Grip Kinetic Profile Variability in Adult Signature Writing

Bassma Ghali1,2, Khondaker Mamun1,2 and Tom Chau1,2*

1Bloordview Research Institute, Holland Bloordview Kids Rehabilitation Hospital, 150 Killgour Road, Toronto, Ontario, Canada
2Institute of Biomaterials and Biomedical Engineering, University of Toronto, 164 College Street, Toronto, Ontario, Canada

Abstract

Previous studies of handwriting grip kinetics have demonstrated the ability to classify writers based on the topography of grip forces associated with signature writing. However, the topographic representation requires a large array of individual sensors in practice. The possibility of differentiating participants on the basis of a summative, temporal force profile is yet unknown. In this study, we investigated the variability of features derived from a time-evolving total grip force profile. Using an instrumented writing utensil, twenty adult participants provided 600 samples of a well-practised bogus signature over a period of 10 days. Deploying a combination of temporal, spectral and information-theoretic features, a linear discriminant analysis classifier outperformed nonparametric and nonlinear classifier alternatives and discriminated among participants with an average classification rate of 5.8% as estimated by cross-validation. These results suggest the existence of a unique kinetic profile for each writer even when generating the same written product. Our findings highlight the potential of using grip kinetics as a biometric measure.

Keywords: Handwriting kinetics; Grip force profile; Signature verification; Classification; Inter-writer variability; Feature selection

Introduction

Signature writing is a well-learned [1] but highly complex perceptual motor task [2-3], invoking the coordinated activation of both proximal (e.g., thenar) and distal (e.g., trapezius) muscles [4], relying upon the central role of proprioception and the secondary role of vision [5], integrating kinesthetic input [6] and harnessing short-term memory [7]. The complexity of these biomechanical and cognitive systems introduces variations between and within individuals [2].

The advent of kinetic utensils that measure the forces applied by the fingers on the pen [8-10] has spawned numerous investigations of handwriting grip kinetics. Most of these studies have been clinical in nature, with the goal of using handwriting grip forces to inform diagnosis and treatment of handwriting disorders [8,9,11,12]. However, studies of handwriting grip kinetic variability in the adult population is very limited [3]. Recently, we studied the variability of grip kinetics associated with signature writing in adults [1,13] and found that the variability of kinetic topographies (i.e., grip shape) between individuals was much higher than the variability within an individual, even when considering signatures collected over several months. These findings encouraged the study of variations in a summative handwriting grip kinetic profile, which is the time series of total force variation over the course of a signature.

In sports such as tennis, baseball and golf, the within- and between-subject variations of kinetic profiles have been studied with the aim of optimizing player performance [14]. The grip force profile of a golf swing was found to be repeatable within a player and distinctive between players [14-16]. Gait studies have found that the kinetics associated with walking in adult participants are repeatable for the same person on multiple days [17,18]. Grip force patterns have also proven to be valuable for biometric verification in gun control applications, which suggests grip consistency within individuals [19,20]. A recent review of keyboarding-based biometrics showed that, in addition to considering the time to type, the latency between keystrokes, and many other features, few studies have considered the keystroke pressure applied by the fingers on the keys, a potential discriminative feature [21]. Salami et al. [22] and Sulong et al. [23] proposed a keyboard embedded with sensors capable of measuring the force applied on each key and the latency between keystrokes as a mean to authenticate a user while typing. Using multiple classifiers, it was found that the combination of pressure and latency yielded better user authentication than that achievable with either feature alone. In similar spirit, finger pressure has been deemed to be more discriminative than hold-time and inter-key duration for the authentication of touch pad users [24].

In the field of handwriting authentication and signature verification, intra- and inter-participant variations of axial pen pressure, spatiotemporal features and kinematic characteristics have been explored [2,25-29]. Ramsay [2] modeled the dynamics of spatiotemporal information of handwriting using a differential equation and classified handwriting samples of different individuals. Lei and Govindaraju [26] examined the consistency and discriminative power of multiple features commonly used in on-line signature verification systems, concluding that the pen-tip coordinates, speed, and angle between the speed vector and the horizontal axis of the writing surface were among the most consistent features. Another study found that the dynamic features (speed, angle, axial pressure, and acceleration) surpassed static features in discriminative capability between genuine writing and skilled forgeries [27]. Bashir and Kempf [30,31] recently identified person-specific features in grip force signals and reported improved writer recognition when a grip force signal was added to the classifier. However, these studies were based on samples collected in one session and did not examine the effect of intra-subject variability over time. The discriminative potential of handwriting grip kinetics has yet to be fully ascertained. This study thus set out to investigate one aspect of this potential, namely, to quantify the variability of grip kinetic profiles between adults while writing the same well-practiced signature over multiple days and multiple times within the same day.

*Corresponding author: Tom Chau, Institute of Biomaterials and Biomedical Engineering, University of Toronto, 164 College Street, Toronto, Ontario, Canada, Tel: +1 416 425 6220 (ext 3515); Fax: +1 416 425 1634; E-mail: tom.chau@utoronto.ca

Received August 05, 2013; Accepted September 16, 2013; Published September 21, 2013


Copyright: © 2013 Ghali B, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Methods

Participants

To generate the required database, we recruited a convenient sample of 20 participants with an age range of 18 to 45 years and a mean of 27 ± 6 years. The sample included 8 males/12 females and 17 right-handed/3 left-handed participants. Individuals with a known history of musculoskeletal injuries or neurological impairments were excluded from the study. The study protocol was approved by the research ethics boards of Holland Bloorview Kids Rehabilitation Hospital and the University of Toronto. An informed written consent was signed by each participant. Demographic information such as age, gender and handedness were also collected from each participant.

Instrumentation

To collect the required data, the instrumentation setup shown in Figure 1 was used. Grip force signals, which are the forces applied by the fingers on the pen barrel, were acquired at 250 Hz via an instrumented writing utensil adorned with an array of Tekscan 9811 force sensors [1,8]. A custom-made data acquisition box transferred these signals to the data collection computer where they were saved for processing. A systematic calibration procedure, detailed in Ghali et al. [1], was performed on the force sensors every 2 to 3 days to derive calibration curves needed during pre-processing of the grip force data. During calibration and data analysis, only the subset of sensors that covered the pen barrel (32 sensors of the 96 sensors array) was considered. The sensor array was replaced when a sensor malfunction due to wear and tear was observed. The axial forces applied by the pen on the writing surface along with the pen tip position and pen angles (twist, altitude and azimuth) were acquired by a Wacom Cintiq 12 WX digitizing LCD display at a frequency of 105 Hz. These latter signals were synchronized with the grip force signals using the developed data acquisition software. A grounding strap was placed on the non-dominant hand of each participant to reduce noise in the grip force signals.

Data collection protocol

The bogus signature shown in Figure 2 was given to all participants. Each participant practiced the bogus signature by writing it repeatedly on paper 25 times a day for two weeks. After the practice period, each participant completed 30 sessions of data collection over 10 days that spanned an average period of 20:4 ± 3:6 days depending on the participant’s availability. On each day, three sessions were performed at different times of the day. In each session, the participant sat comfortably on a chair, donned the grounding strap on the non-dominant hand, held the instrumented pen with the dominant hand and wrote multiple signature samples within the specified area on the digitizing tablet. A researcher was always present to check the force sensors at the beginning of each session and to note any writing mistakes or events that could affect the data. In each session, a 10 second baseline of the force sensors was collected prior to writing to determine the pre-grip value of each sensor. Twenty samples of the bogus signature and twenty samples of the participant’s authentic signature were collected during each session. A total of 600 well-practiced bogus signatures and 600 authentic signatures were obtained from each participant over the 30 sessions. In this study, only the bogus signatures are considered. An analysis of the authentic signatures was reported in Ghali et al. [1].

Data pre-processing

Through visual inspection of the bogus signatures and their associated grip force signals, and a review of the researcher notes taken during each session, some signature samples were identified as possessing a long pause, a mistake or a sensor malfunction while writing. These samples (417 bogus signatures across all participants) were excluded from subsequent analyses. Other signature samples that exhibited extra strokes because of accidental contact with the writing surface before or after writing were salvaged by adjusting the start or end time of the signature, respectively.

The high frequency noise in the grip force signals was removed using a Butterworth low pass filter with a cut off frequency of 10 Hz, below which resided more than 95% of the signal power. Some samples (1105 bogus signatures) had a low frequency oscillating noise that could not be removed and were thus excluded from further analyses. The remaining 10478 signature samples (87.3% of the 12000 samples; average 524 samples per participant) were translated into signals with physical force units (Newtons) [1].

Feature extraction

The total grip force signal of each signature sample was obtained by adding the pre-processed grip force signals of the 32 individual sensors. Figure 3 exemplifies a bogus signature sample, the associated 32 pre-processed grip force signals and the summative total grip force signal.

Two genre of features were extracted from the total grip force signals: (1) signature-exclusive features which are extracted only from the total grip force signal during signature writing and (2) referenced features which represent the closeness of a given writing sample to the reference signal of a participant. These signature-exclusive features include:

- Mean of the total force signal as a measure of location of the signal values
- Maximum of the total force signal as a measure of the grip strength
- Interquartile range of the total force signal as a robust measure of kinetic dispersion

Figure 1: Data collection instrumentation setup.

Figure 2: The bogus signature. Each participant practiced this signature for two weeks prior to the data collection to develop familiarity with the signature.
• Coefficient of variation (CV) of the total force signal as a normalized measure of kinetic dispersion. CV is the ratio of the standard deviation to the mean of the signal.

• Skewness of the total force signal, which is a measure of asymmetry of the signal distribution

• Kurtosis of the total force signal, which is a measure of the peakedness of the signal distribution

• Number of zero crossings of the detrended total force signal as a crude measure of the frequency content of the signal

• Number of peaks in the total force signal

• Centroid frequency of the power spectral density of the total force signal

• Bandwidth of the power spectral density of the total force signal

• Maximum power of the power spectral density of the total force signal

• Sub-band power of the total force signal. Since most of the energy content was below 5 Hz, the sub-band power was calculated in 5 frequency bands each with a 1 Hz window size.

• Entropy rate of the total force signal, which is a measure of the regularity of the signal

A more detailed explanation of many of these features can be found in Lee [32].

For the second group of features, a reference total force signal of each participant was first estimated according to the following steps:

**Time normalization:** The total force signals of all participants were resampled as necessary such that all total grip force signals shared a common length, which was chosen to be the average length across all participants and all signature samples.

**Registration:** To correct any temporal misalignment among the time-normalized total force signals of each participant, the signals were subjected to curve registration as described in Chau et al. [33]. Figure 4 portrays an example of the total force signals before and after registration and the associated mean signals for one of the participants.

The reference signal calculation: The mean total force signal for each participant was estimated as the average of the registered total force signals across signatures of the given participant. Figure 5 depicts the overall mean total force signal based on all 20 participants along with the mean total force signals of 2 participants as examples.

The second group of features measures the similarity between the reference curve of each participant and all other total force signals belonging to that participant (within-participant similarity measure) and belonging to other participants (between-participant similarity measure). These features were:

• Pearson correlation coefficient (NCC) between two signals, which is a measure of the strength of linear dependence (correlation) between two signals.

• Root mean square error (RMSE) between two magnitude-normalized signals (i.e. zero mean and unit standard deviation signals), which reflects the distance between two signals.

• Cost of registering two signals, which is the sum-of-squares criterion function detailed in Chau et al. [33].

![Figure 3: A signature sample, its grip force signals and total grip force profile. Top: a signature sample and the associated timing of selected points; Middle: the pre-processed grip force signals of the 32 force sensors; Bottom: the total grip force signal over the course of a signature, which is the sum of the signals shown in the middle figure.](image)

![Figure 4: Example of total force signals for one participant before (left graph) and after (center graph) registration. To facilitate visualization, only a subset of total force signals is shown. Pre- and post-registration mean total force profiles for the given participant appear in the rightmost graph.](image)

![Figure 5: Examples of mean total force signals. The solid line shows the overall mean total grip force signal based on all participants while each dotted line exemplifies the mean total grip force signal of a participant.](image)
Pattern classification

To ascertain the within-participant consistency of the total force signals and the extent of the between participant variability, a binary classifier was created for each participant. For the $i^{th}$ classifier, $i=1,…,20$, the true signatures class included the bogus signatures that belonged to the $i^{th}$ participant, while the false signatures class entailed an equal number of bogus signatures randomly selected from the other 19 participants. For each classifier, the inputs were the features extracted from the total force signal of each signature and the output was a binary output indicating whether the signature sample belonged to the $i^{th}$ participant (true sample) or not (false sample). For each participant, the mean misclassification rate (MCR) was estimated based on ten iterations of 10 fold cross validation. In each iteration of the cross validation, a different random set of false signature samples were selected.

Several different classifiers were considered including a simple linear classifier, the linear discriminant analysis (LDA) classifier [34], a probabilistic classifier (Naive Bayes), a nonparametric nonlinear classifier (K-nearest neighbor (KNN)) and a parametric nonlinear classifier (support vector machine (SVM) with radial bases function (RBF)). For each classifier, the mean MCR as well as the percentage of false positives (FP) and false negatives (FN) were tabulated based on 100 folds (10 iterations $\times$ 10 folds).

Classification was first performed using all 20 extracted features. To remove potential feature redundancy and to reduce dimensionality, a subset of 9 features was systematically selected according to the procedure below.

1. The sample covariance matrix of the features vectors was calculated. Six features were excluded due to high inter-feature correlations.

2. With the 14 remaining features, weighted sequential feature selection (WSFS) [35] was invoked to find the most discriminatory set of features in each iteration for each participant. Features were ranked based on their individual discriminability. The optimal subset of features for each iteration of 10 folds was identified as if the feature surfaced the most frequently while yielding the lowest MCR.

3. To minimize feature space dimensionality and to hone in on a uniform set of features across participants, only features that were frequently selected across participants were admitted to the final feature set. Specifically, 9 features emerged: mean, CV, skewness, centroid frequency, bandwidth, entropy rate, sub-band power (2 Hz $\leq f<3$Hz), NCC and RMSE.

The above analysis was performed with an unordered (i.e., randomly selected) set of true samples from each participant. Random selection ensured that the training set included samples across sessions. Therefore, both training and testing sets likely contained signature samples from the same session. To determine the effect of training and testing with samples from different sessions on MCR values, the same analysis was repeated with sequentially ordered true samples. In this latter case, the testing set included samples from sessions that were not part of the training set. A Wilcoxon rank sum test was performed for each participant to compare the two groups of MCRs.

The effect of reducing the number of samples on classification performance was also examined. Subsets of decreasing size, from 100% to 10% of the samples available for each participant were considered. For each subset, the mean MCR was calculated based on 10 iterations of 10 fold cross validation.

Results

Figure 6 presents the average MCRs of the four classifiers using all 20 features (unshaded bars) and using 9 selected features (shaded bars). On average, 943 (90%) training samples and 105 (10%) testing samples were used in each fold of cross validation. Only the SVM classifier performance improved with feature selection. By statistically comparing the average MCR, FP and FN obtained with the 20-feature LDA classifier and the same quantities for classifiers of decreasing feature dimension, it was found classification performance is preserved down to 9 features (p=0.126, 0.07, and 0.36 respectively). With only 8 features, FP increased significantly (p=0.044). Likewise, with only 7 features, MCR was significantly higher (p =0.046). Since the simple LDA classifier with all 20 features yielded the best performance overall, it will be the focus of the subsequent analyses.

Figure 7 provides a more detailed breakdown of the LDA classifier performance across participants, in terms of percentage false positives and false negatives. These results arise from considering an unordered full set of true samples with all 20 features.

Unordered and ordered sets of true samples yielded similar MCR values for all participants (p=0.05; Wilcoxon rank sum) except for participant 11 (p=0.0014) where the unordered set yielded a lower MCR value.

The effect of reducing the sample size on mean MRC is illustrated...
in Figure 8. This analysis was performed separately for each participant and the average MCRs across participants are reported in the figure with respect to the number of samples in each fold. Reducing the sample size did not significantly increase the error rate (p=0.125; robust regression test).

Discussion

This study presents the first investigation into the variability of time-evolving grip kinetic profiles of 20 adult participants writing the same, well-practiced signature over multiple days and at various times each day. The signatures studied herein can be considered free-hand or skilled forgeries given the extended practice period [36]. Overall, our findings corroborate previous reports that grip kinetics, albeit analyzed from a topographic perspective, generally do not differ significantly between well-practiced and authentic signatures [37].

A closer examination of Figure 5 reveals some common kinetic fluctuations across participants. Generally, the grip force gradually increases while writing the first name and sharply declines during the completion of the second ‘T’. The notable kinetic dip that follows is due the cessation in writing as the writer moves the pen horizontally to commence the last name. Similar to the first name, the grip force gradually increases as the word is written and tails off sharply with the writing of the last letter.

Despite these general similarities, the mean total force signals varied among writers in terms of their magnitude, their difference between words and their fluctuations within each word. Similar interindividual differences have been reported in the study of grip kinetics associated with golf swings [14]. The between-participant variability in the grip kinetic profile is likely attributable to the personalized nature of handwriting motor skills [38]. Indeed, early study of handwriting kinetics [39] qualitatively reported between-participant variability in finger pressure patterns.

Within-participant variations in the total force profile did exist as well, as exemplified in Figure 4. This variation included a change in the magnitude and shape of the profile over time. Some of the variability might be due to grasp adjustments and pen rotations, which were observed by the researcher during the data collection and retrospectively, through a review of the pictures taken at each session. These changes were particularly evident for participant 6 and explain the high MCR values for this participant in Figure 7. Grasp adjustments change the position and orientation of the sensor array with respect to the hand. Since there is some inevitable ‘dead space’ between neighbouring sensors on the array, grasp adjustments may alter the measurement of total force. Other contributors to within-participant force profile variation may have included circadian fluctuations in the grip strength and writer motivation [38], calibration errors, mental and physical fatigue, as well as changes in body and arm posture while writing [40].

Although all 32 sensors were considered in this study, future research may consider the potential to classify writers on the basis of a smaller subset of sensors. In a previous study [1], it was noted that handwriting grip forces are captured by only a small subset of sensors around the barrel. Future research may thus consider the judicious elimination of uninformative sensors.

In this study, the bogus signature is considered a text-based signature since each letter can be identified. A visual inspection of the authentic signatures of all 20 participants revealed that all but one participant employed a text-based form of signature writing; participant 7 was the only one who had a non-text signature where none of the letters could be identified. However, the MCR of participant 7 was still within the range of MCRs obtained across participants as evident in Figure 7. In a future study, it would be interesting to examine if the present findings would generalize to non-text writing more broadly.

The classification analysis suggested that even in the presence of within-participant variability in the grip kinetic profiles, the between-participant variations tend to be greater, allowing for inter-participant separation. The low MCR with the LDA classifier suggests that the features derived from the grip kinetic profiles are in fact linearly separable. The lack of difference in MCRs for ordered and unordered samples implies that this separability is consistent over time. Finally, the inter-participant separation seemed to be intact even if the pool of signatures was reduced dramatically, an important consideration for signature verification systems [36,41]. Overall, the findings of this study support further investigation of grip kinetic profiles associated with signature writing as a biometric measure.

Conclusion

In this study, we examined the total grip force profile generated while writing multiple iterations of a well-practiced signature by 20 participants. The algorithmic discrimination between writers based on features extracted from the total grip force profile indicated that despite intra-participant variability, each participant had a unique grip kinetic profile. Further, classification performance was robust to reductions in sample size and the temporal ordering of test signatures. Collectively, these findings indicate that the grip kinetic profile may be a valuable measure for signature verification applications.

Acknowledgement

The authors would like to acknowledge the Canada Research Chairs program, Syngrafii Inc., and the Natural Sciences and Engineering Research Council of Canada for supporting this research financially. The authors would also like to acknowledge Ms. Nayanashri Thalanki Anantha, Ms. Jennifer Chan, Ms. Laura Bell, and Ms. Sarah Stoops for their assistance with the data collection for this study.

References