Measuring Linearity of Planar Curves

Joviša Žunić, Jovanka Pantović and Paul L. Rosin

Abstract In this paper we define a new linearity measure which can be applied to open planar curve segments. We have considered the sum of the distances between the curve end points and the curve centroid. We have shown that this sum is bounded from above by the length of the curve segment considered. In addition, we have proven that this sum equals the length of the curve segment only in the case of straight line segments. Exploiting such a nice characterization of straight line segments, we define a new linearity measure for planar curves. The new measure ranges over the interval (0, 1], and produces the value 1 if and only if the measured line is a perfect straight line segment. Also, the new linearity measure is invariant with respect to translations, rotations and scaling transformations.

Keywords Shape · Shape descriptors · Curves · Linearity measure · Image processing.

1 Introduction

Shape descriptors have been employed in many computer vision and image processing tasks (e.g. image retrieval, object classification, object recognition, object identification, etc.). Different mathematical tools have been used to define the shape descriptors: algebraic invariants [1], Fourier analysis [2], morphological operations [3], integral transformations [4], statistical methods [5], fractal techniques [6], logic [7], combinatorial methods [8], multiscale approaches [9], integral invariants

J. Žunić (✉)
Computer Science, University of Exeter, Exeter EX4 4QF, UK
e-mail: j.zunic@ex.ac.uk

J. Žunić · J. Pantović
Mathematical Institute of the Serbian Academy of Sciences and Arts, Belgrade, Serbia

J. Pantović
Faculty of Technical Sciences, University of Novi Sad, 21000 Novi Sad, Serbia

P.L. Rosin
Cardiff University, School of Computer Science, Cardiff CF24 3AA, UK

A. Fred and M. De Marsico (eds.), *Pattern Recognition Applications and Methods*,
Advances in Intelligent Systems and Computing 318,
DOI 10.1007/978-3-319-12610-4_16

[10], etc. Generally speaking, shape descriptors can be classified into two groups: area based descriptors and boundary based ones. Area based descriptors are more robust (i.e. less sensitive to noise or shape deformations) while boundary based descriptors are more sensitive. A preference for either type of descriptor depends on the application performed and the data available. For example low quality data would require robust descriptors (i.e. area based ones) while high precision tasks would require more sensitive descriptors (i.e. boundary based ones). In the literature so far, more attention has been paid to the area based descriptors, not only because of their robustness but also because they are easier to be efficiently estimated when working with discrete data. Due to the recent proliferation of image verification, identification and recognition systems there is a strong demand for shape properties that can be derived from their boundaries [10–12]. It is worth mentioning that some objects, like human signatures for example, are curves by their nature and area based descriptors cannot be used for their analysis.

In this paper we deal with linearity measures that should indicate the degree to which an open curve segment differs from a perfect straight line segment. Several linearity measures for curve segments are already considered in the literature [13–16].

Perhaps the simplest way to define the linearity measure of an open curve segment is to consider the ratio between the length of the curve considered and the distance between its end points. This is a natural and simple definition which is also called the *straightness index* [17]. It satisfies the following basic requirements for a linearity measure of open curve segments.

- The straightness index varies through the interval $(0, 1]$;
- The straightness index equals 1 only for straight line segments;
- The straightness index is invariant with respect to translation, rotation and scaling transformation on a curve considered.

Also, the straightness index is simple to compute and its behavior can be clearly predicted, i.e. we can see easily which curves have the same linearities, measured by the straightness index. It is obvious that those curves whose end points and the length coincide, have the same straightness index. But the diversity of such curves is huge and the straightness index cannot distinguish among them, which could be a big drawback in certain applications. Some illustrations using simple polygonal curves are shown in Fig. 1.

Fig. 1 Five displayed curves (*solid lines*) have different linearities measured by $\mathcal{L}(\mathcal{C})$. The straightness index has the same value for all five curves

In this paper we define a new linearity measure $\mathcal{L}(\mathcal{C})$ for open curve segments. The new measure satisfies the basic requirements (listed above) which are expected to be satisfied for any curve linearity measure. Since it considers the distance of the end points of the curve to the centroid of the curve, the new measure is also easy to compute. The fact that it uses the curve centroids implies that it takes into account a relative distribution of the curve points.

The paper is organized as follows. Section 2 gives basic definitions and denotations. The new linearity measure for planar open curve segments is in Sect. 3. Several experiments which illustrate the behavior and the classification power of the new linearity measure are provided in Sect. 4. Concluding remarks are in Sect. 5.

2 Definitions and Denotations

Without loss of generality, throughout the paper, it will be assumed (even if not mentioned) that every curve $\mathcal{C}$ has length equal to 1 and is given in an arc-length parametrization. I.e., planar curve segment $\mathcal{C}$ is represented as:

$$x = x(s), \qquad y = y(s), \quad \text{where } s \in [0, 1].$$

The parameter s measures the distance of the point $(x(s), y(s))$ from the curve start point $(x(0), y(0))$, along the curve $\mathcal{C}$.

The centroid of a given (unit length) planar curve $\mathcal{C}$ will be denoted by $(x_{\mathcal{C}}, y_{\mathcal{C}})$ and computed as

$$(x_{\mathcal{C}}, y_{\mathcal{C}}) = \left(\int_{\mathcal{C}} x(s)\, ds, \int_{\mathcal{C}} y(s)\, ds \right). \tag{1}$$

Taking into account that the length of $\mathcal{C}$ is assumed to be equal to 1, we can see that the coordinates of the curve centroid, as defined in (1), are the average values of the curve points.

As usual,

$$d_2(A, B) = \sqrt{(x-u)^2 + (y-v)^2}$$

will denote the Euclidean distance between the points $A = (x, y)$ and $B = (u, v)$.

As mentioned, we introduce a new linearity measure $\mathcal{L}(\mathcal{C})$ which assigns a number from the interval $(0, 1]$. The curve $\mathcal{C}$ is assumed to have the length 1. More precisely, any appearing curve will be scaled by the factor which equals the length of it before the processing. So, an arbitrary curve $\mathcal{C}_a$ would be replaced with the curve $\mathcal{C}$ defined by

$$\mathcal{C} = \frac{1}{\int_{\mathcal{C}_a} ds} \cdot \mathcal{C}_a = \left\{ \left(\frac{x}{\int_{\mathcal{C}_a} ds}, \frac{y}{\int_{\mathcal{C}_a} ds} \right) \mid (x, y) \in \mathcal{C}_a \right\}.$$

Shape descriptors/measures are very useful for discrimination among the objects—in this case open curve segments. Usually the shape descriptors have a clear geometric meaning and, consequently, the shape measures assigned to such descriptors have a predictable behavior. This is an advantage because the suitability of a certain measure to a particular shape-based task (object matching, object classification, etc.) can be predicted to some extent. On the other hand, a shape measure assigns to each object (here curve segment) a single number. In order to increase the performance of shape based tasks, a common approach is to assign a graph (instead of a number) to each object. E.g. such approaches define *shape signature* descriptors, which are also 'graph' representations of planar shapes, often used in shape analysis tasks [18, 19], but they differ from the idea used here and in [16].

We will apply a similar idea here as well. To compare objects considered we use *linearity plots* (the approach is taken from [16] where more details can be found) to provide more information than a single linearity measurement. The idea is to compute linearity incrementally, i.e. to compute linearity of sub-segments of $\mathcal{C}$ determined by the start point of $\mathcal{C}$ and another point which moves along the curve $\mathcal{C}$ from the beginning to the end of $\mathcal{C}$. The linearity plot $\mathrm{P}(\mathcal{C})$, associated with the given curve $\mathcal{C}$ is formally defined as follows.

Definition 1 Let $\mathcal{C}$ be a curve given in an arc-length parametrization: $x = x(s)$, $y = y(s)$, and $s \in [0, 1]$. Let $\mathbf{A}(s)$ be the part of the curve $\mathcal{C}$ bounded by the starting point $(x(0), y(0))$ and by the point $(x(s), y(s)) \in \mathcal{C}$. Then, for a linearity measure $\mathcal{L}$, the linearity plot $\mathrm{P}(\mathcal{C})$ is defined by:

$$\mathrm{P}(\mathcal{C}) = \{(s, \mathcal{L}(\mathbf{A}(s)) \mid \ s \in [0, 1]\}. \tag{2}$$

We will also use the *reverse linearity plot* $\mathrm{P}_{rev}(\mathcal{C})$ defined as:

$$\mathrm{P}_{rev}(\mathcal{C}) = \{(s, \mathcal{L}(\mathbf{A}_{rev}(1 - s)) \mid \ s \in [0, 1]\}, \tag{3}$$

where $\mathbf{A}_{rev}(1 - s)$ is the segment of the curve $\mathcal{C}$ determined by the end point $(x(1), y(1))$ of $\mathcal{C}$ and the point which moves from the end point of $\mathcal{C}$, to the start point of $\mathcal{C}$, along the curve $\mathcal{C}$. In other words, $\mathrm{P}_{rev}(\mathcal{C})$ is the linearity plot of the curve $\mathcal{C}'$ which coincides with the curve $\mathcal{C}$ but the start (end) point of $\mathcal{C}$ is the end (start) point of $\mathcal{C}'$. A parametrization of $\mathcal{C}'$ can be obtained by replacing the parameter s, in the parametrization of $\mathcal{C}$, by a new parameter s' such that $s' = 1 - s$. Obviously such a defined s' measures the distance of the point $(x(s'), y(s'))$ from the starting point $(x(s' = 0), y(s' = 0))$ of $\mathcal{C}'$ along the curve $\mathcal{C}'$, as s' varies through the interval $[0, 1]$.

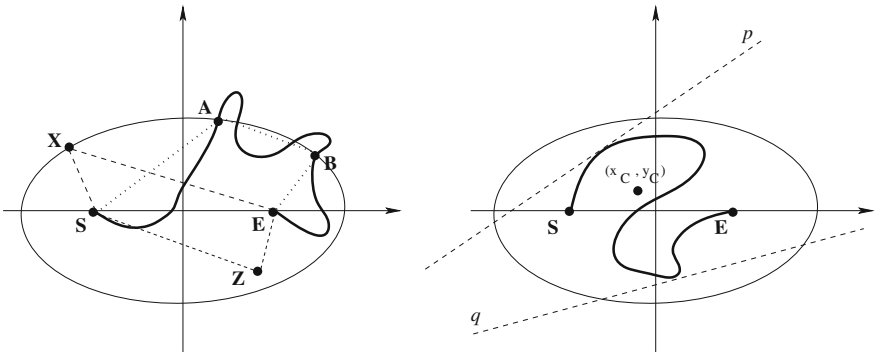

Fig. 2 Terms in the proof of Theorem 1 are illustrated above

3 New Linearity Measure for Open Curve Segments

In this section we introduce a new linearity measure for open planar curve segments.

We start with the following theorem which says that amongst all curves having the same length, straight line segments have the largest sum of distances between the curve end points to the curve centroid. This result will be exploited to define the new linearity measure for open curve segments.

Theorem 1 *Let $\mathcal{C}$ be an open curve segment given in an arc-length parametrization $x = x(s)$, $y = y(s)$, and $s \in [0, 1]$. The following statements hold:*

(a) *The sum of distances of the end points $(x(0), y(0))$ and $(x(1), y(1))$ from the centroid $(x_{\mathcal{C}}, y_{\mathcal{C}})$ of the curve $\mathcal{C}$ is bounded from above by 1, i.e.:*

$$d_2((x(0), y(0)), (x_{\mathcal{C}}, y_{\mathcal{C}})) + d_2((x(1), y(1)), (x_{\mathcal{C}}, y_{\mathcal{C}})) \leq 1. \tag{4}$$

(b) *The upper bound established by the previous item is reached by the straight line segment and, consequently, cannot be improved.*

Proof Let $\mathcal{C}$ be a curve given in an arc-length parametrization: $x = x(s)$ and $y = y(s)$, with $s \in [0, 1]$, and let $S = (x(0), y(0))$ and $E = (x(1), y(1))$ be the end points of $\mathcal{C}$. We can assume, without loss of generality, that the curve segment $\mathcal{C}$ is positioned such that

- the end points S and E belong to the x-axis (i.e. $y(0) = y(1) = 0$), and
- S and E are symmetric with respect to the origin (i.e. $-x(0) = x(1)$),

as illustrated in Fig. 2. Furthermore, let

$$\mathcal{E} = \{X = (x, y) \mid d_2(X, S) + d_2(X, E) = 1\}$$

be an ellipse which consists of points whose sum of distances to the points S and E is equal 1. Now, we prove (a) in two steps.

(i) First, we prove that the curve $\mathcal{C}$ and the ellipse $\mathcal{E}$ do not have more than one intersection point (i.e. $\mathcal{C}$ belongs to the closed region bounded by $\mathcal{E}$).
This will be proven by a contradiction. So, let us assume the contrary, that $\mathcal{C}$ intersects $\mathcal{E}$ at k ($k \geq 2$) points:

$$(x(s_1), y(s_1)),\ (x(s_2), y(s_2)), \ldots,\ (x(s_k), y(s_k)),$$

where $0 < s_1 < s_2 < \cdots < s_k < 1$. Let $A = (x(s_1), y(s_1))$ and $B = (x(s_k), y(s_k))$. Since the sum of the lengths of the straight line segments $[SA]$ and $[AE]$ is equal to 1, the length of the polyline $SABE$ is, by the triangle inequality, bigger than 1. Since the length of the arc $\widehat{SA}$ (along the curve $\mathcal{C}$) is not smaller than the length of the edge $[SA]$, the length of the arc $\widehat{AB}$ (along the curve $\mathcal{C}$) is not smaller than the length of the straight line segment $[AB]$, and the length of the arc $\widehat{BE}$ (along the curve $\mathcal{C}$) is not smaller than the length of the straight line segment $[BE]$, we deduce that the curve $\mathcal{C}$ has length bigger than 1, which is a contradiction. A more formal derivation of the contradiction $1 < 1$ is

$$\begin{aligned} 1 &= d_2(S, A) + d_2(A, E) \\ &< d_2(S, A) + d_2(A, B) + d_2(B, E) \\ &\leq \int_{\widehat{SA}} ds + \int_{\widehat{AB}} ds + \int_{\widehat{BE}} ds = \int_{\mathcal{C}} ds \\ &= 1. \end{aligned} \tag{5}$$

So, $\mathcal{C}$ and $\mathcal{E}$ do not have more than one intersection point, implying that $\mathcal{C}$ lies in the closed region bounded by $\mathcal{E}$.

(ii) Second, we prove that the centroid of $\mathcal{C}$ does not lie outside of $\mathcal{E}$.
The proof follows easily:

- the convex hull $CH(\mathcal{C})$ of $\mathcal{C}$ is the smallest convex set which includes $\mathcal{C}$ and, consequently, is a subset of the region bounded by $\mathcal{E}$;
- The centroid of $\mathcal{C}$ lies in the convex hull $CH(\mathcal{C})$ of $\mathcal{C}$ because it belongs to every half plane which includes $\mathcal{C}$ (the intersection of such half planes is actually the convex hull of $\mathcal{C}$ (see [20]));
- the two items above give the required:

$$(x_{\mathcal{C}}, y_{\mathcal{C}}) \in CH(\mathcal{C}) \subset region\_bounded\_by\_\mathcal{E}.$$

Finally, since the centroid of $\mathcal{C}$ does not lie outside $\mathcal{E}$, the sum of the distances of the centroid $(x_{\mathcal{C}}, y_{\mathcal{C}})$ of $\mathcal{C}$ to the points S and E may not be bigger than 1, i.e.

$$\begin{aligned} & d_2((x(0), y(0)), (x_{\mathcal{C}}, y_{\mathcal{C}})) + d_2((x(1), y(1)), (x_{\mathcal{C}}, y_{\mathcal{C}})) \\ & = d_2(S, (x_{\mathcal{C}}, y_{\mathcal{C}})) + d_2(E, (x_{\mathcal{C}}, y_{\mathcal{C}})) \\ & \leq 1. \end{aligned}$$

This establishes (a).

To prove (b) it is enough to notice that if $\mathcal{C}$ is a straight line segment of length 1, then the sum of its end points to the centroid of $\mathcal{C}$ is 1. □

Now, motivated by the results of Theorem 1, we give the following definition for a new linearity measure $\mathcal{L}(\mathcal{C})$ for open curve segments.

Definition 2 Let $\mathcal{C}$ be an open curve segment, whose length is 1. Then, the linearity measure $\mathcal{L}(\mathcal{C})$ of $\mathcal{C}$ is defined as the sum of distances between the centroid $(x_{\mathcal{C}}, y_{\mathcal{C}})$ of $\mathcal{C}$ and the end points of $\mathcal{C}$. I.e.:

$$\mathcal{L}(\mathcal{C}) = \sqrt{(x(0) - x_{\mathcal{C}})^2 + (y(0) - y_{\mathcal{C}})^2} + \sqrt{(x(1) - x_{\mathcal{C}})^2 + (y(1) - y_{\mathcal{C}})^2}$$

where $x = x(s)$, $y = y(s)$, $s \in [0, 1]$ is an arc-length representation of $\mathcal{C}$.

The following theorem summarizes desirable properties of $\mathcal{L}(\mathcal{C})$.

Theorem 2 *The linearity measure $\mathcal{L}(\mathcal{C})$ has the following properties:*

(i) $\mathcal{L}(\mathcal{C}) \in (0, 1]$, *for all open curve segments $\mathcal{C}$;*
(ii) $\mathcal{L}(\mathcal{C}) = 1 \Leftrightarrow$ *$\mathcal{C}$ is a straight line segment;*
(iii) $\mathcal{L}(\mathcal{C})$ *is invariant with respect to the similarity transformations.*

Proof Item (i) is a direct consequence of Theorem 1.

To prove (ii) we will use the same notations as in the proof of Theorem 1 and give a proof by contradiction. So, let us assume the following:

- the curve $\mathcal{C}$ differs from a straight line segment, and
- the sum of distances between the end points, and the centroid of $\mathcal{C}$ is 1.

Since $\mathcal{C}$ is not a straight line segment, then $\mathbf{d}_2(S, E) < 1$, and the centroid $(x_{\mathcal{C}}, y_{\mathcal{C}})$ lies on the ellipse $\mathcal{E} = \{X = (x, y) \mid \mathbf{d}_2(X, S) + \mathbf{d}_2(X, E) = 1\}$. Further, it would mean that there are points of the curve $\mathcal{C}$ belonging to both hyperplanes determined by the tangent on the ellipse $\mathcal{E}$ passing through the centroid of $\mathcal{C}$. This would contradict the fact that $\mathcal{C}$ and $\mathcal{E}$ do not have more than one intersection point (which was proven as a part of the proof of Theorem 1).

To prove item (iii) it is enough to notice that translations and rotations do not change the distance between the centroid and the end points. Since we assume that $\mathcal{C}$ is represented by using an arc-length parametrization: $x = x(s)$, $y = y(s)$, with the parameter s varying through $[0, 1]$, the new linearity measure $\mathcal{L}(\mathcal{C})$ is invariant with respect to scaling transformations as well. □

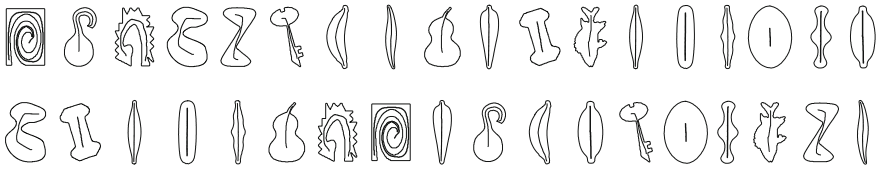

Fig. 3 The shapes are ranking by their axes in increasing order according to linearity $\mathcal{L}$ (*top row*) and sigmoidality S_3 (*bottom row*)

4 Experiments

In this section we provide several experiments in order to illustrate the behavior and efficiency of the linearity measure introduced here.

First Experiment: Illustration. The first example shows in Fig. 3 the results of ranking a set of shapes by the properties of their axes. In the upper row the shapes are ordered according to linearity $\mathcal{L}$ while for comparison in the lower row the shapes are ordered according to Rosin's S_3 sigmoidality measure [21].

Second Experiment: Illustration. The second example shows how the linearity measure can be used as an error measure for polygonal approximation in the same manner as in [16]. For each curve segment the error (i.e. its deviation from linearity) is calculated as $(1 - \mathcal{L}(\mathbf{C}_{ij})) \cdot m_{0,0}(\mathbf{C}_{ij})$, where $\mathbf{C}_{ij}$ denotes the section of curve between $\mathbf{C}(i)$ and $\mathbf{C}(j)$, while $m_{0,0}(\mathbf{C}_{ij})$ is the length of $\mathbf{C}_{ij}$. The optimal polygonal approximation which minimises the summed error over the specified number of curve segments can then be determined using dynamic programming. Results are shown in Fig. 4 of polygonal approximations obtained using different error measures, namely the standard L_2 and L_∞ errors norms on the distances between the curve segment and the corresponding straight line segment as well as an error term based on the linearity measure ($\mathcal{L}_0$) in [16]. It can be seen that there are differences between the polygonal approximations produced by the various error terms, although they are relatively small except for very coarse approximations.

Third Experiment: Illustration. To demonstrate how various shapes produce a range of linearity values, Fig. 5 shows the first two samples of each handwritten digit (0–9) from the training set captured by Alimoğlu and Alpaydin [22] plotted in a $2D$ feature space of linearity $\mathcal{L}(\mathcal{C})$ versus rectilinearity $\mathcal{R}_1(\mathcal{C})$ [23].

Despite the variability of hand writing, most pairs of the same digit are reasonably clustered. The major separations occur for:

- "2" since only one instance has a loop in the middle;
- "4" since the instance next to the pair of "7"s is missing the vertical stroke;
- "5" since the uppermost right instance is missing the horizontal stroke.

The full data set consists of 7,485 digits for training and 3,493 digits for testing. A nearest neighbour classifier using Mahalanobis distances was trained on the training

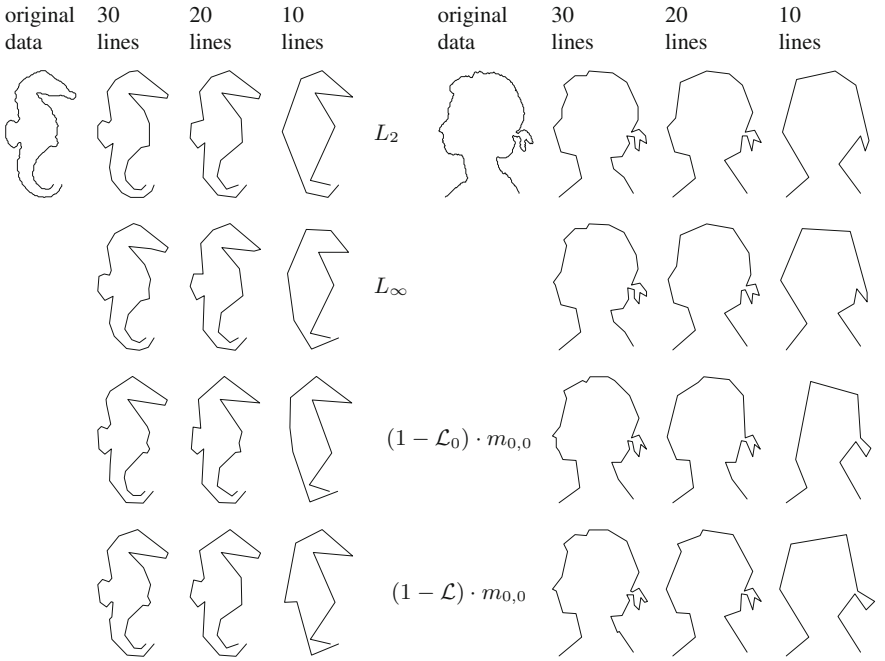

Fig. 4 Two shapes with their optimal polygonal approximations using the L_2 error function (*top row*), L_∞ error function (*second row*), the $(1 - \mathcal{L}_0) \cdot m_{0,0}$ error function (*third row*), and the $(1 - \mathcal{L}) \cdot m_{0,0}$ error function (*fourth row*). Approximations are determined using 30, 20 and 10 line segments

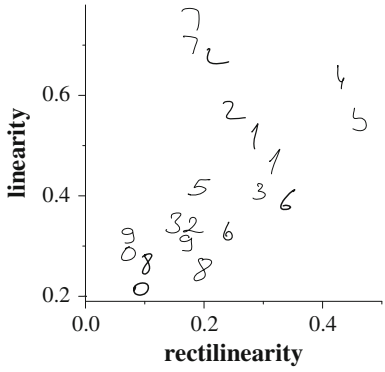

Fig. 5 Handwritten digits ordered by linearity and rectilinearity

data with just linearity as a single feature, and was applied to the test set, producing an accuracy of 21.39 %. This value is low since a richer feature set is required for discrimination. As a step towards obtaining this, the digits were simplified by applying Ramer's polygonal approximation [24] at various thresholds ({2, 4, 8, 16, 32}); some examples are shown in Fig. 6. Linearity was computed at each scale to

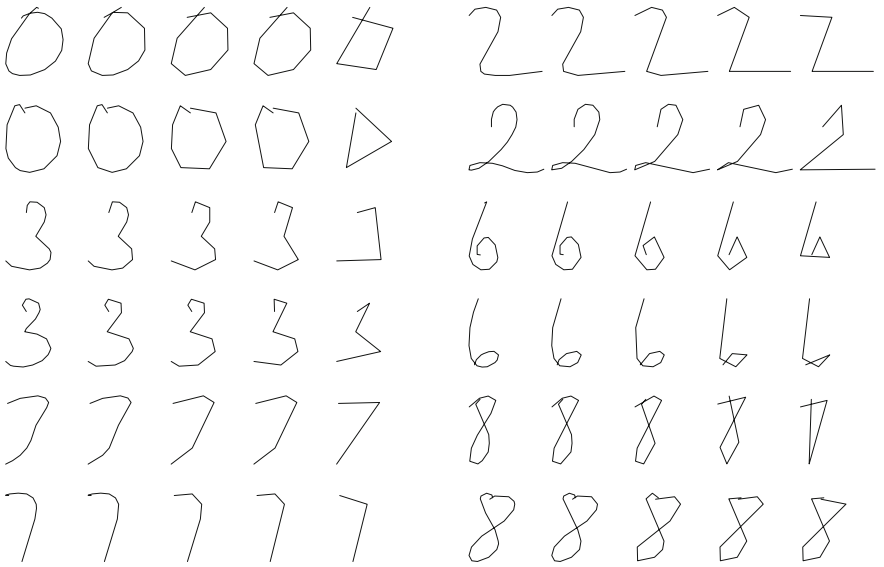

Fig. 6 Examples of digits simplified at five scales

obtain five features per digit, and this improved accuracy to 33.61 %. Finally, by further extending the feature vector to include the first seven Hu moment invariants [1] and six further moment invariants designed for character recognition [25], accuracy was increased to 88.38 %.

Fourth Experiment: Filtering Edges. Figure 7 shows the application of the linearity to filtering edges. The edges were extracted from the images using the Canny detector [26], connected into curves, and then thresholded according to total edge magnitude and length [27]. Linearity was measured in local sections of curve of length 25, and sections above (or below) a linearity threshold were retained. It can be seen that retaining sections of curve with $\mathcal{L}(\mathcal{C}) < 0.5$ finds small noisy or corner sections. Keeping sections of curve with $\mathcal{L}(\mathcal{C}) > 0.9$ or $\mathcal{L}(\mathcal{C}) > 0.95$ identifies most of the significant structures in the image.

Experiments are also shown in which Poisson image reconstruction is performed from the image gradients [28]. In the middle column of Fig. 8 all the connected edges with minimum length of 25 pixels (shown in the first column in Fig. 7) are used as a mask to eliminate all other edges before reconstruction. Some fine detail is removed as expected since small and weak edges have been removed in the pre-processing stage.

When linearity filtering is applied, and only edges corresponding to sections of curve with $\mathcal{L}(\mathcal{C}) > 0.95$ are used (see the fourth column in Fig. 7) then the image reconstruction (Fig. 8 right column) retains only regions that are locally linear structures (including sections of large curved objects).

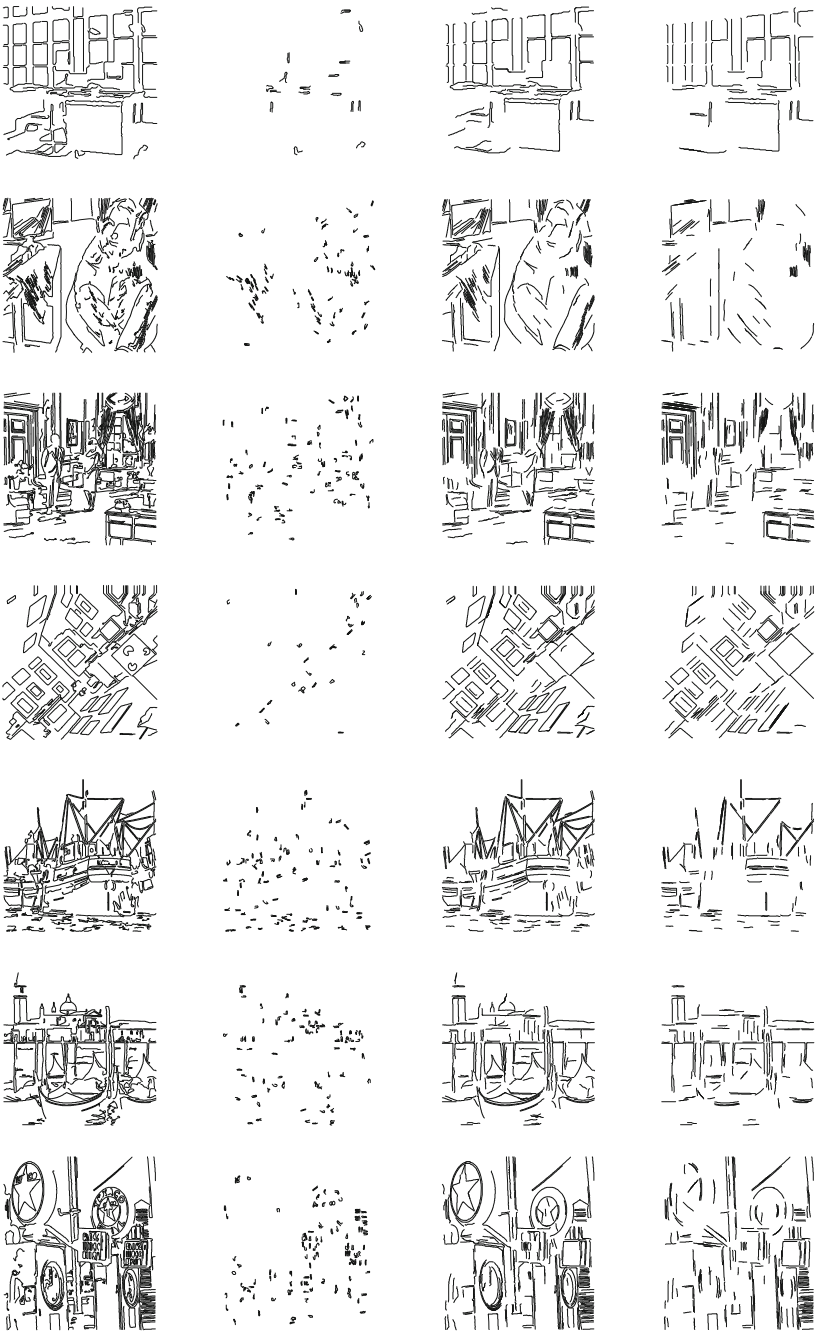

Fig. 7 Filtering connected edges by linearity. *Left/first column* connected edges (minimum length: 25 pixels); *second column* sections of curve with $\mathcal{L}(C) < 0.5$; *third column* sections of curve with $\mathcal{L}(C) > 0.9$; *fourth column* sections of curve with $\mathcal{L}(C) > 0.95$

Fig. 8 Reconstructing the image from its filtered edges. (*left column*) original intensity image; (*middle column*) image reconstructed using all connected edges (minimum length: 25 pixels); (*right column*) image reconstructed using sections of curve with $\mathcal{L}(\mathcal{C}) > 0.95$

Fig. 9 Examples of genuine *(first three columns)* and forged (*last three columns*) signatures

Fifth Experiment: Signature Verification. For this application we use data from Munich and Perona [29] to perform signature verification. The data consists of pen trajectories for 2,911 genuine signatures taken from 112 subjects, plus five forgers provided a total of 1,061 forgeries across all the subjects. Examples of corresponding genuine and forged signatures are shown in Fig. 9. To compare signatures we use the linearity plots defined by (2) and (3) to provide more information than a single linearity measurement. Linearity plot examples are in Fig. 10.

The quality of match between signatures $\mathcal{C}_1$ and $\mathcal{C}_2$ is measured by the similarity between the linearity plots $\mathrm{P}(\mathcal{C}_1)$ and $\mathrm{P}(\mathcal{C}_2)$. This similarity is measured by the area bounded by the linearity plots $\mathrm{P}(\mathcal{C}_1)$ and $\mathrm{P}(\mathcal{C}_2)$ and by the vertical lines $s = 0$ and

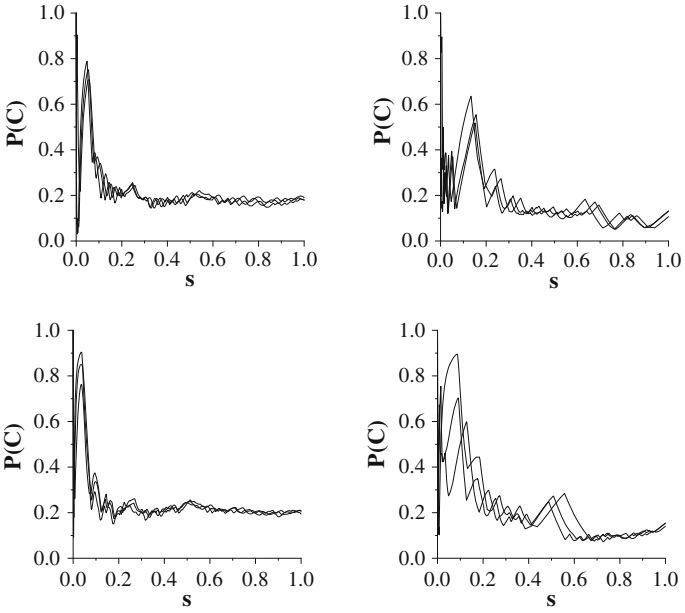

Fig. 10 Examples of linearity plots for the genuine signatures (*top row*) and the forged signatures (*bottom row*) in Fig. 9

$s = 1$. Figure 10 demonstrates the linearity plots for the signatures shown in Fig. 9. The plots in the first (second respectively) row contain the three genuine (forged respectively) signatures from Fig. 9.

Nearest neighbour matching is then performed on all the data using the leave-one-out strategy. Signature verification is a two class (genuine or fake) problem. Since the identity of the signature is already known, the nearest neighbour matching is only applied to the set of genuine and forged examples of the subject's signature. Computing linearity of the signatures using $\mathcal{L}(\mathcal{C})$ produces 96.9 % accuracy. This improves the results obtained by using the linearity measure defined in [16] which achieved 93.1 % accuracy.

5 Conclusions

This paper has described a new shape measure $\mathcal{L}(\mathcal{C})$ for computing the linearity of open curve segments. For a given unit length curve $\mathcal{C}$, its assigned linearity measure $\mathcal{L}(\mathcal{C})$ is computed as the sum of the distances of the end points of $\mathcal{C}$ to the centroid of $\mathcal{C}$. Of course, if the curve considered has an arbitrary length, then the assigned linearity measure is computed as the ratio of the sum of distances of the curve end

points to the curve centroid and the curve length. Such a defined open curve linearity measure $\mathcal{L}(\mathcal{C})$ satisfies the basic requirements for a linearity measure:

- $\mathcal{L}(\mathcal{C})$ is in the interval (0, 1];
- $\mathcal{L}(\mathcal{C})$ equals 1 only for straight line segments;
- $\mathcal{L}(\mathcal{C})$ is invariant with respect to translation, rotation and scaling transformations on the curve.

In addition $\mathcal{L}(\mathcal{C})$ is both extremely simple to implement and efficient to compute.

The effectiveness of the new linearity measure is demonstrated on a variety of tasks. Since the linearity measure $\mathcal{L}(\mathcal{C})$ is a single number, in order to increase the discrimination power in object classification tasks, based on a use of $\mathcal{L}(\mathcal{C})$, we have employed two methods. The first one is based on a use of *linearity plots*, where the quantity $\mathcal{L}(\mathcal{C})$ is replaced with a graph. The second one is based on an idea from [30]: (i) Several shapes (i.e. open curves) were computed from an object (i.e. digit curves in the presented example) by applying a tunable polygonal approximation algorithm; (ii) $\mathcal{L}(\mathcal{C})$ values, assigned to each of such shapes/curves, were used for the classification.

Acknowledgments This work is partially supported by the Serbian Ministry of Science and Technology/project III44006/OI174008.

References

1. Hu, M.: Visual pattern recognition by moment invariants. IRE Trans. Inf. Theory **8**(2), 179–187 (1962)
2. Bowman, E., Soga, K., Drummond, T.: Particle shape characterisation using Fourier descriptor analysis. Geotechnique **51**(6), 545–554 (2001)
3. Ruberto, C.D., Dempster, A.: Circularity measures based on mathematical morphology. Electron. Lett. **36**(20), 1691–1693 (2000)
4. Rahtu, E., Salo, M., Heikkilä, J.: A new convexity measure based on a probabilistic interpretation of images. IEEE Trans. Patt. Anal. Mach. Intell. **28**(9), 1501–1512 (2006)
5. Melter, R., Stojmenović, I., Žunić, J.: A new characterization of digital lines by least square fits. Pattern Recognit. Lett. **14**(2), 83–88 (1993)
6. Imre, A.: Fractal dimension of time-indexed paths. Appl. Math. Comput. **207**(1), 221–229 (2009)
7. Schweitzer, H., Straach, J.: Utilizing moment invariants and Gröbner bases to reason about shapes. Comput. Intell. **14**(4), 461–474 (1998)
8. Acketa, D., Žunić, J.: On the number of linear partitions of the (m, n)-grid. Inf. Process. Lett. **38**(3), 163–168 (1991)
9. Direkoglu, C., Nixon, M.: Shape classification via image-based multiscale description. Pattern Recognit. **44**(9), 2134–2146 (2011)
10. Manay, S., Cremers, D., Hong, B.W., Yezzi, A., Soatto, S.: Integral invariants for shape matching. IEEE Trans. Pattern Anal. Mach. Intell. **28**(10), 1602–1618 (2006)
11. Mio, W., Srivastava, A., Joshi, S.: On shape of plane elastic curves. Int. J. Comput. Vis. **73**(3), 307–324 (2007)
12. Stojmenović, M., Žunić, J.: Measuring elongation from shape boundary. J. Math. Imaging Vis. **30**(1), 73–85 (2008)

13. Gautama, T., Mandić, D., Hull, M.V.: A novel method for determining the nature of time series. IEEE Trans. Biomed. Eng. **51**(5), 728–736 (2004)
14. Gautama, T., Mandić, D., Hulle, M.V.: Signal nonlinearity in fMRI: a comparison between BOLD and MION. IEEE Trans. Med. Images **22**(5), 636–644 (2003)
15. Stojmenović, M., Nayak, A., Žunić, J.: Measuring linearity of planar point sets. Pattern Recognit. **41**(8), 2503–2511 (2008)
16. Žunić, J., Rosin, P.: Measuring linearity of open planar curve segments. Image Vis. Comput. **29**(12), 873–879 (2011)
17. Benhamou, S.: How to reliably estimate the tortuosity of an animal's path: straightness, sinuosity, or fractal dimension? J. Theoret. Biol. **229**(2), 209–220 (2004)
18. El-ghazal, A., Basir, O., Belkasim, S.: Farthest point distance: a new shape signature for Fourier descriptors. Signal Process. Image Commun. **24**(7), 572–586 (2009)
19. Zhang, D., Lu, G.: Study and evaluation of different Fourier methods for image retrieval. Image and Vision Computing **23**(1), 3349 (2005)
20. Valentine, F.: Convex Sets. McGraw-Hill, New York (1964)
21. Rosin, P.: Measuring sigmoidality. Pattern Recognit. **37**(8), 1735–1744 (2004)
22. Alimoğlu, F., Alpaydin, E.: Combining multiple representations for pen-based handwritten digit recognition. ELEKTRIK: Turk. J. Electr. Eng. Comput. Sci. **9**(1), 1–12 (2001)
23. Žunić, J., Rosin, P.: Rectilinearity measurements for polygons. IEEE Trans. Patt. Anal. Mach. Intell. **25**(9), 1193–3200 (2003)
24. Ramer, U.: An iterative procedure for the polygonal approximation of plane curves. Comput. Graph. Image Process. **1**, 244–256 (1972)
25. Pan, F., Keane, M.: A new set of moment invariants for handwritten numeral recognition. In: IEEE International Conference on Image Processing, pp. 154–158 (1994)
26. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **8**(6), 679–698 (1986)
27. Rosin, P.: Edges: saliency measures and automatic thresholding. Mach. Vis. Appl. **9**(4), 139–159 (1997)
28. Pérez, P., Gangnet, M., Blake, A.: Poisson image editing. ACM Trans. Graph. **22**(3), 313–318 (2003)
29. Munich, M., Perona, P.: Visual identification by signature tracking. IEEE Trans. Patt. Anal. Mach. Intell. **25**(2), 200–217 (2003)
30. Aktaş, M., Žunić, J.: A family of shape ellipticity measures for galaxy classification. SIAM J. Imaging Sci. **6**(2), 765–781 (2013)