as a "via stroke". A via stroke corresponds to a pair of "ballistic strokes". Minimization of the mean squared jerk with a via-point constraint yields 5th-order spline functions:

x(t) = ax0 + ax1*t + ax2*t**2 + ... + ax5*t**5 + px*max(0,t-t1)**5,

where ax0 ... ax5 and px are the spline parameters and t1 is the optimized time duration to reach the via point. A similar function follows for y(t).

Cursive-script patterns can be simulated using a limited set of basic via strokes:

Hook (like cursive "i" without dot), cup (like cursive "v"), gamma (like cursive "l"), and oval (like cursive "o"). However, the minimum-jerk model with only via-point constraints did not allow simulation of the cup, gamma, and oval [24]. It was unclear where extra via points were needed apart from the natural via point near the point of maximum curvature. In order to generate the latter patterns, the direction (i.e., dy/dx) at the beginning and endpoints of the via stroke was added as a constraint. Under the border condition of zero velocities and accelerations at the endpoints, the directions are undefined. According to the mathematical "de l'Hopital" rule, the next nonzero derivative has to be constrained, which is jerk. When jerk needs to be constrained, mean squared jerk should not be minimized, but the mean squared fourth time derivative (i.e., snap). This yields a 7th-order spline function, analogous to the 5th-order spline function above.

The minimum-snap model employs 18 parameters per via stroke: ax0, ... ax7, px, ay0, ..., t1), however, most parameters, and even duration t1 to reach the via point, are fixed by the boundary conditions. The boundary conditions are that the beginning and end of the via strokes should have a smooth connection to the adjacent via strokes, i.e., equal horizontal and vertical positions, velocities, accelerations, and jerks. The two boundary conditions of the via point, fix 4 parameters per via stroke. The minimization equation for the mean squared snap fixes two more parameters for x and y axes. In total, 6 parameters per via stroke (or only 3 parameters per ballistic stroke) are required: x and y positions at beginning and via points, and direction and amplitude of the jerk at the beginning point.

7.4. Symbolical Models

A symbolic model simulates arbitrary texts of a particular writer's handwriting style [48]. Interesting features of this model are the implementation of visual feedback, and the writer-specific motor memory containing "symbolic letter descriptions". Visual feedback monitors the baseline and the lineation levels. When monitoring a departure from the baseline or the lineation levels, an exponentially decaying lineation memory is used as a reference to adjust vertical sizes of future strokes.

It is interesting to note that the generated handwriting patterns are not based on curve fitting of a particular pattern, but contain departures of the same magnitude as those between individual replications. The correlations of the horizontal and vertical stroke sizes and durations between the original patterns appeared similar as those between the simulated pattern and the original ones. Therefore, the simulation model is sufficiently accurate.

The symbolic letter descriptions of the allographs were trained on the basis of a corpus of a writer’s handwriting. Thus, each allograph is represented by a sequence of strokes, where each stroke is represented by 4 stroke parameters: dX (horizontal displacement per stroke), dY (vertical displacement per stroke), T (compound stroke duration), and C (stroke-shape factor). These parameters have been based upon the average of several manually selected replications of the same allograph in various contexts. At the symbolic module, dX/dY was quantized into close, normal, or far, and dY was quantized into descender, descender-plus, base-minus, base, base-plus, body-minus, body, body-plus, ascender-minus, and ascender.

Although there is no evidence that durations are represented in motor memory [10, 30], the "compound stroke duration" T is used as a parameter to conveniently interpret the required "stroke-shape factor" C. Unfortunately, the SNR of the stroke-shape factor appears to be very low. Nevertheless, this model is useful as an example of parsimonious coding of cursive script. The interval between successive zero crossings in the x velocity component (i.e., t1 (vx=0) and t2 (vx=0), respectively) can be defined as the x-stroke duration, and similarly for the y-stroke duration. The compound stroke duration is defined as the average of the x and the y-stroke durations:

$$T = ((t2\ (vx=0) - t1\ (vx=0)) + (t2\ (vy=0) - t1\ (vy=0))) / 2$$

T ranges between about 50 and 150 ms. The stroke-shape factor C is defined as the time interval between two nearby zero crossings of x and y velocities, relative to the compound stroke duration T:

$$C = (t1\ (vx=0) - t1\ (vy=0)) / T.$$

The shape factor is a generalized phase difference between x and y velocity-versus-time functions. If the x velocity component is ahead of the y component, then the stroke shape will form (part of) a counterclockwise loop (i.e., $-1.5 < C < 0$). In the opposite case, the stroke shape will form (part of) a clockwise loop (i.e., $0 < C < 1.5$). In the special case that the x and y zero crossings occur simultaneously ($C = 0$), there will be a sharp stroke ending, followed by a movement reversal.

The procedure to translate the allographic-code into the required movement patterns is as follows: In the symbolic module, specific connecting strokes have to be inserted between pairs of allographs and punctuation signs. The connecting stroke depends upon the final stroke of the preceding allograph and the initial stroke of the subsequent allograph. The parameters of the connecting stroke have been estimated by the average of the replications in a corpus of handwriting. The "cursive connections grammar" contains the generic rules prescribing the connecting strokes; for example, the input "an ad..." is expanded to: "(pendown) (a) (base to midline, clockwise, close progression) (n) (base to base-plus, sharp ending, close progression) (penup) (space) (pendown) (a) (base to midline, sharp ending, normal progression) (d) ...". Note that penup and pendown signals are part of the coding.

At the quantitative level, the strokes per allograph are selected from a letter data base. The "quantitative letter descriptors" describe the strokes in terms of dX, dY, T, and C. The compound stroke duration T, and the form factor C, allow the approximation of the moments in time where x and y velocities change sign. In order to generate the kinematics of a handwriting trajectory, a general form of the velocity pattern is selected. For convenience, a sinusoid velocity time function is selected to fit between successive zero crossings of x and y velocities, which appear to approximate handwriting movement patterns relatively well [35].

8. Summary

1. This paper provides a guide for finding and verifying appropriate handwriting features for on-line cursive-script recognition. Appropriate features for on-line cursive-script recognition systems can be derived from handwriting production models and will be only slightly sensitive to the within-writer motor-noise variations. A macroscopic model of handwriting production, having several modules, was presented. The highest module is a movement memory containing the parsimonious and abstract description of the handwriting letters. In the lower modules, the abstract description is translated into individual motor actions.

2. In order to extract features, handwriting patterns are split into a sequence of motor actions such as ballistic strokes. For that purpose, an invariant segmentation procedure is proposed, using a mixture of motor-based rules (segment at absolute-velocity minima; segmentations at least 60 ms apart) and empirical rules (weigh the horizontal velocity component less than the vertical velocity component when estimating the absolute-velocity; concatenate short strokes and parallel strokes).

3. A spectral analysis shows that the frequency bandwidth increases with writing speed. The bandwidth of 10 Hz and stroke durations of 0.1 s suggests that four sample values are enough to describe both components of an individual stroke. Ideally, four features per stroke are needed in order to disambiguate strokes.

4. Invariant features are the ones showing a high signal-to-noise ratio (SNR). In order to estimate the SNR various replications of a writing pattern are required. SNR values of invariant features are of the order of 4.

5. Invariant features may be closely related to the movement information stored

in motor memory, whereas less invariant features may be derived at the lower levels. The top-down hierarchy describes the causal relations between higher-level and lower-level features. The most invariant features are graphical features, e.g., stroke sizes and directions. Less invariant features are the dynamical parameters, e.g., forces and durations.

Variations in one stroke may affect neighboring strokes. This is described by the sequence hierarchy. Stroke sizes show a sequence hierarchy but stroke duration does not. This supports that the segmentation into strokes yields a sequence of motor actions, but, that strokes do not vary individually. Strokes should be considered in their context. Features which change proportionally or linearly for all strokes are subject to lower-level scale transformations and have to be normalized.

6. Context of stroke size and stroke curvature mainly affects durations and speeds. In normal handwriting, stroke size is varied by both duration and accelerative force. Absolute velocity reaches minima at points of high curvature.

7. The most important features for cursive-script recognition are stroke sizes and stroke directions as they have high SNRs. Computational models suggest complete sets of features which allow legible reconstruction of writing patterns. Movement optimization models allow the reconstruction of the handwriting movement from static images and, therefore, lead to optical recognition of cursive script.

References

1. Ellis, A.W. (1982). Spelling and writing (and reading and speaking). In A.W. Ellis (Ed.), Normality and pathology in cognitive functions (pp. 113-146). London: Academic Press.
2. Margolin, D.I. (1984). The neuropsychology of writing and spelling: Semantic, phonological, motor and perceptual processes. Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 36A, 459-489.
3. Van Galen, G.P., & Teulings, H.L. (1983). The independent monitoring of form and scale parameters in handwriting. Acta Psychologica, 54, 9-22.
4. Van Galen, G.P. (1991). Handwriting: Issues for a psychomotor theory. Human Movement Science, 10, 165-191.
5. Teulings, H.L., Thomassen, A.J.W.M., & Van Galen, G.P. (1983). Preparation of partly precued handwriting movements: The size of movement units in writing. Acta Psychologica, 54, 165-177.
6. Denier van der Gon, J.J., & Thuring, J.Ph. (1965). The guiding of human writing movements. Biological Cybernetics, 2, 145-148.

7. Teulings, H.L., Thomassen, A.J.W.M., & Maarse, F.J. (1989). A description of handwriting in terms of main axes. In R. Plamondon, C.Y. Suen, & M. Simner (Eds.), Computer recognition and human production of handwriting (pp. 193-211). Singapore: World Scientific.
8. Pick, H.L., & Teulings, H.L. (1983). Geometric transformations of handwriting as a function of instruction and feedback. Acta Psychologica, 54, 327-340.
9. Guiard, Y. (1993). On Fitts's and Hooke's laws: Simple harmonic movement in upper-limb cyclical aiming. Acta Psychologica, 82, 139-160.
10. Teulings, H.L., & Schomaker, L.R.B. (1993). Invariant properties between stroke features in handwriting. Acta Psychologica, 82, 69-88.
11. Hollerbach, J.M. (1981). An oscillation theory of handwriting. Biological Cybernetics, 39, 139-156.
12. Morasso, P., Mussa Ivaldi, F.A., & Ruggiero, C. (1983). How a discontinuous mechanism can produce continuous patterns in trajectory formation and handwriting. Acta Psychologica, 54, 83-98.
13. Tappert, C.C. (1986). An adaptive system for handwriting recognition. In H.S.R. Kao, G.P. Van Galen, R. Hoosain (Eds.), Graphonomics: Contemporary research in handwriting (pp. 185-198). Amsterdam: North-Holland.
14. Viviani, P., & Cenzato, M. (1985). Segmentation and coupling in complex movements. Journal of experimental psychology: Human perception and performance, 11, 828-845.
15. Plamondon, R. (1989). Handwriting Control: A functional model. In R.M.J. Cotterill (Ed.), Models of brain function (pp. 563-574). London: Cambridge University Press.
16. Kadirkamanathan, M. & Rayner, P.J.W. (1990). A scale-space filtering approach to stroke segmentation of cursive script. In R. Plamondon, & G. Leedham (Eds.), Computer Processing of Handwriting (pp. 133-166). Singapore: World Scientific.
17. Higgins, C.A., & Ford, D.M. (1992). A new segmentation method for cursive script recognition. In S. Impedovo and J.C. Simon (Eds.), From pixels to features III: Frontiers in handwriting recognition (pp. 75-86). Amsterdam: North-Holland.
18. Wright, Ph.T. (1988). Design and implementation of a handwriting recognizer. Plessey New Technology, 2, 1988.
19. Helsper, E., Schomaker, L.R.B., Teulings, H.L., & Abbink., G.H. (1993). Feature description of optically scanned handwritten words. Proceedings of the Sixth International Conference on Handwriting and Drawing (pp. 270-272). Paris: Telecom.
20. Wing, A.M. (1978). Response timing in handwriting. In G.E. Stelmach (Ed.), Information processing in motor control and learning (pp.153-172). New York: Academic Press.
21. Abbink, G.H., Teulings, H.L., & Schomaker, L.R.B. (1993). Description of on-line

script using Hollerbach's generation model. Pre-Proceedings of the Third International Workshop on Frontiers in Handwriting Recognition (pp. 217-224). Buffalo, NY: Cedar.

22. Edelman, S., Flash, T., & Ullman, S. (1990). Reading cursive handwriting by alignment of letter prototypes. International Journal of Computer Vision, 5, 303-331.
23. Guberman, Sh.A., & Rozentsveig, V.V. (1976). Algorithm for the recognition of handwritten text. Avtomatika i Telemekhanika, 5, 122-129.
24. Edelman, S., & Flash, T. (1987). A model of handwriting. Biological Cybernetics, 57, 25-36.
25. Van Galen, G.P., & Schomaker, L.R.B. (1992). Fitts' law as a low-pass filter effect of muscle stiffness. Human Movement Science, 11, 11-21.
26. Teulings, H.L., & Maarse, F.J. (1984). Digital recording and processing of handwriting movements. Human Movement Science, 3, 193-217.
27. Teulings, H.L., Schomaker, L.R.B., Abbink, G.H., & Helsper, E. (1993). Invariant segmentation of on-line cursive script. Proceedings of the Sixth International Conference on Handwriting and Drawing (pp. 198-200). Paris: Telecom.
28. Jerri, A.J. (1977). The Shannon sampling theorem -- Its various extensions and applications: A tutorial review. Proceedings of the IEEE, 65, 1565-1596.
29. Langolf, G.D., Chaffin, D.B., & Foulke, J.A. (1976). An investigation of Fitts' law using a wide range of movement amplitudes. Journal of Motor Behavior, 8, 113-128.
30. Teulings, H.L., Thomassen, A.J.W.M., & Van Galen, G.P. (1986). Invariants in handwriting: The information contained in a motor program. In H.S.R. Kao, G.P. Van Galen, & R. Hoosain (Eds.), Graphonomics: Contemporary research in handwriting (pp. 305-315). Amsterdam: North-Holland.
31. Schomaker, L.R.B., & Teulings, H.L. (1992). Stroke- versus character-based recognition of on-line, connected cursive script. In S. Impedovo and J.C. Simon (Eds.), From pixels to features III: Frontiers in handwriting recognition (pp. 313-325). Amsterdam: North-Holland.
32. Thomassen, A.J.W.M., & Schomaker, L.R.B. (1986). Between-letter context effects in handwriting trajectories. In H.S.R. Kao, G.P. Van Galen, R. Hoosain (Eds.), Graphonomics: Contemporary research in handwriting (pp. 273-288). Amsterdam: North- Holland.
33. Seni, G., Ng, S. (1993). Towards an on-line cursive word recognizer. Pre-Proceedings of the Third International Workshop on Frontiers in Handwriting Recognition (pp. 449-454). Buffalo, NY: Cedar.
34. Thomassen, A.J.W.M., & Teulings, H.L. (1985). Time, size, and shape in handwriting: Exploring spatio-temporal relationships at different levels. In J.A. Michon & J.B. Jackson (Eds.), Time, mind, and behavior (pp. 253-263).

Heidelberg: Springer.

35. Plamondon, R., & Maarse, F.J. (1989). An evaluation of motor models of handwriting. IEEE Transactions on Systems, Man, and Cybernetics, 19, 1060-1072.

36. Michel, F. (1971). Etude experimentale de la vitesse du geste graphique. Neuropsychologia, 9, 1-13.

37. Thomassen, A.J.W.M., & Teulings, H.L. (1983b). The development of handwriting. In M. Martlew (Ed.), The psychology of written language: Developmental and educational perspectives (pp. 179-213). New York: Wiley.

38. Greer, K.L., & Green, D.W. (1983). Context and motor control in handwriting. Acta Psychologica, 54-205-215.

39. Wing, A.M. (1980). The height of handwriting. Acta Psychologica, 46, 141-152.

40. Lacquaniti, F., Terzuolo, C., & Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. Acta Psychologica, 54, 115-130.

41. Thomassen, A.J.W.M., & Teulings, H.L. (1983a). Constancy in stationary and progressive handwriting. Acta Psychologica, 54, 179-196.

42. Maarse, F.J., Thomassen, A.J.W.M. (1983). Produced and perceived writing slant: difference between up and down strokes. Acta Psychologica, 54, 131-147.

43. Meeks, M.L., & Kuklinski, T.T. (1990). Measurement of dynamic digitizer performance. In R. Plamondon, & G. Leedham (Eds.), Computer Processing of Handwriting (pp. 89-110). Singapore: World Scientific.

44. Schomaker, L.R.B., & Teulings, H.L. (1990). A handwriting recognition system based on properties of the human motor system. In C.Y. Suen (Ed.), Frontiers in handwriting recognition (pp. 195-209). Montreal: Cenparmi. ISBN: 1-895193-00-1.

45. Schomaker, L.R.B., & Plamondon, R. (1991). The relation between pen force and pen point kinematics in handwriting. Biological Cybernetics, 63, 277-289.

46. Singer, Y., & Tishby, N. (1993). Cursive handwriting word spotting using a discrete dynamic approach. Pre-Proceedings of the Third International Workshop on Frontiers in Handwriting Recognition (pp. 373-378). Buffalo, NY: Cedar.

47. Flash, T., & Hogan, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model. Journal of Neuroscience, 5, 1788-1703.

48. Schomaker, L.R.B., Thomassen, A.J.W.M., & Teulings, H.L. (1989). A computational model of cursive handwriting. In R. Plamondon, C.Y. Suen, & M. Simner (Eds.), Computer recognition and human production of handwriting (pp. 153-177). Singapore: World Scientific.