

# A categorisation and implementation of digital pen features for behaviour characterisation

Alexander Prange      Michael Barz      Daniel Sonntag

DFKI

## Abstract

In this paper we provide a categorisation and implementation of digital ink features for behaviour characterisation. Based on four feature sets taken from literature, we provide a categorisation in different classes of syntactic and semantic features. We implemented a publicly available framework to calculate these features and show its deployment in the use case of analysing cognitive assessments performed using a digital pen.

## 1 Introduction

The research described in this paper is motivated by the development of applications for the behaviour analysis of handwriting and sketch input. Our goal is to provide other researchers with a reproducible, categorised set of features that can be used for behaviour characterisation in different scenarios. We use the term feature to describe properties of strokes and gestures which can be calculated based on the raw sensor input from capture devices, such as digital pens or tablets.

In this paper, a large number of features known from the literature are presented and categorised into different subsets. For better understanding and reproducibility we formalised all features either using mathematical notations or pseudo code and summarised them in the appendix section of this paper. Furthermore, we created a open-source python reference implementation of these features, which is publicly available<sup>1</sup>.

The presented ink features can be used in a variety of ways. Most commonly they are used to perform character and gesture recognition based on machine learning techniques. Here we describe their use for automated behaviour characterisation in the use case of cognitive assessments. Traditionally these tests are performed using pen and paper with manual evaluation by the therapist. We show how ink features can be used in that context to provide additional feedback about the cognitive state of the patient. Finally, we explain how digital ink can be used as an input modality in multimodal, multisensor interfaces.

---

<sup>1</sup>Download is available at GitHub <https://github.com/DFKI-Interactive-Machine-Learning/ink-features>

## 2 Digital Ink

Over the past few years the availability of digital pen hardware has increased drastically, and there is a wide variety of devices to choose from if dealing with handwriting analysis. Several different technologies are used to record handwriting, e.g., accelerometer-based digital pens convert the movement of the pen on the surface whereas active pens transmit their location, pressure and other functionalities to the built-in digitiser of the underlying device. Positional pens, most often encountered in graphic tablets, have a surface that is sensitive to the pen tip. A special, nearly invisible, dot pattern can be printed on regular paper, so that camera-based pens detect where the stylus contacts the writing surface.

In this work we focus on the similarities between the most commonly used hardware devices for sketch recognition. As not all technologies deliver the same type of sensor data, we identified a subset that is covered by the majority of input devices. We refer to it as digital ink, a set of time-series data containing coordinates and pressure at each timestamp. For the remainder of this paper we use the following notation:

$$x, y : \text{coordinates} \tag{1}$$

$$p : \text{pressures} \tag{2}$$

$$t : \text{timestamps} \tag{3}$$

A series  $S$  of  $n$  sample points between a pen down and pen up event is called a stroke and can be represented as a series of tuples

$$S = (x_0, y_0, p_0, t_0), (x_1, y_1, p_1, t_1), \dots, (x_{n-1}, y_{n-1}, p_{n-1}, t_{n-1}) \tag{4}$$

where  $x_i$  represents the x coordinate of the  $i$ -th sample point within the series, with  $0 \leq i < n$ . The tuple itself may be referenced by  $s_i$ . Timestamps are measured in milliseconds, it is insignificant if they are absolute or relative to the first point.

## 3 Features

We refer to individual, measurable properties or characteristics of digital ink as features. Features are calculated directly from the input sample points and represented by a numerical value. Therefore a feature can be seen as a function:

$$f : S \mapsto \mathbb{R} \tag{5}$$

Depending on the feature,  $S$  can be a set of strokes (gesture level), a single stroke (stroke level) or a subset of sample points. Usually a vector of features

$$F = [f_1, \dots, f_m] \tag{6}$$

is extracted from the input gesture and can then be used in a classifier.

### 3.1 Feature Sets

Traditionally stroke level features are most often used for statistical gesture recognition. One of the most prominent set of features was presented by Dean Rubine in 1991 [35]. It contains a total of 13 features that have been designed to reflect the visual appearance of strokes in order to be used in a gesture recogniser. More recent work by Don J.M. Willems and Ralph Niels [42] defines a total of 89 features using formal mathematical descriptions and algorithms. Adrien Delaye and Eric Anquetil introduced the HBF49 Feature Set [11], which contains 49 features and was specifically designed for different sets of symbols and as reference for evaluating symbol recognition systems. In previous work we used 14 features described by Sonntag et al. [38] to distinguish between written text and other types of gestures in online handwriting recognition.

#### 3.1.1 Common Features

Due to the nature of sketched or handwritten input there are a few features and concepts that the above mentioned publications have in common. The most prominent example is the length of a stroke, here we use the Euclidean distance to measure the distance between sampling points.

Given two sampling points  $q = (x_q, y_q)$  and  $r = (x_r, y_r)$  their distance is calculated as follows:

$$\|qr\| = \|r - q\| = \sqrt{(x_r - x_q)^2 + (y_r - y_q)^2} \quad (7)$$

The length of a stroke (a sequence of sampling points) is given by the sum of distances between the sampling points:

$$f_{StrokeLength} = \sum_{i=1}^{n-1} \|s_i - s_{i-1}\| \quad (8)$$

A bounding box (see figure 1) around a set of strokes describes the smallest enclosing area containing the entire set of points. Its size is determined by the minimum and maximum sample points:

$$x_{min} = \min_{0 \leq i < n} x_i \quad (9)$$

$$y_{min} = \min_{0 \leq i < n} y_i \quad (10)$$

$$x_{max} = \max_{0 \leq i < n} x_i \quad (11)$$

$$y_{max} = \max_{0 \leq i < n} y_i \quad (12)$$

The area of the bounding box is then given by:

$$f_{BoundingBoxArea} = (x_{max} - x_{min}) \cdot (y_{max} - y_{min}) \quad (13)$$

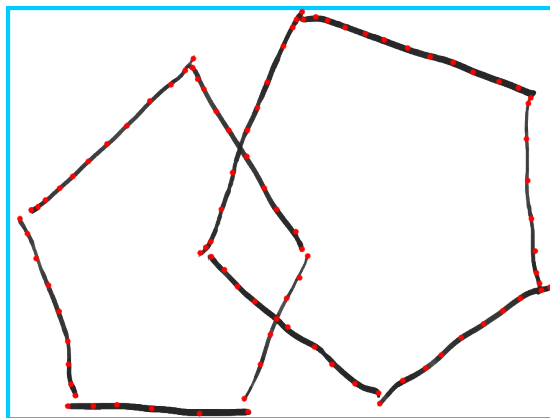


Figure 1: The rectangular bounding box (cyan) around a set of strokes (black) given by a set of sample points (red).

## 3.2 Feature Categories

We have chosen the above described sets of features which are formalised in a reproducible way. As the features describe different aspects of the digital ink we decided to sort them into different categories. We distinguish each feature to be either a syntactic or semantic feature. Syntactic features reflect task independent characteristics about the geometry of the input, whereas semantic features describe closely task related knowledge. In this work we introduce 7 categories of syntactic features:

### 3.2.1 Angle Based

Angle based features are calculated from angles between sample points (e.g., curvature, perpendicularity, rectangularity).

$$\theta_i = \arccos \left\{ \frac{\overrightarrow{s_{i-1}s_i} \cdot \overrightarrow{s_i s_{i+1}}}{\|\overrightarrow{s_{i-1}s_i}\| \cdot \|\overrightarrow{s_i s_{i+1}}\|} \right\} \quad (14)$$

$$f_{Curvature} = \sum_{i=1}^{n-2} \theta_i \quad (15)$$

$$f_{Perpendicularity} = \sum_{i=1}^{n-2} \sin^2(\theta_i) \quad (16)$$

### 3.2.2 Space Based

Space based features depend on the distances between samples (e.g., convex hull area, principal axes, compactness). The area  $A$  of a gesture is usually

derived from the area of the convex hull around all sample points, which can be calculated using Graham’s algorithm [18].

$$f_{ConvexHullArea} = A \quad (17)$$

With the area of the convex hull and the length of its perimeter  $l$  we get a feature called compactness. The closer the sample points are together, the smaller the compactness will be. Handwritten texts, e.g., will have a larger compactness than geometric symbols, such as rectangles [42].

$$f_{Compactness} = \frac{l^2}{A} \quad (18)$$

Related to the bounding box of a figure, we use its side length to calculate the eccentricity. Note that we are using the co-ordinate axes instead of the principal axes (which are rotated with the pen gesture).

$$f_{Eccentricity} = \sqrt{1 - \frac{b^2}{a^2}} \quad (19)$$

### 3.2.3 Centroidal

Centroidal features describe relations between sample points and the overall centroid (e.g., centroid offset, deviation, average radius).

Using the dimensions of the bounding box we calculate the center point  $c$ :

$$c = \begin{pmatrix} x_{center} \\ y_{center} \end{pmatrix} = \begin{pmatrix} x_{min} + 0.5 \cdot (x_{max} - x_{min}) \\ y_{min} + 0.5 \cdot (y_{max} - y_{min}) \end{pmatrix} \quad (20)$$

The average distance of sample points to the center point is another feature:

$$f_{MeanCentroidDistance} = \frac{1}{n} \sum_{i=0}^{n-1} \|s_i - \mathbf{c}\| \quad (21)$$

### 3.2.4 Temporal

Temporal features are derived from timestamps of sample points (e.g., duration, speed, acceleration). The velocity  $v$  between sample points is defined as:

$$v_i = \frac{s_{i+1} - s_{i-1}}{t_{i+1} - t_{i-1}} \quad (22)$$

From which the feature of average velocity is calculated:

$$f_{AverageVelocity} = \frac{1}{n-2} \sum_{i=1}^{n-2} \|v_i\| \quad (23)$$

The acceleration is calculated as follows:

$$a_i = \frac{v_{i+1} - v_{i-1}}{t_{i+1} - t_{i-1}} \quad (24)$$

And the average acceleration is then given by:

$$f_{AverageAcceleration} = \frac{1}{n-4} \sum_{i=2}^{n-3} \|\mathbf{a}_i\| \quad (25)$$

### 3.2.5 Pressure Based

Pressure based features are computed from hardware sensors capturing applied pressure (e.g., average pressure, standard deviation). The most intuitive and obvious features are the average pressure and the standard deviation in pressure:

$$f_{AveragePressure} = \frac{1}{n} \sum_{i=0}^{n-1} p_i \quad (26)$$

$$f_{StandardPressureDeviation} = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (p_i - f_{AveragePressure})^2} \quad (27)$$

### 3.2.6 Trajectory Based

Trajectory based features reflect the visual appearance of strokes (e.g., closure, average stroke direction).

The path length from one sample point to another is denoted  $L$  and is calculated as follows:

$$L_{i,j} = \sum_{k=i}^{j-1} \begin{cases} 0 \\ \|s_k s_{k+1}\| \end{cases} \quad (28)$$

$L = L_{0,n-1}$  is the total length of  $S$ . Whereas the first to last point vector and its length is:

$$v = \overrightarrow{s_1 s_n}, \quad \|v\| = \|s_1 s_n\| \quad (29)$$

Typical trajectory based features are closure and average direction:

$$f_{Closure} = \frac{\|v\|}{L} \quad (30)$$

$$f_{AverageDirection} = \frac{1}{n-1} \sum_{i=0}^{n-2} \arctan\left(\frac{y_{i+1} - y_i}{x_{i+1} - x_i}\right) \quad (31)$$

### 3.2.7 Meta

Meta features are higher level features and relations between components (e.g., number of strokes, inter-connections, crossings, straight line ratio). One intuitive example would be the number of straight lines ( $f_{\#StraightLines}$ ) or to be more precise the number of straight segments. We use a sliding window with a threshold to calculate sets of connected points which have minimal curvature

between them. The size sliding window and threshold can be either dynamically adjusted to the length of the stroke or be a fixed value depending on the task.

The feature called *connected components* ( $f_{ConnectedComponents}$ ) [42] describes the number of segments which are interconnected with other segments, e.g., crossings between strokes.

Geometric Features	
Angle Based	Space Based
Circular Variance	Stroke Length
Rectangularity	Gesture Length
Curvature	Perimeter Length
Average Curvature	Compactness
SD of Curvature	Eccentricity
Angles after Resampling	Principal Axes
Cosine of First to Last Point Vector	First Point X
Sine of First to Last Point Vector	First Point Y
Cosine Initial Vector	Last Point X
Sine Initial Vector	Last Point Y
Bounding Box Diagonale Angle	First to Last Point Vector
Perpendicularity	2D Histogramm
Average Perpendicularity	Ratio of Axes
SD of Perpendicularity	Ratio of Principal Axes
Signed Perpendicularity	Length of First Principal Axis
K-Perpendicularity	SD of Stroke Length
Maximum k-Angle	Sample Ratio Octants
Absolute Directional Angle	Convex Hull Area
Relative Angle Histogram	Convex Hull Compactness
Principal Axis Orientation (sin)	Distance of First to Last Point
Principal Axis Orientation (cos)	Average Length of Straight Lines
Maximum Angular Difference	Initial Horizontal Offset
Circular Variance	Final Horizontal Offset
Sum of Absolute Values of Angles	Initial Vertical Offset
Sum of Angles	Final Vertical Offset
Sum of Squared Angles	
Macro Perpendicularity	
Average Macro Perpendicularity	
SD of Macro Perpendicularity	
Absolute Curvature	
Squared Curvature	
Centroidal	
Deviation	
Centroid Offset	
Average Centroidal Radius	
SD of Centroidal Radius	
Hu moments	

Temporal Features
Maximum Speed (Squared) Duration of Gesture Pen Up/Pen Down Ratio Average Velocity SD of Velocity Maximum Velocity Average Acceleration SD of Acceleration Maximum Acceleration Maximum Deceleration
Pressure Based Features
Average Pressure SD of Pressure
Trajectory Features
Closure Inflexion X Inflexion Y Proportion of Downstroke Trajectory Ratio between Half-Perimeter and Trajectory Average Stroke Direction Cup Count Last Cup Offset First Cup Offset Number of Pen Down Events Sin Chain Code Cos Chain Code SD of Stroke Direction
Meta Features
Number of Strokes Number of Straight Lines SD of Straight Lines Straight Line Ratio Largest Straight Line Ratio Number of Connected Components Number of Crossings

Table 1: Categorisation of syntactic features into classes.



### 3.3 Semantic/Task Based Features

Depending on the task, additional features can be deduced from the task itself. As these features describe higher level semantic concepts about the sketched contents, we often refer to them as semantic features. Semantic features highly depend on the given context and therefore vary noticeably between different tasks. Such features usually cannot be transferred easily to other tasks, as they are often hard-coded per task.

Figure 2 shows the visualisation of a selected semantic feature set in the context of the Clock Drawing Test, a widely used pen and paper screening test used for more than 50 years as a screening tool for cognitive impairment. Participants are asked to draw a clock face with the time set to 10 past 11 o'clock. The drawn clock is then examined by a trained physician and rated based on a predefined scoring scheme, reflecting the visual appearance and integrity of the clock using a numerical score. In this example we deduced the following features based on the traditional scoring system:

- $c$  denotes the center point of the clock (centroid), the closer it is to the center of the clock's circle, the more points are awarded.
- $L_h$  and  $L_m$  represent the lengths of the hour and minute hands respectively. If the clock is well drawn, the hour hand should be shorter than the minute hand.
- The angle between the hour and minute hands is denoted as  $\alpha$ , together with the orientation of the hands it can be used to determine if the correct time was set.
- $\Delta_9$  is the displacement of clock face digits relative to their ideal location. In this example it is the vertical offset of digit number 9 to its correct center position.

## 4 Related Work

One of the first reproducible ink feature sets was presented by Rubine [35] in 1991. He described how to use these features in a trainable single-stroke recogniser for gestures. Willems and Niels [42] presented a set of 89 ink features which they used for iconic and multi-stroke gesture recognition. The HBF49 feature set was presented by Delaye and Anquetil [11] to be used in online symbol recognition. Sonntag et al. [38] used ink features to distinguish between writing and sketching in online mode detection of handwriting input.

Ink features can not only be used for gesture or sketch recognition, but also for characterisation of handwriting behaviour. Drotar et al. [13] have shown that the analysis of in-air movement can be used as a marker for Parkinson's disease. The kinematic analysis of handwriting movements can be used to distinguish between certain forms of dementia [36].

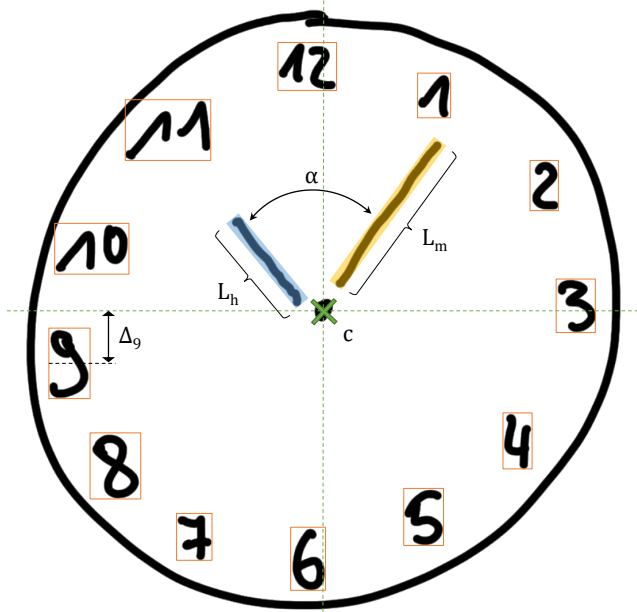


Figure 2: Visualisation of semantic features in the context of the Clock Drawing Test.

Digitalising popular existing cognitive assessments, such as the Clock Drawing Test (CDT), has been topic of recent debate. There are clear benefits resulting from digitalisation, such as increased diagnostic accuracy [29]. Davis et al. only recently presented their work on how to infer cognitive status from subtle behaviours observed in digital ink [10]. Based on such ink features machine learning models can be trained [39], which can also be explained by existing, validated scoring schemes [40]. Examples of complex digitalised cognitive assessments include the Rey-Osterrieth Complex Figure test [7], which can be used for various purposes, such as diagnosing the periphery [8].

Behaviour characterisation can be also used in different settings, e.g., to gain feedback about cognitive load of the writer. Luria and Rosenblum [27] conducted a study to determine the effect of mental workload on handwriting behaviour. Yu et al. [46] showed that online writing features can be used for mental workload classification, such as cognitive load evaluation [45]. Ink features can be also used in multimodal scenarios [33], where they may enhance the prediction of cognitive and emotional states [47].

## 5 Use Case: Cognitive Assessments

On use case where we apply our feature set is the analysis of handwriting behaviour for dementia screening tools in the Interakt project [37]. Dementia is a

name	time needed	pen input	symbols
AKT [17]	15 min	100%	cross-out
CDT [16]	2-5 min	100%	clock, digits, lines
CERAD [28]	30-45 min	20%	(see figure 6)
DemTect [23]	6-8 min	20%	numbers, words
MMSE [15]	5-10 min	9%	pentagrams
MoCA [30]	10 min	17%	clock, digits, lines
ROCF [7]	15 min	100%	circles, rectangles, triangles, lines
TMT [34]	3-5 min	100%	lines

Table 2: Comparison of the most widely used cognitive assessments

general term for a decline in mental ability severe enough to interfere with daily life. In 2018, the Alzheimer’s Association documented that approximately 10-20% of the population over 65 years of age suffer from some form of dementia [2]. Screening tests for dementia have been the subject of recent debate because there are limitations when they are conducted using pen and paper. For example, the collected material is monomodal (written form) and there is no direct digitalisation for further and automatic processing, the results can be biased. We selected the assessments based on feedback from domain experts and a recent market analysis of existing, most widely used, cognitive assessments conducted by Niemann et al. [31]. Our selected and implemented paper and pencil tests are shown in table 2, namely Age-Concentration (AKT) [17], Clock Drawing Test (CDT) [16], CERAD Neuropsychological Battery [28], Dementia Detection (DemTect) [23], Mini-Mental State Examination (MMSE) [15], Montreal Cognitive Assessment (MoCA) [30], Rey-Osterrieth Complex Figure (ROCF) [7], and Trail Making Test (TMT) [34]. The selection of the tests accounts for a variety of patient populations and test contexts.

One of the most prominent example is the internationally used Mini-Mental State Examination (MMSE) [15], a 30-point questionnaire, which is extensively used in medicine and research to measure cognitive impairment. Depending on the experience of the physician and the cognitive state of the patient the administration of the test takes between 5 and 10 minutes and examines functions including awareness, attention, recall, language, ability to follow simple commands and orientation [41]. Due to its standardisation, validity, short administration period and ease of use, it is widely used as a reliable screening tool for dementia [19]. The MMSE also includes several tasks which involve handwriting input by the participant, e.g., writing a complete sentence and copying a geometric figure.

The Clock Drawing Test (CDT) [16] is another popular cognitive assessment, where the patient is asked to draw a clock with a specified time on a piece of paper, see figure 3. Based on the completeness and appearance of the clockface and the arrangement of the digits a score is calculated. The CDT and MMSE are perfect examples for illustrating the two categories of traditional

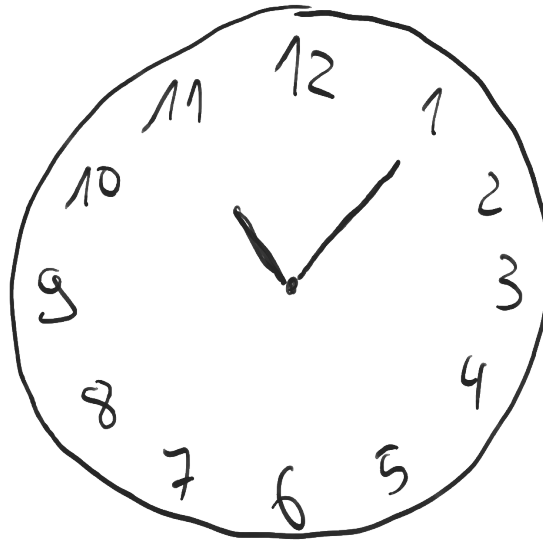


Figure 3: Clock Drawing Test (CDT).

paper and pencil cognitive testing. There are assessments, like the CDT, which rely solely on handwriting and sketch input to produce a score, whereas there are others, such as the MMSE, which also include other modalities, such as speech for instance. Depending on the assessment, the handwriting input has a different weight for the overall scoring of the test. Table 2 shows the absolute percentages of the test questions where the a pen is used to answer them. Tasks in the MMSE containing pen input include writing a sentence and copying a figure of two overlapping pentagrams (see figure 4). Out of 22 possible points in the scoring of the MMSE, the pen input related task add up to 2 points, resulting in an overall 9% of the entire test to be scored through analysis of pen input. Regarding task design the Montreal Cognitive Assessment (MoCA) [30] is comparable with the MMSE and CDT, e.g., it also includes copying a figure and drawing a clock. The CERAD Neuropsychological Battery [28] is a collection of several tests (including the MMSE and TMT), where amongst others the subject has to copy several shapes depicted in figure 6. In the Trail Making Test (TMT) [34] the subject has to connect numbers and letters in ascending order. A more complex example of a test that is rated entirely based on pen input is the Rey-Osterrieth complex figure test (ROCF) [7], where subjects are required to copy the figure three times, once while looking at the template, once directly after that, but without seeing the template, and once from recall 30 minutes later. The Age-Concentration Test (AKT) [17] asks subjects to cross out a specific shape from a set of similar, yet varying shapes in a limited amount of time. Handwritten words and digits are contained in the DemTect [23], where subjects translate numbers into words and vice versa.

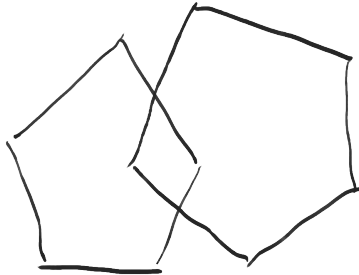


Figure 4: Mini Mental Status Exam (MMSE): Copy pentagram figure task.

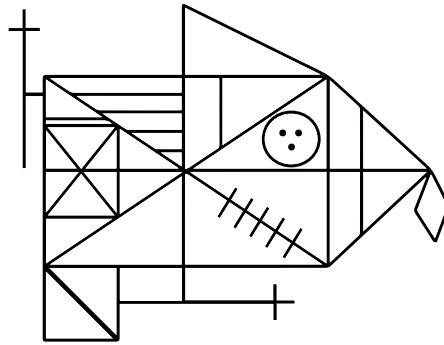


Figure 5: The Rey-Osterrieth complex figure (ROCF).

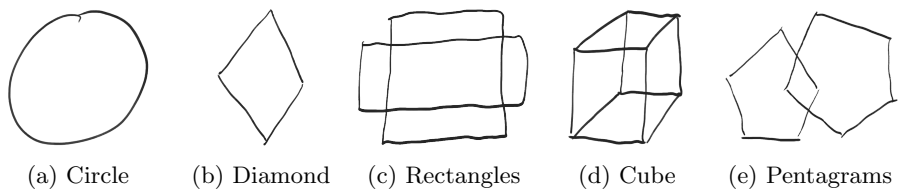


Figure 6: Symbols used in the CERAD neuropsychological battery.

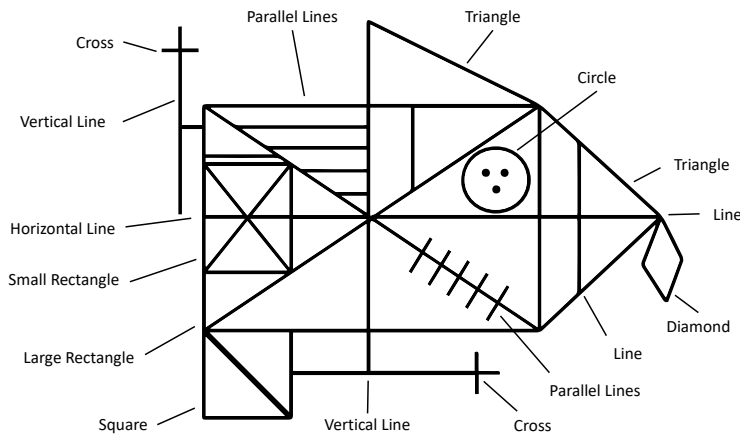


Figure 7: The Rey-Osterrieth complex figure (ROCF) is composed of several sub-shapes.

## 5.1 Symbols data set

Based on the design of sketching tasks in cognitive testing, we created a set of 11 gestures, which are commonly found in different cognitive assessments. We focused on the geometric shapes of which tests are composed, e.g., the Clock Drawing Test contains a circle (clockface) and lines (hands). The CERAD battery, MMSE and MoCA contain several shapes like pentagrams, diamonds and rectangles. Single shapes in turn compose parts of other assessments, such as the ROCF depicted in figure 7, which contains several sub-shapes, such as, triangles, rectangles, lines and circles. As depicted in figure 8 a total of 8 shapes were chosen from the most commonly used cognitive assessments: arrow, circle, rectangle, triangle, circle, diamond, overlapping rectangles, cube and pentagrams. We chose 3 additional gestures based on a previously conducted user study, where we asked participants to specify gestures that they would use to indicate that they are finished with the current handwriting task. Our symbol data set consists of 11 classes (shapes) with 100 samples per class. The 7 subjects have provided a total of 7700 handwritten samples.

## 5.2 Interakt Architecture

In the Interakt use case the patient performs a digitalised cognitive assessment using a digital pen, which captures handwriting data in real-time. Figure ?? shows the technical architecture, in which the digital ink is analysed using the previously described syntactic and semantic features. Completing the cognitive assessment results in raw pen data being streamed to the backend service, where a document is created and indexed based on the performed test. This document contains semantic information about the areas of the test (e.g., text fields, figures etc.) and the digital ink data. We store the documents in a file format called

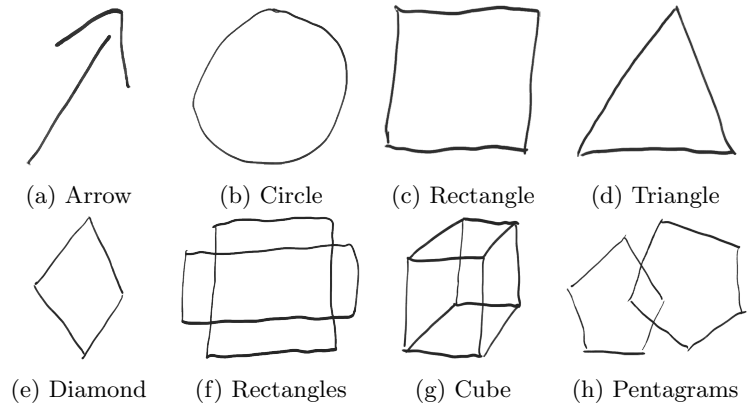


Figure 8: Set of gestures chosen from cognitive assessments.

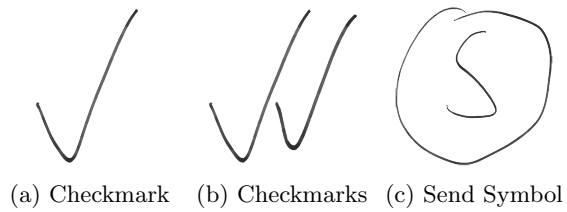


Figure 9: Set of symbols.

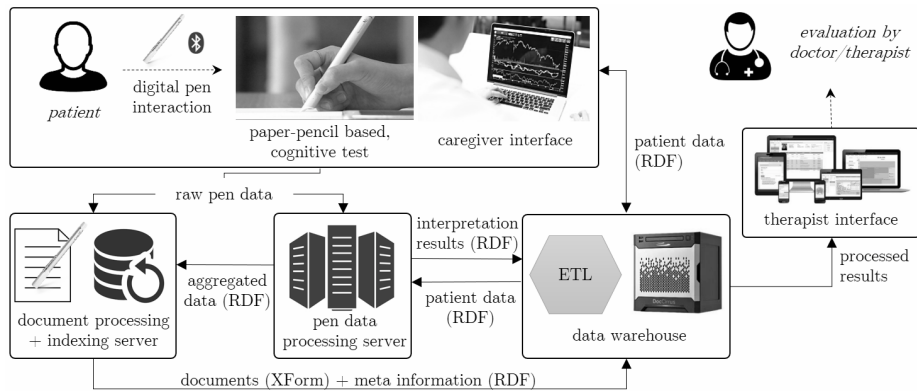


Figure 10: System Architecture in the dementia screening use case.

*XForm*, which is either a JSON or XML based structured description of the test and the captured ink. With this format a visual representation of the completed test can be reconstructed and the doctor can retrace the patient’s input using a playback functionality that replays the strokes in real-time as they were recorded. Based on the respective assessment different sets of syntactic and semantic features are used by the *pen data processing server* to analyse the handwritten and sketched contents of the test and deliver aggregated evaluation results that can be presented to the therapist. Depending on the situation the analysis of the assessment may also involve additional patient data or previous test results, which are obtained from the *data warehouse*. The processed and evaluated assessment is finally also stored in the data warehouse, from where the doctor can access the results of the assessment in the *therapist interface*. The entire evaluation process takes place in real-time.

## 6 Multimodality

In this section, we describe how additional modalities, beyond pen-based features, can help in the analysis of observed user behaviour, when interacting with a tablet computer and relying on the built-in sensors only. For instance, researchers in the medical domain investigated “observable differences in the communicative behaviour of patients with specific psychological disorders” [12], e.g., the detection of depression from facial actions and vocal prosody [9], which can be realised using the camera and microphone of a tablet device. Including additional modalities can help with the disambiguation of signal- or semantic-level information in one error-prone recognition modality by using partial information supplied by another modality [32]. We consider the digital pen signal as primary modality for behaviour characterisation in combination with additional sensors and modalities as indicated in Figure 11: eye tracking and facial expression analysis via the video signal of the front-facing camera, natural speech



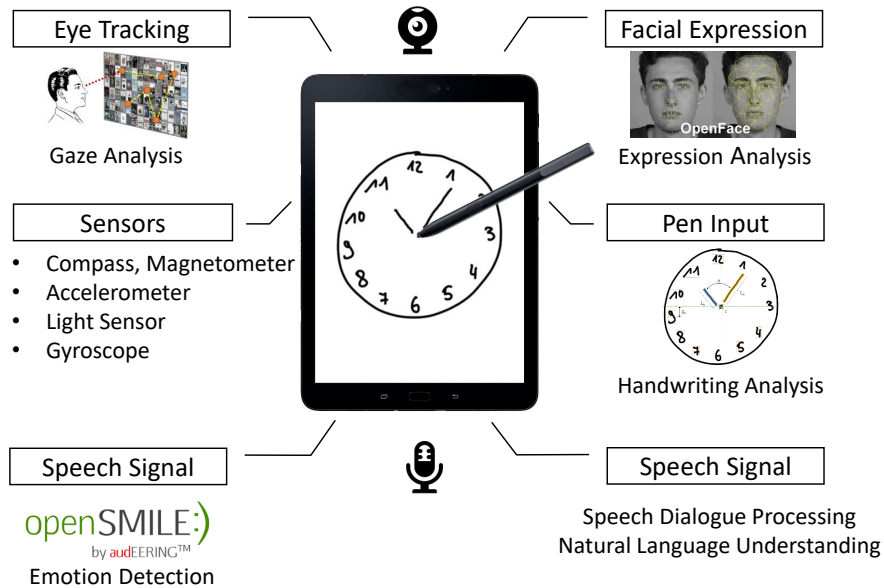


Figure 11: Multimodal interaction architecture on mobile device.

processing via the built-in microphone and additional sensor inputs of modern tablet devices.

## 6.1 Eye Tracking

Eye tracking can improve human behaviour analysis, because human gaze is related to cognitive processes. For instance, gaze trajectories can be used for inferring a user's task [44], for differentiating between novices and experts [5] and to model human visual attention [6]. Further, the number and duration of fixations and the transitions between different contents provide information about a user's cognitive engagement [26] and its cognitive load [24]. To augment pen signals, it is interesting that users pro-actively control their gaze behaviour to gather visual information for guiding movements across different activities [22] including hand movements [25]. This relation suggests that pen and gaze signals can be analysed jointly for improving behaviour characterisation. For multisensory behaviour analysis on unmodified mobile devices, RGB-based eye tracking is most interesting, because it's deployable using the built-in front-facing camera [43]. Compared to professional eye tracking equipment, the tracking quality is significantly lower [4]. However, the form-factor can be essential for certain use cases, e.g., in dementia day hospitals that require non-obtrusive devices due to

the patient’s cognitive abilities.

## 6.2 Facial Expressions

OpenFace<sup>2</sup> [3] is an open source toolkit for facial behaviour analysis using the stream of an RGB-webcam. It provides state-of-the-art performance in facial landmark and head pose tracking, as well as facial action unit recognition which can be used to infer emotions. These observations can be used for affective user interaction. Further, it enables webcam-based eye tracking.

## 6.3 Speech Signal

The openSMILE toolkit<sup>3</sup> [14] provides methods for speech-based behaviour analysis and is distributed under an open source license. It offers an API for low-level feature extraction from audio signals and pre-trained classifiers for voice activity detection, speech-segment detection and speech-based emotion recognition in real-time. The toolkit can be used on top of speech-based interaction frameworks to add a valence to users’ utterances.

## Acknowledgement

This research is part of the Intera-KT project, which is supported by the German Federal Ministry of Education and Research (BMBF) under grant number 16SV7768.

## References

- [1] ALMAZ’N, J., FORNES, A., AND VALVENY, E. A non-rigid feature extraction method for shape recognition. In *2011 International Conference on Document Analysis and Recognition* (Sept 2011), pp. 987–991.
- [2] ALZHEIMER’S ASSOCIATION. 2018 Alzheimer’s disease facts and figures. *Alzheimer’s & Dementia: The Journal of the Alzheimer’s Association* 14, 3 (Mar 2018), 367–429.
- [3] BALTRUSAITIS, T., ZADEH, A., LIM, Y. C., AND MORENCY, L. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)* (May 2018), pp. 59–66.
- [4] BARZ, M., POLLER, P., AND SONNTAG, D. Evaluating Remote and Head-worn Eye Trackers in Multi-modal Speech-based HRI. In *Proceedings of the*

---

<sup>2</sup><https://github.com/TadasBaltrusaitis/OpenFace/>

<sup>3</sup><https://audeering.com/technology/opensmile/>

- Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (New York, NY, USA, 2017), B. Mutlu, M. Tscheligi, A. Weiss, and J. E. Young, Eds., ACM, pp. 79–80.
- [5] BEDNARIK, R. Expertise-dependent visual attention strategies develop over time during debugging with multiple code representations. *Int. J. Hum.-Comput. Stud.* 70, 2 (Feb. 2012), 143–155.
- [6] BORJI, A., AND ITTI, L. State-of-the-Art in Visual Attention Modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35, 1 (jan 2013), 185–207.
- [7] CANHAM, R., SMITH, S., AND TYRRELL, A. *Automated scoring of a neuropsychological test: The Rey Osterrieth Complex Figure*. IEEE COMPUTER SOC, 2000, pp. A406–A413.
- [8] COATES, D. R., WAGEMANS, J., AND SAYIM, B. Diagnosing the periphery: Using the rey–osterrieth complex figure drawing test to characterize peripheral visual function. In *i-Perception* (2017).
- [9] COHN, J. F., KRUEZ, T. S., MATTHEWS, I., YANG, Y., NGUYEN, M. H., PADILLA, M. T., ZHOU, F., AND LA TORRE, F. D. Detecting depression from facial actions and vocal prosody. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops* (Sept 2009), pp. 1–7.
- [10] DAVIS, R., LIBON, D. J., AU, R., PITMAN, D., AND PENNEY, D. L. Think: Inferring cognitive status from subtle behaviors. *AI Magazine* 36, 3 (2015), 49–60.
- [11] DELAYE, A., AND ANQUETIL, E. HBF49 feature set: A first unified baseline for online symbol recognition. *Pattern Recognition* 46, 1 (Jan. 2013), 117–130.
- [12] DEVULT, D., ARTSTEIN, R., BENN, G., DEY, T., FAST, E., GAINER, A., GEORGILA, K., GRATCH, J., HARTHOLT, A., LHOMMET, M., LUCAS, G., MARSELLA, S., MORBINI, F., NAZARIAN, A., SCHERER, S., STRATOU, G., SURI, A., TRAUM, D., WOOD, R., XU, Y., RIZZO, A., AND MORENCY, L.-P. SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support. In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems* (Richland, SC, 2014), AAMAS '14, International Foundation for Autonomous Agents and Multiagent Systems, pp. 1061–1068.
- [13] DROTÁR, P., MEKYSKA, J., REKTOROVÁ, I., MASAROVÁ, L., SMÉKAL, Z., AND FAUNDEZ-ZANUY, M. Analysis of in-air movement in handwriting: A novel marker for parkinson’s disease. *Computer Methods and Programs in Biomedicine* 117, 3 (2014), 405 – 411.

- [14] EYBEN, F., WENINGER, F., GROSS, F., AND SCHULLER, B. Recent developments in opensmile, the munich open-source multimedia feature extractor. In *Proceedings of the 21st ACM International Conference on Multimedia* (New York, NY, USA, 2013), MM '13, ACM, pp. 835–838.
- [15] FOLSTEIN, M., FOLSTEIN, S., AND MCHUGH, P. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 12, 3 (Nov 1975), 189–198.
- [16] FREEDMAN, M., LEACH, L., KAPLAN, E., WINOCUR, G., SHULMAN, K., AND DELIS, D. *Clock Drawing: A Neuropsychological Analysis*. Oxford University Press, 1994.
- [17] GATTERER, G., FISCHER, P., SIMANYI, M., AND DANIELCZYK, W. The A-K-T ("Alters-Konzentrations-Test") a new psychometric test for geriatric patients. *Funct. Neurol.* 4, 3 (1989), 273–276.
- [18] GRAHAM, R. L. An efficient algorithm for determining the convex hull of a finite planar set. *Inf. Process. Lett.* 1, 4 (1972), 132–133.
- [19] HARRELL, L. E., MARSON, D., CHATTERJEE, A., AND PARRISH, J. A. The severe mini-mental state examination: A new neuropsychologic instrument for the bedside assessment of severely impaired patients with alzheimer disease. *Alzheimer Disease & Associated Disorders* 14, 3 (2000).
- [20] HU, M.-K. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory* 8, 2 (February 1962), 179–187.
- [21] IMPEDOVO, S., PIRLO, G., MODUGNO, R., AND FERRANTE, A. Zoning methods for hand-written character recognition: An overview. In *2010 12th International Conference on Frontiers in Handwriting Recognition* (Nov 2010), pp. 329–334.
- [22] JOHANSSON, R. S., WESTLING, G., BÄCKSTRÖM, A., AND FLANAGAN, J. R. Eye–Hand Coordination in Object Manipulation. *Journal of Neuroscience* 21, 17 (2001), 6917–6932.
- [23] KALBE, E., KESSLER, J., CALABRESE, P., SMITH, R., PASSMORE, A. P., BRAND, M., AND BULLOCK, R. DemTect: a new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *International Journal of Geriatric Psychiatry* 19, 2 (Feb 2004), 136–143.
- [24] KORBACH, A., BRÜNKEN, R., AND PARK, B. Differentiating Different Types of Cognitive Load: a Comparison of Different Measures. *Educational Psychology Review* (mar 2017), 1–27.
- [25] LAND, M., MENNIE, N., AND RUSTED, J. The roles of vision and eye movements in the control of activities of daily living. *Perception* 28, 11 (1999), 1311–1328.

- [26] LEMAIGNAN, S., GARCIA, F., JACQ, A., AND DILLENBOURG, P. From Real-time Attention Assessment to "With-me-ness" in Human-Robot Interaction. *The Eleventh ACM/IEEE International Conference on Human Robot Interaction* (2016), 157–164.
- [27] LURIA, G., AND ROSENBLUM, S. A computerized multidimensional measurement of mental workload via handwriting analysis. *Behavior Research Methods* 44, 2 (Jun 2012), 575–586.
- [28] MORRIS, J., MOHS, R., ROGERS, H., FILLENBAUM, G., AND HEYMAN, A. Consortium to establish a registry for Alzheimer’s disease (CERAD) clinical and neuropsychological assessment of Alzheimer’s disease. *Psychopharmacol Bull.* 24, 4 (1988), 641–52.
- [29] MULLER, S., PREISCHE, O., HEYMANN, P., ELBING, U., AND LASKE, C. Increased Diagnostic Accuracy of Digital vs. Conventional Clock Drawing Test for Discrimination of Patients in the Early Course of Alzheimer’s Disease from Cognitively Healthy Individuals. *Front Aging Neurosci* 9 (2017), 101.
- [30] NASREDDINE, Z. S., PHILLIPS, N. A., ET AL. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 53, 4 (2005), 695–699.
- [31] NIEMANN, M., PRANGE, A., AND SONNTAG, D. Towards a Multimodal Multisensory Cognitive Assessment Framework. In *31st IEEE International Symposium on Computer-Based Medical Systems, CBMS 2018, Karlstad, Sweden, June 18-21, 2018* (2018), pp. 24–29.
- [32] OVIATT, S., AND COHEN, P. R. *The Paradigm Shift to Multimodality in Contemporary Computer Interfaces*. Morgan & Claypool Publishers, 2015.
- [33] OVIATT, S., SCHULLER, B., COHEN, P. R., SONNTAG, D., POTAMIANOS, G., AND KRÜGER, A., Eds. *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 1*, vol. Volume 1. Association for Computing Machinery and Morgan & Claypool, New York, NY, USA, 2017.
- [34] REITAN, R. *Trail Making Test*. Reitan Neuropsychology Laboratory, 1992.
- [35] RUBINE, D. Specifying gestures by example. In *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1991), SIGGRAPH ’91, ACM, pp. 329–337.
- [36] SCHROTER, A., MERGL, R., BURGER, K., HAMPEL, H., MOLLER, H. J., AND HEGERL, U. Kinematic analysis of handwriting movements in patients with Alzheimer’s disease, mild cognitive impairment, depression and healthy subjects. *Dement Geriatr Cogn Disord* 15, 3 (2003), 132–142.

- [37] SONNTAG, D. Interakt - A multimodal multisensory interactive cognitive assessment tool. *CoRR abs/1709.01796* (2017).
- [38] SONNTAG, D., WEBER, M., CAVALLARO, A., AND HAMMON, M. Integrating digital pens in breast imaging for instant knowledge acquisition. *AI Magazine* 35, 1 (2014), 26–37.
- [39] SOUILLARD-MANDAR, W., DAVIS, R., RUDIN, C., AU, R., LIBON, D. J., SWENSON, R., PRICE, C. C., LAMAR, M., AND PENNEY, D. L. Learning classification models of cognitive conditions from subtle behaviors in the digital clock drawing test. *Mach. Learn.* 102, 3 (Mar. 2016), 393–441.
- [40] SOUILLARD-MANDAR, W., DAVIS, R., RUDIN, C., AU, R., AND PENNEY, D. Interpretable Machine Learning Models for the Digital Clock Drawing Test. *ArXiv e-prints* (Jun 2016), arXiv:1606.07163.
- [41] TUIJL, J. P., SCHOLTE, E. M., CRAEN, A. J., AND MAST, R. C. Screening for cognitive impairment in older general hospital patients: comparison of the six-item cognitive impairment test with the mini-mental state examination. *International Journal of Geriatric Psychiatry* 27, 7 (2012), 755–762.
- [42] WILLEMS, D., AND NIELS, R. Definitions for features used in online pen gesture recognition. Tech. rep., NICI, Radboud University Nijmegen, 2008.
- [43] WOOD, E., AND BULLING, A. EyeTab: model-based gaze estimation on unmodified tablet computers. In *Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA '14* (New York, New York, USA, 2014), ACM Press, pp. 207–210.
- [44] YARBUS, A. L. Eye movements and vision. *Neuropsychologia* 6, 4 (1967), 222.
- [45] YU, K., EPPS, J., AND CHEN, F. Cognitive load evaluation of handwriting using stroke-level features. In *Proceedings of the 16th International Conference on Intelligent User Interfaces* (New York, NY, USA, 2011), IUI '11, ACM, pp. 423–426.
- [46] YU, K., EPPS, J., AND CHEN, F. Mental workload classification via online writing features. In *2013 12th International Conference on Document Analysis and Recognition* (Aug 2013), pp. 1110–1114.
- [47] YU, K., EPPS, J., AND CHEN, F. Q. Cognitive load measurement with pen orientation and pressure. In *MMCogEms: Inferring Cognitive and Emotional States from Multimodal Measures* (2011).

## A Sonntag/Weber Features

The features described in this section are implementations based on the 14 features described by Sonntag et al. [38]. For this section we use the following notations:

A stroke is a sequence  $S$  of samples,

$$S = \{\vec{s}_i | i \in [0, n - 1], t_i < t_{i+1}\} \quad (32)$$

where  $n$  is the number of recorded samples. A sequence of strokes is indicated by

$$\mathcal{S} = \{S_i | i \in [0, m - 1]\} \quad (33)$$

where  $m$  is the number of strokes.

The centroid is defined as

$$\vec{\mu} = \frac{1}{n} \sum_{i=0}^{n-1} \vec{S}_i \quad (34)$$

where  $n$  is the number of samples used for the classification, the mean radius (standard deviation) as

$$\mu_r = \frac{1}{n} \sum_{i=0}^{n-1} \|\vec{s}_i - \vec{\mu}\| \quad (35)$$

and the angle as

$$\varphi_{s_i} = \cos^{-1} \left\{ \frac{(s_i - s_{i-1}) \cdot (s_{i+1} - s_i)}{\|s_i - s_{i-1}\| \|s_{i+1} - s_i\|} \right\} \quad (36)$$

### A.1 Number of Strokes

$$f_1 = |\mathcal{S}| \quad (37)$$

### A.2 Length

$$f_2 = \sum_{i=0}^{n-2} \|s_i - s_{i+1}\| \quad (38)$$

### A.3 Area

The area covered by the sequence of strokes  $\mathcal{S}$  is defined as the area of the bounding box that results from a sequence of strokes. We calculate the area of the convex hull  $A$  based on Graham's algorithm[18]

$$f_3 = Area(Conv(\mathcal{S})) = A \quad (39)$$

#### A.4 Perimeter Length

The length of the path around the convex hull

$$f_4 = \|\text{Conv}(\mathcal{S})\| \quad (40)$$

#### A.5 Compactness

$$f_5 = \frac{\|\text{Conv}(\mathcal{S})\|^2}{A} \quad (41)$$

#### A.6 Eccentricity

Let  $a$  and  $b$  denote the length of the major or minor axis of the convex hull, respectively

$$f_6 = \sqrt{1 - \frac{b^2}{a^2}} \quad (42)$$

#### A.7 Principal Axes

$$f_7 = \frac{b}{a} \quad (43)$$

#### A.8 Circular Variance

Let  $\mu_r$  denote the mean distance of the samples to the centroid  $\mu$ . The circular variance is then computed as follows

$$f_8 = \frac{1}{n \cdot \mu_r^2} \sum_{i=0}^{n-1} (\|s_i - \mu\| - \mu_r)^2 \quad (44)$$

#### A.9 Rectangularity

$$f_9 = \frac{A}{a \cdot b} \quad (45)$$

#### A.10 Closure

$$f_{10} = \frac{\|s_0 - s_n\|}{f_4} \quad (46)$$

#### A.11 Curvature

Let  $\varphi(s_i)$  be the angle between the  $\overline{s_{i-1}s_i}$  and  $\overline{s_i s_{i+1}}$  segments at  $s_i$ .

$$f_{11} = \sum_{i=1}^{n-2} \varphi(s_i) \quad (47)$$



### A.12 Perpendicularity

$$f_{12} = \sum_{i=1}^{n-2} \sin(\varphi(s_i))^2 \quad (48)$$

### A.13 Signed Perpendicularity

$$f_{13} = \sum_{i=1}^{n-2} \sin(\varphi(s_i))^3 \quad (49)$$

### A.14 Angles after Equidistant Resampling

For this feature we do an equidistant resampling with 6 line segments. The five angles between succeeding lines are considered to make the features scale and rotation invariant (normalisation of writing speed).

$$f_{14} = \sum_{i=0}^4 \sin(\alpha_i), \sum_{i=0}^4 \cos(\alpha_i) \quad (50)$$

## B Rubine's Features

Features from this section are implementations of the described features by Rubine [35].

### B.1 Cosine of initial angle

$$f_1 = \frac{(x_2 - x_0)}{\sqrt{(x_2 - x_0)^2 + (y_2 - y_0)^2}} \quad (51)$$

### B.2 Sine of initial angle

$$f_2 = \frac{(y_2 - y_0)}{\sqrt{(x_2 - x_0)^2 + (y_2 - y_0)^2}} \quad (52)$$

### B.3 Length of bounding box diagonal

$$f_3 = \sqrt{(x_{\max} - x_{\min})^2 + (y_{\max} - y_{\min})^2} \quad (53)$$

### B.4 Angle of the bounding box diagonal

$$f_4 = \arctan \frac{y_{\max} - y_{\min}}{x_{\max} - x_{\min}} \quad (54)$$

### B.5 Distance between first and last point

$$f_5 = \sqrt{(x_{n-1} - x_0)^2 + (y_{n-1} - y_0)^2} \quad (55)$$

### B.6 Cosine of the angle between first and last point

$$f_6 = \frac{(x_{n-1} - x_0)}{f_5} \quad (56)$$

### B.7 Sine of the angle between first and last point

$$f_7 = \frac{(y_{n-1} - y_0)}{f_5} \quad (57)$$

### B.8 Total gesture length

Let  $\Delta x_i = x_{i+1} - x_i$ ,  $\Delta y_i = y_{i+1} - y_i$

$$f_8 = \sum_{i=0}^{n-2} \sqrt{\Delta x_i^2 + \Delta y_i^2} \quad (58)$$

### B.9 Total angle traversed

Let

$$\theta_i = \arctan \frac{\Delta x_i \Delta y_{i-1} - \Delta x_{i-1} \Delta y_i}{\Delta x_i \Delta x_{i-1} + \Delta y_i \Delta y_{i-1}} \quad (59)$$

$$f_9 = \sum_{i=1}^{n-2} \theta_i \quad (60)$$

### B.10 Sum of the absolute value of the angle at each point

$$f_{10} = \sum_{i=1}^{n-2} |\theta_i| \quad (61)$$

### B.11 Sum of the squared value of the angle at each point

$$f_{11} = \sum_{i=1}^{n-2} \theta_i^2 \quad (62)$$

### B.12 Maximum speed (squared) of the gesture

Let  $\Delta t_i = t_{i+1} - t_i$

$$f_{12} = \max_{i=0}^{n-2} \frac{\Delta x_i^2 + \Delta y_i^2}{\Delta t_i^2} \quad (63)$$

### B.13 Duration of the gesture

$$f_{13} = t_{n-1} - t_0 \quad (64)$$

## C Features by Willems and Niels

The features described in this section are implementations based on the feature set described by Willems and Niels [42]. For this section we use the following notations:

Let  $c$  be center of the bounding box around the gesture defined by the co-ordinate axes.

$$c = \begin{pmatrix} x_{center} \\ y_{center} \end{pmatrix} = \begin{pmatrix} x_{min} + \frac{1}{2}(x_{max} - x_{min}) \\ y_{min} + \frac{1}{2}(y_{max} - y_{min}) \end{pmatrix} \quad (65)$$

While the ratio of the co-ordinate axis is not rotation independent, the ratio of the principal axes is. To determine the principal axes *Principal Component Analysis* is used [42]. Let  $p_1$  and  $p_2$  be the normalised principal component vectors of the set  $S$ . And let  $c$  be the center of the box enclosing the trajectory and along the principal component vectors. The lengths of the major axes along the principal component vectors are

$$\alpha = 2 \max_{0 \leq i < n} |p_2 \cdot (c - s_i)|, \quad \beta = 2 \max_{0 \leq i < n} |p_1 \cdot (c - s_i)| \quad (66)$$

### C.1 Length of the gesture

$$f_1 = \sum_{i=0}^{n-2} \|s_{i+1} - s_i\| \quad (67)$$

### C.2 Area

The area around the gesture is calculated using Graham's convex hull algorithm [18].

$$f_2 = A \quad (68)$$

### C.3 Compactness

Let  $l$  be the the length of the perimeter of the convex hull, then compactness is defined as

$$f_3 = \frac{l^2}{A} \quad (69)$$

#### C.4 Ratio between co-ordinate axes

The lengths along the two co-ordinate axes (a, along the x-axis, and b along the y-axis) as given as

$$a = \max_{0 \leq i < j < n} |x_i - x_j| \quad (70)$$

$$b = \max_{0 \leq i < j < n} |y_i - y_j| \quad (71)$$

Eccentricity is a measure for the ratio between the co-ordinate axes.

$$f_4 = \sqrt{1 - \frac{b'^2}{a'^2}} \quad (72)$$

where  $a' = a \wedge b' = b$  if  $a > b$  else  $a' = b \wedge b' = a$ .

#### C.5 Ratio between co-ordinate axes

The ratio of the co-ordinate axes, which is very much related to eccentricity, is denoted as follows

$$f_5 = \frac{b'}{a'} \quad (73)$$

#### C.6 Closure

$$f_6 = \frac{f_{64}}{f_1} = \frac{\sum_{i=0}^{n-2} \|s_{i+1} - s_i\|}{\|s_{n-1} - s_0\|} \quad (74)$$

#### C.7 Circular variance

$$f_7 = \frac{\sum_{i=0}^{n-1} (\|s_i - \mu\| - f_{68})^2}{n \cdot f_{68}^2} \quad (75)$$

#### C.8 Curvature

Let angle between sequenced samples be:

$$\psi_{s_i} = \arccos \left\{ \frac{(s_i - s_{i-1}) \cdot (s_{i+1} - s_i)}{\|s_i - s_{i-1}\| \|s_{i+1} - s_i\|} \right\} \quad (76)$$

then curvature will be:

$$f_8 = \sum_{i=1}^{n-2} \psi_{s_i} \quad (77)$$

#### C.9 Average curvature

$$f_9 = \frac{1}{n-2} \sum_{i=1}^{n-2} \psi_{s_i} \quad (78)$$

### C.10 Standard deviation in curvature

$$f_{10} = \sqrt{\frac{1}{n-2} \sum_{i=1}^{n-2} (\psi_{s_i} - f_9)^2} \quad (79)$$

### C.11 Pen up/down ratio

Let  $\mathcal{G} = \{S_0, S_1, \dots, S_{n-1}\}$  be the set of strokes composing a gesture of  $n$  strokes. The duration of stroke  $S$  of length  $m$  is given by

$$\phi(S) = t_{m-1}(S) - t_0(S) \quad (80)$$

where  $t_j(S)$  is the  $j$ -th timestamp of stroke  $S$ . The duration of all strokes  $\mathcal{S}$  with  $|\mathcal{S}| = n$  is then defined as

$$\phi(\mathcal{S}) = \sum_{i=0}^{n-1} \phi(S_i) \quad (81)$$

and the duration of the entire gesture is given by

$$\phi(\mathcal{G}) = t_{|S_{n-1}|-1}(S_{n-1}) - t_0(S_0) \quad (82)$$

The ratio of pen up/down is the ratio between the time spent writing (pen down) and in air (pen up)

$$f_{11} = \frac{\text{airtime}}{\text{writetime}} = \frac{\phi(\mathcal{G}) - \phi(\mathcal{S})}{\phi(\mathcal{S})} \quad (83)$$

### C.12 Average direction

$$f_{12} = \frac{1}{n-1} \sum_{i=0}^{n-2} \arctan \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \quad (84)$$

### C.13 Perpendicularity

$$f_{13} = \sum_{i=1}^{n-2} \sin^2 \psi_{s_i} \quad (85)$$

### C.14 Average perpendicularity

$$f_{14} = \frac{1}{n-2} \sum_{i=1}^{n-2} \sin^2 \psi_{s_i} \quad (86)$$

### C.15 Standard deviation in perpendicularity

$$f_{15} = \sqrt{\frac{1}{n-2} \sum_{i=1}^{n-2} (\sin^2 \psi_{s_i} - f_{14})^2} \quad (87)$$

### C.16 Centroid offset

The principal axes are used to calculate the centroid offset:

$$f_{16} = |\mathbf{p}_2 \cdot (\boldsymbol{\mu} - \mathbf{c})| \quad (88)$$

### C.17 Length of first principal axis

Based on the principal axis, its length is another feature

$$f_{17} = \alpha \quad (89)$$

### C.18 Sine orientation of principal axis

The orientation of the principal axis  $\psi$  is given by

$$f_{18} = \sin \psi = p_{1_y} \quad (90)$$

### C.19 Cosine orientation of principal axis

$$f_{19} = \cos \psi = p_{1_x} \quad (91)$$

### C.20 Rectangularity

Based on the lengths of the major axes along the principal component vectors and the area of the convex hull  $A$ , the rectangularity is defined as:

$$f_{20} = \frac{A}{\alpha \cdot \beta} \quad (92)$$

### C.21 Maximum angular difference

$$f_{21} = \max_{1+k \leq i \leq n-k} \psi_{s_i}^k \quad (93)$$

### C.22 Average pressure

$$f_{22} = \frac{1}{n} \sum_{i=0}^{n-1} p_i \quad (94)$$

### C.23 Standard deviation of pressure

$$f_{23} = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (p_i - f_{22})^2} \quad (95)$$

### C.24 Duration

$$f_{24} = t_{n-1} - t_0 \quad (96)$$

**C.25 Average velocity**

$$\mathbf{v}_i = \frac{\|s_{i+1} - s_i\| + \|s_i - s_{i-1}\|}{t_{i+1} - t_{i-1}} \quad (97)$$

$$f_{25} = \frac{1}{n-2} \sum_{i=1}^{n-2} \|\mathbf{v}_i\| \quad (98)$$

**C.26 Standard deviation of velocity**

$$f_{26} = \sqrt{\frac{1}{n-2} \sum_{i=1}^{n-2} (\|\mathbf{v}_i\| - f_{25})^2} \quad (99)$$

**C.27 Maximum velocity**

$$f_{27} = \max_{1 \leq i \leq n-2} \|\mathbf{v}_i\| \quad (100)$$

**C.28 Average acceleration**

$$\mathbf{a}_i = \frac{\mathbf{v}_{i+1} - \mathbf{v}_{i-1}}{t_{i+1} - t_{i-1}} \quad (101)$$

$$f_{28} = \frac{1}{n-4} \sum_{i=2}^{n-3} \|\mathbf{a}_i\| \quad (102)$$

**C.29 Standard deviation of acceleration**

$$f_{29} = \sqrt{\frac{1}{n-4} \sum_{i=2}^{n-3} (\|\mathbf{a}_i\| - f_{28})^2} \quad (103)$$

**C.30 Maximum acceleration**

$$f_{30} = \max_{2 \leq i \leq n-3} \|\mathbf{a}_i\| \quad (104)$$

**C.31 Minimum acceleration**

$$f_{31} = \min_{2 \leq i \leq n-3} \|\mathbf{a}_i\| \quad (105)$$

**C.32 Number of cups**

$$f_{32} = nCups \quad (106)$$

**C.33 Offset of the first cup**

$$f_{33} = \text{lastCupOffset} \quad (107)$$

**C.34 Offset of the last cup**

$$f_{34} = \text{firstCupOffset} \quad (108)$$

**C.35 Initial horizontal offset**

$$f_{35} = \frac{x_0 - x_{\min}}{a} \quad (109)$$

**C.36 Final horizontal offset**

$$f_{36} = \frac{x_{n-1} - x_{\min}}{a} \quad (110)$$

**C.37 Initial vertical offset**

$$f_{37} = \frac{y_0 - y_{\min}}{b} \quad (111)$$

**C.38 Final vertical offset**

$$f_{38} = \frac{y_{n-1} - y_{\min}}{b} \quad (112)$$

**C.39 Number of straight lines**

Based on the definition of straight lines by Willems and Niels[42], we denote the set of straight lines  $L_i$  inside a gesture as  $\mathcal{L} = \{L_0, L_1, \dots, L_{i-1}\}$  and the number of straight lines as

$$f_{39} = |\mathcal{L}| \quad (113)$$

**C.40 Average length of straight lines**

Let  $\|L_i\|$  be the length of a straight line, then the average length of straight lines is calculated as

$$f_{40} = \frac{1}{|\mathcal{L}|} \sum_{i=0}^{|\mathcal{L}|-1} \|L_i\| \quad (114)$$

**C.41 Standard deviation of straight line length**

$$f_{41} = \sqrt{\frac{1}{|\mathcal{L}|} \sum_{i=0}^{|\mathcal{L}|-1} (\|L_i\| - f_{40})^2} \quad (115)$$



#### C.42 Straight line ratio

$$f_{42} = \sum_{i=0}^{|\mathcal{L}|-1} \frac{\|L_i\|}{\sum_{j=1}^{n-1} \|s_j - s_{j-1}\|} = \frac{1}{f_1} \sum_{i=0}^{|\mathcal{L}|-1} \|L_i\| \quad (116)$$

#### C.43 Largest straight line ratio

$$f_{43} = \max_{0 \leq i \leq n-1} \frac{\|L_i\|}{\sum_{j=1}^{n-1} \|s_j - s_{j-1}\|} = \frac{1}{f_1} \max_{0 \leq i < n < |\mathcal{L}|} \|L_i\| \quad (117)$$

#### C.44 Number of pen down events

Let  $\mathcal{G} = \{S_0, S_1, \dots, S_{i-1}\}$  be the set of strokes composing a gesture of  $n$  strokes. The number of pen down events equals the number of strokes

$$f_{44} = |\mathcal{G}| \quad (118)$$

#### C.45 Octants

$$f_{44+o} = \frac{1}{n-1} \sum_{i=0}^{n-1} \omega_{io} \quad (119)$$

where

$$\omega_{io} = \begin{cases} 1 & \text{if } \frac{\pi}{4}(o-1) \leq \nu_i < \frac{\pi}{4}o \\ 0 & \text{if } \nu_i < \frac{\pi}{4}(o-1) \vee \nu_i \geq \frac{\pi}{4}o \end{cases} \quad (120)$$

and where

$$\nu_i = \arctan \frac{dy_i}{dx_i} \quad (121)$$

and

$$\begin{aligned} dx_i &= x_i - x_{center} \\ dy_i &= y_i - y_{center} \end{aligned}$$

#### C.46 Number of connecting strokes

We define the set of connected components as  $\mathcal{C} = \{C_0, C_1, \dots, C_{i-1}\}$ . A connected component  $C_i$  is a part of a gesture that consists of one or more strokes that touch each other, and that do not touch any other strokes [42].

$$f_{53} = |\mathcal{C}| \quad (122)$$

**C.47 Number of crossings**

$$f_{54} = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \kappa_{ij} \quad (123)$$

where

$$\kappa_{ij} = \begin{cases} 1 & \text{if } s_i \rightarrow s_{i+1} \cap s_j \rightarrow s_{j+1} \neq \emptyset \\ 0 & \text{if } s_i \rightarrow s_{i+1} \cap s_j \rightarrow s_{j+1} = \emptyset \end{cases} \quad (124)$$

**C.48 Cosine of initial angle**

$$f_{55} = \frac{x_2 - x_0}{\|s_2 - s_0\|} \quad (125)$$

**C.49 Sine of initial angle**

$$f_{56} = \frac{y_2 - y_0}{\|s_2 - s_0\|} \quad (126)$$

**C.50 Length of the bounding box diagonal**

Given the two co-ordinate axes the length of the bounding box is given as

$$f_{57} = \sqrt{a^2 + b^2} \quad (127)$$

**C.51 Angle of the bounding box diagonal**

$$f_{58} = \tan \frac{b}{a} \quad (128)$$

**C.52 Length between first and last point**

$$f_{59} = \|s_{n-1} - s_0\| \quad (129)$$

**C.53 Cosine of first to last point**

$$f_{60} = \frac{x_{n-1} - x_0}{\|s_{n-1} - s_0\|} \quad (130)$$

**C.54 Sine of first to last point**

$$f_{61} = \frac{y_{n-1} - y_0}{\|s_{n-1} - s_0\|} \quad (131)$$

**C.55 Absolute curvature**

$$f_{62} = \sum_{i=1}^{n-2} |\psi_{s_i}| \quad (132)$$

### C.56 Squared curvature

$$f_{63} = \sum_{i=1}^{n-2} \psi_{s_i}^2 \quad (133)$$

### C.57 Macro perpendicularity

Let angle between sampled points be:

$$\psi_{s_i}^k = \arccos \left\{ \frac{(s_i - s_{i-k}) \cdot (s_{i+k} - s_i)}{\|s_i - s_{i-k}\| \|s_{i+k} - s_i\|} \right\} \quad (134)$$

then macro perpendicularity will be:

$$f_{64} = \sum_{i=1+k}^{n-k} \sin^2 \psi_{s_i}^k \quad (135)$$

### C.58 Average macro perpendicularity

$$f_{65} = \frac{1}{n-2k} \sum_{i=1+k}^{n-k} \sin^2 \psi_{s_i}^k \quad (136)$$

### C.59 Standard deviation in macro perpendicularity

$$f_{66} = \sqrt{\frac{1}{n-2k} \sum_{i=1+k}^{n-k} (\sin^2 \psi_{s_i}^k - f_{65})^2} \quad (137)$$

### C.60 Ratio of principal axes

Based on the lengths of the major axes along the principal component vectors, the ratio of the principal axes becomes:

$$f_{67} = \frac{\beta}{\alpha} \quad (138)$$

### C.61 Average centroidal radius

The average distance of sample points from the centroid is a feature called average centroidal radius.

$$f_{68} = \frac{1}{n} \sum_{i=0}^{n-1} \|s_i - \mu\| \quad (139)$$

### C.62 69 Standard deviation of the centroidal radius

$$f_{69} = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (\|s_i - \mu\| - f_{68})^2} \quad (140)$$

### C.63 Chain codes

Let the chain code be defined as

$$C_s = \begin{cases} 1 & \text{if } 0 \leq \psi_s < \frac{\pi}{4} \\ 2 & \text{if } \frac{\pi}{4} \leq \psi_s < \frac{\pi}{2} \\ 3 & \text{if } \frac{\pi}{2} \leq \psi_s < \frac{3\pi}{4} \\ 4 & \text{if } \frac{3\pi}{4} \leq \psi_s < \pi \\ 5 & \text{if } \pi \leq \psi_s < \frac{5\pi}{4} \\ 6 & \text{if } \frac{5\pi}{4} \leq \psi_s < \frac{3\pi}{2} \\ 7 & \text{if } \frac{3\pi}{2} \leq \psi_s < \frac{7\pi}{4} \\ 8 & \text{if } \frac{7\pi}{4} \leq \psi_s < 2\pi \end{cases} \quad (141)$$

then the average angle of the chain code will be

$$\psi_{C_s} = \frac{(C_s - \frac{1}{2})\pi}{4}$$

then

$$f_{68+2s} = \sin \psi_{C_s}$$

and

$$f_{69+2s} = \cos \psi_{C_s}$$

### C.64 Average stroke length

If  $S_i \in \mathcal{S}$  is a stroke with  $n$  sample points, then let  $L_i$  be the length of that stroke:

$$L_i = \sum_{j=0}^{n-2} \|s_{j+1} - s_j\| \quad (142)$$

Assuming  $|\mathcal{S}| = m$  the average stroke length is given by

$$f_{86} = \frac{1}{m} \sum_{i=0}^{m-1} L_i \quad (143)$$

### C.65 Standard deviation in stroke length

$$f_{87} = \sqrt{\frac{1}{m} \sum_{i=0}^{m-1} (L_i - f_{86})^2} \quad (144)$$

### C.66 Average stroke direction

If  $S_i \in \mathcal{S}$  is a stroke with  $n$  sample points, then let  $\varphi_i$  be the direction of that stroke:

$$\varphi_i = \frac{1}{n-1} \sum_{j=0}^{n-2} \arctan \frac{y_{j+1} - y_j}{x_{j+1} - x_j} \quad (145)$$

Assuming  $|\mathcal{S}| = m$  the average stroke direction is given by

$$f_{88} = \frac{1}{m} \sum_{i=0}^{m-1} \varphi_i \quad (146)$$

### C.67 Standard deviation in stroke direction

$$f_{89} = \sqrt{\frac{1}{m} \sum_{i=0}^{m-1} (\varphi_i - f_{88})^2} \quad (147)$$

## D HBF49 Features

The features described in this section are implementations based on the feature set HBF49 described by Delaye and Anquetil [11]. For this section we use the following notations:

Let  $B$  be the rectangular bounding box defined by  $x_{min}, x_{max}, y_{min}, y_{max}$ . The width  $w$  and height  $h$  of this box are defined as

$$w = x_{max} - x_{min}, h = y_{max} - y_{min} \quad (148)$$

Coordinates  $c_x$  and  $c_y$  are the coordinates of the center point  $c$  of bounding box  $B$ .

Let  $L_{i,j}$  be the length of the path between sample points  $s_i$  and  $s_j$ , then  $L$  is the total length of the path of the gesture.

Let  $\mathcal{S}$  be the set of strokes composing the gesture.

#### D.1 Horizontal position of first point

Let  $l = \max(h, w)$  be the side of a square box centered on  $c$ . The normalised position of the first point is then given by

$$f_1 = \frac{x_0 - c_x}{l} + \frac{1}{2} \quad (149)$$

#### D.2 Vertical position of first point

$$f_2 = \frac{y_0 - c_y}{l} + \frac{1}{2} \quad (150)$$

### D.3 Horizontal position of last point

$$f_3 = \frac{x_{n-1} - c_x}{l} + \frac{1}{2} \quad (151)$$

### D.4 Vertical position of last point

$$f_4 = \frac{y_{n-1} - c_y}{l} + \frac{1}{2} \quad (152)$$

### D.5 First point to last point vector length

$$v = \overrightarrow{s_0 s_{n-1}} \quad (153)$$

$$f_5 = \|v\| \quad (154)$$

### D.6 Sine of first point to last point vector

Let  $u_x$  be the unit vector codirectional with the x axis.

$$f_6 = \frac{v_x \cdot u_x}{f_5} \quad (155)$$

### D.7 Cosine of first point to last point vector

Let  $u_y$  be the unit vector codirectional with the y axis.

$$f_7 = \frac{v_y \cdot u_y}{f_5} \quad (156)$$

### D.8 Closure

$$f_8 = \frac{\|v\|}{L}. \quad (157)$$

### D.9 Sine of initial angle

Our initial vector between the first and third point is given by  $w = \overrightarrow{s_0 s_2}$ .

$$f_9 = \frac{w_x \cdot u_x}{\|w\|} \quad (158)$$

### D.10 Cosine of initial angle

$$f_{10} = \frac{w_y \cdot u_y}{\|w\|} \quad (159)$$

### D.11 Horizontal inflexion

Let  $s_m$  be the middle-path point with respect to the middle point of segment  $s_0 s_{n-1}$ .

$$f_{11} = \frac{1}{w} \left( x_m - \frac{x_1 + x_n}{2} \right) \quad (160)$$

### D.12 Vertical inflexion

$$f_{12} = \frac{1}{h} \left( y_m - \frac{y_1 + y_n}{2} \right) \quad (161)$$

### D.13 Downstroke proportion

Downstrokes are portions of drawing trajectories oriented towards the bottom of the writing surface, i.e. oriented towards increasing values in dimension y.

$$f_{13} = \sum_{k=0} L_k \quad \{\forall k \in \mathcal{S} \mid k \text{ is a downstroke}\} \quad (162)$$

### D.14 Number of strokes

$$f_{14} = |\mathcal{S}| \quad (163)$$

### D.15 Angle of the bounding box diagonal

$$f_{15} = \arctan \frac{h}{w} \quad (164)$$

### D.16 Trajectory length

$$f_{16} = L \quad (165)$$

### D.17 Ratio between bounding box and trajectory length

$$f_{17} = \frac{w + h}{L} \quad (166)$$

### D.18 Deviation

$$f_{18} = \frac{1}{n} \sum_{i=0}^{n-1} \|s_i \mu\| \quad (167)$$

### D.19 Average direction

$$f_{19} = \frac{1}{n-1} \sum_{i=0}^{n-2} \arctan \left( \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \right) \quad (168)$$

### D.20 Curvature

$$\theta_i = \arccos \frac{\overrightarrow{s_{i-1}s_i} \cdot \overrightarrow{s_i s_{i+1}}}{\|\overrightarrow{s_{i-1}s_i}\| \|\overrightarrow{s_i s_{i+1}}\|} \quad (169)$$

$$f_{20} = \sum_{i=1}^{n-2} \theta_i \quad (170)$$

### D.21 Perpendicularity

$$f_{21} = \sum_{i=1}^{n-2} \sin^2(\theta_i) \quad (171)$$

### D.22 k-Perpendicularity

$$\theta_i^k = \arccos \frac{\overrightarrow{s_{i-k} s_i} \cdot \overrightarrow{s_i s_{i+k}}}{\|\overrightarrow{s_{i-k} s_i}\| \|\overrightarrow{s_i s_{i+k}}\|} \quad (172)$$

$$f_{22} = \sum_{i=k}^{n-k-1} \sin^2(\theta_i^k) \quad (173)$$

### D.23 k-Angle

$$f_{23} = \max_{i=k}^{n-k-1} \theta_i^k \quad (174)$$

### D.24 Dominant direction

Let  $n_a$  be the number of segments in  $\mathcal{S}$  ( $n_a = n - K$ , with  $K$  the number of strokes).

$$f_{24} = \frac{h_1 + h_5}{n_a} \quad (175)$$

$$f_{25} = \frac{h_2 + h_6}{n_a} \quad (176)$$

$$f_{26} = \frac{h_3 + h_7}{n_a} \quad (177)$$

$$f_{27} = \frac{h_4 + h_8}{n_a} \quad (178)$$

### D.25 Local changes in direction

Local angle benefit from smoothing by linear combination of  $\theta_i$  and  $\theta_i^k$  (refer to  $f_{14}$  and  $f_{16}$ ):

$$\psi_i^k = \gamma \theta_i + (1 - \gamma) \theta_i^k \quad (179)$$

In the HBF49 feature set [11] the values are set empirically to  $\gamma = 0.25$  and  $k = 2$ . The contributions of  $\psi_i^k$  angles are accumulated in four histogram bins uniformly distributed in  $[0, \pi]$ . Contributions to the histogram are weighted by the inverse of their angular distance with the central direction of the two neighboring bins. The features are obtained from the histogram  $h$  divided by  $n_a$  (see  $f_{24}$ - $f_{27}$ ).

$$f_{28} = \frac{h[0]}{n_a} \quad (180)$$



$$f_{29} = \frac{h[1]}{n_a} \quad (181)$$

$$f_{30} = \frac{h[2]}{n_a} \quad (182)$$

$$f_{31} = \frac{h[3]}{n_a} \quad (183)$$

## D.26 2D Histogram

For the 2D histogram we divide the rectangular bounding box around the figure into  $3 \times 3$  partitions of equal size. Sampling points are sorted into the 9 cells resulting from partitioning [21]. For each sample point a fuzzy weighted contribution to the 4 neighboring cells is computed, where the weights depend on the distance from the point to the cell centers [1].

$$f_{32} = \frac{1}{n} \sum_{i=0}^{n-1} \mu_{11}(s_i) \quad (184)$$

$$f_{33} = \frac{1}{n} \sum_{i=0}^{n-1} \mu_{12}(s_i) \quad (185)$$

$$f_{34} = \frac{1}{n} \sum_{i=0}^{n-1} \mu_{13}(s_i) \quad (186)$$

$$f_{35} = \frac{1}{n} \sum_{i=0}^{n-1} \mu_{21}(s_i) \quad (187)$$

$$f_{36} = \frac{1}{n} \sum_{i=0}^{n-1} \mu_{22}(s_i) \quad (188)$$

$$f_{37} = \frac{1}{n} \sum_{i=0}^{n-1} \mu_{23}(s_i) \quad (189)$$

$$f_{38} = \frac{1}{n} \sum_{i=0}^{n-1} \mu_{31}(s_i) \quad (190)$$

$$f_{39} = \frac{1}{n} \sum_{i=0}^{n-1} \mu_{32}(s_i) \quad (191)$$

$$f_{40} = \frac{1}{n} \sum_{i=0}^{n-1} \mu_{33}(s_i) \quad (192)$$

## D.27 Hu moments

Let  $\mu = (\mu_x, \mu_y)$  be the center of gravity, then the central inertia moments are computed as follows

$$m_{pq} = \sum_{i=1}^n (x_i - \mu_x)^p (y_i - \mu_y)^q, \quad \text{for } 0 \leq p, q \leq 3 \quad (193)$$

In order to guarantee scale independence the moments are normalised:

$$\nu_{pq} = \frac{m_{pq}}{m_{00}^\gamma}, \quad \text{with } \gamma = 1 + \frac{p+q}{2} \quad (194)$$

The seven Hu moments [20] are computed as:

$$f_{41} = \nu_{02} + \nu_{20} \quad (195)$$

$$f_{42} = (\nu_{20} - \nu_{02})^2 + 4\nu_{11}^2 \quad (196)$$

$$f_{43} = (\nu_{30} - 3\nu_{12})^2 + (3\nu_{21} - \nu_{03})^2 \quad (197)$$

$$f_{44} = (\nu_{30} + \nu_{12})^2 + (\nu_{21} + \nu_{03})^2 \quad (198)$$

$$\begin{aligned} f_{45} = & (\nu_{30} - 3\nu_{12})^2 (\nu_{30} + \nu_{03}) \left[ (\nu_{30} + \nu_{12})^2 - 3(\nu_{21} + \nu_{03})^2 \right] \\ & + (3\nu_{21} - \nu_{03}) (\nu_{21} + \nu_{03}) \left[ 3(\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2 \right] \end{aligned} \quad (199)$$

$$\begin{aligned} f_{46} = & (\nu_{20} - \nu_{02}) \left[ (\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2 \right] \\ & + 4\nu_{11} (\nu_{30} + \nu_{12}) (\nu_{21} + \nu_{03}) \end{aligned} \quad (200)$$

$$\begin{aligned} f_{47} = & (3\nu_{21} - \nu_{03}) (\nu_{30} + \nu_{12}) \left[ (\nu_{30} + \nu_{12})^2 - 3(\nu_{21} + \nu_{03})^2 \right] \\ & - (\nu_{30} - 3\nu_{12}) (\nu_{21} + \nu_{03}) \left[ 3(\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2 \right] \end{aligned} \quad (201)$$

## D.28 Normalised convex hull area

Let  $H$  be the convex hull around the gesture, then  $A_H$  denotes the area of the convex hull.

$$f_{48} = \frac{A_H}{w * h} \quad (202)$$

## D.29 Compactness

$$f_{49} = \frac{L^2}{A_H} \quad (203)$$