RESEARCH   **Open Access**

Computer-assisted analysis of painting brushstrokes: digital image processing for unsupervised extraction of visible features from van Gogh's works

Fabrizio Lamberti, Andrea Sanna and Gianluca Paravati[*]

Abstract

The automatic extraction of objective features from paintings, like brushstroke distribution, orientation, and shape, could be particularly useful for different artwork analyses and management tasks. In fact, these features contribute to provide a unique signature of the artists' style and can be effectively used for artist identification and classification, artwork examination and retrieval, etc. In this paper, an automatic technique for unsupervised extraction of individual brushstrokes from digital reproductions of van Gogh's paintings is presented. Through the iterative application of segmentation, characterization, and validation steps, valid brushstrokes complying with specific area and shape constraints are identified. On the extracted brushstrokes, several representative features can then be calculated, like orientation, length, and width. The accuracy of the devised method is evaluated by comparing numerical results obtained on a dataset of digital reproductions of van Gogh's oil-on-canvas works with observations made by human subjects and with another recent approach for automatic brushstroke analysis. Experimental tests showed that the devised methodology produces results that are rather close to those obtained by human subjects and, for some of the metrics considered, can provide improved performances with respect to alternative techniques.

Keywords: Painting analysis; Brushstroke extraction; Visual judgements; Image processing; Unsupervised techniques; Automatic systems

1 Introduction

Analysis of paintings by art historians for appreciation, attribution, and conservation purposes has been supported by science and technology since the early eighteenth century [1]. As a matter of example, chemical tests have been exploited to assess pigment composition and compare paints used by different artists and in different periods, whereas dendrochronology and radiocarbon dating have been used for examining wooden panels and paint layers as well as for supporting the estimation of painting age.

In more recent years, image processing techniques (encompassing ultraviolet fluorescence, infrared reflectography, stereo microscopy, x-radiography, etc.) have

been proposed as further tools supporting the job of art experts and conservation specialists, e.g., to identify restoring interventions, to reveal underdrawings beneath the paint surface, and to provide insights into artists' techniques and intentions [2,3]. Despite their potential, the application of such methods, which rely on low-level features of paintings under analysis, has not gained yet a widespread adoption. Furthermore, the devised approaches do not, in general, allow researchers to escape from subjectivity and to achieve unambiguous results. Thus, they often lead to intense debates and raise criticism on the applicability of such technological supports in the considered domain [4].

Nonetheless, in the last decade, digital image processing has been experimented as an effective means for performing an in-depth visual analysis of higher-level features related to brushstroke distribution, orientation,

*Correspondence: gianluca.paravati@polito.it
Dip. di Automatica ed Informatica, Politecnico di Torino, C.so Duca degli Abruzzi 24, Torino I-10129, Italy

and shape. In fact, according to art experts, these features provide a unique 'signature' of the artist's style and can therefore be extremely useful in achieving objective and quantifiable results [5]. By automating a task that is, in general, extremely time-consuming when performed manually (e.g., by art experts), statistical measures could be gathered and exploited for different purposes, e.g., for supporting artist identification, for retrieving information from pictorial databases, etc.

Despite the relevance of the above aspects, the number of image processing techniques specifically targeted to art investigation is still quite limited and, although the approaches that have been presented in the literature achieved results that are very promising, they often lack a comprehensive evaluation. This is partially due to the fact that, in general, the testing datasets are not widely available [6]. Furthermore, the experimentation of new methods often requires overcoming the scepticism of involved actors towards the introduction of computer-assisted technologies in a field that is traditionally pervaded by a connoisseurship spirit founded on a profound knowledge of the history of art, on the understanding of the artist's body of work, on painting background, and on inner reactions to painting observation [7].

In this paper, a novel technique for unsupervised extraction of individual brushstrokes from digital reproductions of impressionist paintings is presented. The study focuses on oil-on-canvas works by Vincent van Gogh as, in this case, quantitative information concerning brushstroke characteristics plays an essential role in the identification and description of the artist's hand [8]. Specifically, the experimental dataset is made up of a number of image patches extracted from digital reproductions of van Gogh's works at different resolutions, which were normalized to the size of the lowest-resolution image. Some representative patches from the above dataset are reported in Figure 1.

Due to the lack of ground-truth data providing exact information about position and shape of the artist's brushstrokes, the performances of the devised technique have been evaluated by comparison with manually extracted information as well as with results obtained by another automatic technique that has been recently proposed by Li et al. [9]. In particular, numerical results in terms of brushstroke orientation, length, and width have been first extracted and compared with visual observations by non-expert human subjects by extending the metrics and the methodology adopted in [10]. Then, the same metrics have been exploited again on data used in [9] by comparing brushstrokes obtained by the method presented in this paper both with brushstrokes extracted by the technique proposed in the above work as well as with brushstrokes annotated manually by art experts identified by the same authors.

Experimental results show that the designed technique performs at a level that is comparable to that of visual investigation, thus providing a further demonstration of the applicability of image processing techniques in the considered domain. The proposed method can be regarded as an effective means to support different kinds of scenarios possibly requiring automatic brushstroke characterization. In fact, depending on the resolution of the images under consideration, numerical data could be possibly exploited in both artwork analysis processes and in knowledge management-based tasks, e.g., for painting image classification, search, compression, etc.

The remainder of the paper is organized as follows. In Section 2, the literature regarding the application of image processing methods to brushstroke analysis is reviewed. In Section 3, the proposed technique for unsupervised brushstroke extraction and characterization is described, whereas in Section 4, the strategy that has been developed in this work for gathering human judgements is presented. Section 5 illustrates the metrics that have been defined for assessing performances. Section 6 presents the procedure for algorithm configuration. Section 7 deals with the experimental setup and reports on the comparison of results obtained by the proposed method with human observations as well as with the outcomes of the alternative automatic technique considered. Finally, Section 8 concludes the paper by pointing out future developments that could be possibly envisaged for the current work.

2 Background

Image processing techniques for the automatic or semi-automatic analysis of paintings represent a valuable tool for supporting a number of challenging tasks, ranging from painting classification, comparison, and search and retrieval, to artist identification, artwork dating, and authentication [11].

Among these tasks, classification is aimed at assigning a given painting to a specific category based on the extracted features. Categories are useful to allow for a comparison of different types of works and to investigate the authorship as well as identify possible copies. For instance, a classification algorithm based on a stroke detector is presented in [12], where the ending shape of a painter's strokes is extracted for each painting considered. Stroke detectors can be also combined in order to complement their strengths. As a matter of example, in [13] two detectors, namely, a model-based stroke operator and a more generic multi-layer feedforward neural network, are used to study the characteristics of miniatures from the 18th and early 19th centuries. In [14], an artist classifier based on aggregate features is proposed to distinguish among works by Rembrandt, van Gogh, Picasso, Magritte, and Dalí. Features are extracted by globally analyzing painting patches rather than working at the level of
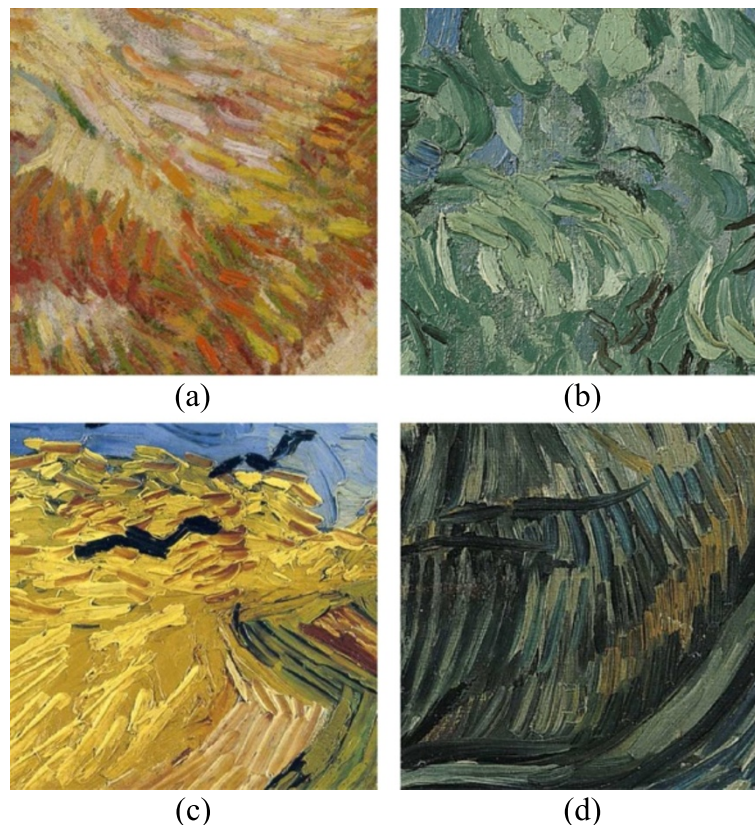
Figure 1 Patches of 512 × 512 pixels extracted from van Gogh paintings of the image set used for the study. These were used with the permission of the Van Gogh Museum Amsterdam (Vincent van Gogh Foundation). **(a)** *Self-portrait*, 1887 to 1888. **(b)** *Olive grove*, June 1889. **(c)** *Wheatfield with crows*, July 1890. **(d)** *A pair of leather clogs*, 1889.

single brushstrokes. The classification process is designed to assign to each patch a possible artist. Then, the final decision on painting authorship is made by applying a majority rule on artists assigned to each particular patch.

Most of the classification methods proposed in the literature need to analyze a huge amount of data. Thus, many algorithms include a simplification step that is devoted to narrowing the search space. In [15], a comprehensive analysis of pre-processing strategies aimed to cope with the above issue is reported.

As anticipated, classification algorithms are aimed at finding generic categories (e.g., based on artist, period, etc.) to describe the artworks under analysis. These categories can then be exploited to perform *ad hoc* searches in pictorial databases. However, experts of the sector and art historians may need to search for and retrieve paintings by using queries that are more elaborated with respect to those used by, e.g., the general public. To this aim, a framework for artwork retrieval based on semantic search is presented in [16]. The framework takes into account several information levels to meet the requirements of different kinds of users. Thus, in the considered approach,

the search categories can be identified as visual attributes (e.g., colors, composition, etc.) and art-historical information (e.g., artist, style, and period) as well as abstract concepts about the artwork (e.g., warm, expressive, etc.).

One of the most challenging applications where the automatic analysis of paintings can be applied into is artwork authentication. In this respect, several techniques have been proposed and tested on reproductions of van Gogh's painting. For instance, color and texture analysis is exploited in [17] to enhance parallel marks of brushstrokes. The devised algorithm proceeds by processing the images with a circular filter, filling closed contours, skeletonizing them, and finally fitting the thinned brushstrokes to a high order polynomial. The resulting polynomial coefficients for each brushstroke are plotted as coordinate values to create a cloud of points, whose structure is considered as capable of providing a feature (proof of authenticity) that is unique for a given artist. Similarly, the technique proposed in [18] aims at discriminating authentic drawings from copies.

Independently of the final application of the specific algorithms (i.e., classification, search and retrieval,

authentication, etc.), the most commonly adopted frameworks for the analysis of paintings are based on the extraction of global and local features. Algorithms extracting global features analyze patches as a whole and produce aggregate results, without providing detailed information for each brushstroke. Algorithms extracting local features work at the level of single brushstrokes by identifying attributes like length, width, orientation, etc.

Global features extraction is addressed in [10], where image processing methods encompassing band-pass filtering and multi-level thresholding are used to automatically determine the prevailing orientation of image patches extracted from van Gogh's works. Automatic results are compared with human observations as well as with the outcomes of other computer-based techniques. A method that is often adopted by the works dealing with global feature extraction is wavelet-based analysis. For instance, the approaches presented in [6] and [19] are aimed at measuring the stylistic proximity of a painting to a given artist through wavelet-based analysis and multi-fractal classification, which exploits information deriving from the analysis of image textures to describe the regularity of the area being considered. In these cases, brushstroke description includes attributes such as average length, mean curvature, and prevailing orientation. Among wavelet-based analysis methods, learning-based approaches can be identified. In general, learning-based approaches work with training images from which distinctive features are extrapolated. An example can be found in [20], where a metric to describe the similarity among artistic styles is developed based on statistical information extracted from a painting's background.

Other works address the issue of automatic analysis by extracting local information. For instance, in [21], the authors create digital signatures of paintings capable of distinguishing one artist from another by working with stroke detectors and hidden Markov models. Curve models are considered in [22] to describe brushstrokes, with the goal of defining a set of attributes used to animate Chinese artworks. The extraction process is based on segmentation, fitting, and refinement steps. Each single brushstroke is described by its shape (i.e., its contour) and texture information. Its curvature is obtained first by identifying a set of control points and, afterwards, by building up a smoothed skeleton by approximation with a curve, which is finally moved during the animation phase. Since in [22], the final aim is to apply to identified brushstrokes a motion effect rather than to study them, the requirements in terms of precision are not so tight, and a loose approximation of the real brushstrokes is considered as acceptable. The De-pict algorithm in [23] follows a different approach but still maintaining the objective of extracting individual brushstrokes. It is based on the analysis of consecutive brushstroke layers. In a first step,

clearly visible brushstrokes at the top of the painting are identified. Then, the algorithm tries to iteratively reconstruct the deeper layers by removing the topmost layer of identified brushstrokes. The fundamental assumption for the application of this methodology is that the underlying brushstrokes should be at least partially visible. This work has been afterwards refined in [24], with the introduction of a more accurate curve model for brushstroke description that takes into account also chromaticity values. Another work encompassing the extraction of individual brushstrokes is reported in [9]. Here, the interest is on finding attributes of van Gogh's paintings able to categorize them by time period. The methodology is based on the extraction of brushstrokes by means of a combination of edge detection and segmentation techniques. The authors extract and statistically analyze several features identified with the help of art historians in order to provide evidence of the unique brushstroke style of van Gogh. The methodology is capable of distinguishing paintings by the artist in different time periods from those of his contemporaries.

The work proposed in this paper moves in the direction of the approaches operating at the local level. In fact, as in [9], the designed algorithm aims to extract individual brushstrokes. However, the method in [9] exploits edge detection to identify brushstrokes, whereas the proposed technique relies on repeated thresholding and validation steps. A characteristic of edge-based techniques is that, because of texture subtleties, in some cases they could split originally homogeneous regions into sub-regions, thus causing extracted brushstrokes to be smaller than the ones of the artist. In this respect, results achieved by the method reported in this paper are compared with those obtained by applying the algorithm in [9], confirming the above behavior both by a qualitative and quantitative point of view. With respect to approaches based on global feature extraction, in this paper, the methodology presented in [10] for comparing automatic measures about patches' prevailing orientation with measures gathered by human subjects is extended to assess the performances of the proposed technique in terms of individual brushstrokes orientation, length, and width, showing a good agreement between automatically and manually obtained results.

3 Automatic brushstroke extraction

The proposed technique aims at isolating individual brushstrokes from grayscale image patches selected from paintings being considered.

The algorithm has been designed as an iterative procedure, where the shape of a brushstroke is refined in consecutive steps. The overall procedure is made up of four main phases, i.e., seeds generation, brushstroke segmentation, brushstroke characterization, and brushstroke

validation. Each phase is described in the following sub-sections, whereas the pseudo-code of the overall algorithm is reported in Algorithm 1.

Algorithm 1: Pseudo-code of the proposed algorithm

Data: *image patch*
Result: *extracted brushstrokes*
1. *seeds* = generate set of seeds (*image patch*);
2. **forall the** *seeds* **do**
    initialize threshold (T);
    valid brushstroke = false;
    while *!valid brushstroke* **do**
        valid area = false;
        while *!valid area* **do**
            candidate brushstroke = segmentation (*seed, T*);
            valid area = check area (*candidate brushstroke*);
            if *!valid area* **then**
                update threshold (T);
            end
        end
        3. compute orientation (*candidate brushstroke*);
        compute length (*candidate brushstroke*);
        compute width (*candidate brushstroke*);
        4. *valid brushstroke* = check shape (*candidate brushstroke*);
    end
end

According to the algorithm, a number of points, referred to as *seeds*, specify the initial locations where brushstrokes should be searched for in the image patch under consideration. Seeds are used to compute the so called *candidate brushstrokes*, i.e., brushstrokes that need to be further processed in order to be possibly validated by the algorithm itself based on area and shape constraints. If the candidate brushstroke can be validated, its main features (namely, orientation, width, and length) are extracted and stored for later use. Otherwise, an iterative procedure tries in consecutive steps to automatically adjust and validate again the brushstroke. If the procedure fails, the brushstroke is finally discarded and the next seed is considered.

3.1 Seed generation

As anticipated, the seed generation phase aims at creating a set of points in the image patch that will represent the starting points for the segmentation-based candidate brushstroke extraction phase. Seeds are generated in a quasi-random fashion by using the Sobol algorithm [25], thus achieving an even coverage of the image space.

3.2 Brushstroke segmentation

The brushstroke segmentation phase takes as input each initial point defined in the seed generation phase and aims at isolating a candidate brushstroke, i.e., a homogenous region around that point by means of a segmentation technique. Qualitative results achieved by processing a single seed are shown in Figure 2, whereas results obtained with multiple seeds on a given patch belonging to the considered dataset are illustrated in Figure 3.

Starting from seed coordinates, neighboring pixels are visited by exploiting a region-growing-based approach with a variable threshold T, which is initialized at a pre-defined value[a]. All the pixels in the immediate neighborhood of the seed are visited to check if their intensity value falls in the range from $I-T$ to $I+T$ (where I is the intensity value of the seed). Each pixel satisfying this requirement is added to the segmented region. The procedure is repeated for each pixel added to the segmented region. The output of the region growing procedure represents a candidate brushstroke.

The region of the candidate brushstroke produced using the above approach is considered against *ad hoc* minimum and maximum *area constraints*, which are controlled by variables r_m and r_M, respectively, and can be expressed as

$$\delta^2 \times r_m < \sigma_r < \delta^2 \times r_M \tag{1}$$

where σ_r represents the area of the region and δ indicates the number of dot per inch (dpi) of the image under consideration. In this work, image patches considered have been normalized to a common resolution of 86.1 dpi. Hence, unless otherwise stated, it will be $\delta = 86.1$. It can be easily observed that the effect of (1) strongly depends on the value assigned to coefficients r_m and r_M, whose configuration will be discussed in Section 6.

The output of the described procedure is a binary map indicating which pixels belong to the candidate brushstroke. If the area constraints are not satisfied, the threshold T is automatically decreased and region growing is performed again until a valid region is found.

Figure 4 exemplifies the above process. Starting from a high value of T, the region-growing procedure selects an initial number of adjacent pixels (Figure 4a). Typically, this process generates an irregular shape that might comprise various brushstrokes. The region found in this step is validated against the area constraints. If the constraints are not satisfied, a new value for the threshold is computed (i.e., T is reduced) and a new iteration begins (Figure 4b,c,d,e,f,g,h,i). This process ends when either a candidate brushstroke with an area within the given range is identified (Figure 4j) or no valid region is found, i.e., the area of the candidate brushstroke is too small or T cannot be further reduced. In the latter case, the extracted brushstroke and the current seed are discarded. When a candidate brushstroke with an acceptable area is found
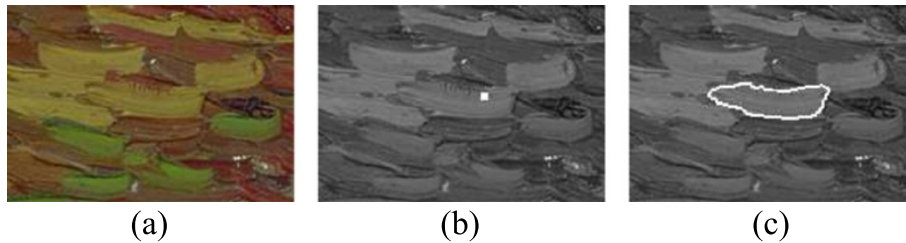
Figure 2 Unsupervised brushstroke extraction. (a) Portion of an image patch in the dataset extracted from the van Gogh painting *The garden of Saint-Paul's Hospital*, November 1889, used with permission of the van Gogh Museum Amsterdam (Vincent van Gogh Foundation). **(b)** Grayscale image and location of a possible seed (e.g., randomly selected). **(c)** Extracted brushstroke.

(e.g., Figure 4j), possible holes in the shape need to be filled. To this aim, a closing morphological operator, consisting in a dilation step followed by an erosion phase, is applied to the extracted brushstroke.

3.3 Brushstroke characterization

The characterization phase is aimed at obtaining relevant information for each candidate brushstroke. Although many parameters and attributes could be taken into account (e.g., viscosity, texture subtleties, occlusions, color distribution, etc.), the proposed algorithm focuses on shape features. In particular, as said, candidate brushstrokes are described by means of quantitative information about their orientation, length, and width.

Based on the analysis of the area covered, the candidate brushstroke shape is approximated by two line segments indicating its main orientation (length) and thickness (width). Segments are calculated by applying a least square fitting technique [26]. In particular, the algorithm computes the ellipse that fits the selected region, and the major and minor axes found are assumed to represent the length and width of the brushstroke.

Information extracted is also used in the next step of the algorithm for validation purposes. As it will be shown in the following, the approach pursued proved to be capable of supporting a quite accurate extraction of brushstrokes constituting the paintings under analysis.

3.4 Brushstroke validation

Once the characterization has been completed, it is possible that some of the candidate brushstrokes extracted by applying the methodology presented in the previous sub-sections show an irregular shape (e.g., due to over- or under-segmentation). Thus, a validation step needs to be performed in order to distinguish between valid and non-valid brushstrokes. To this aim, features extracted in the characterization phase are used to design two *ad hoc* validation rules (or *shape constraints*, using the terminology adopted in Section 3.2), which need to be satisfied in order for a candidate brushstroke to be considered as valid.

The first validation rule is based on the assumption that a generic brushstroke can be approximated with an elliptical shape. In this way, an irregular shape (e.g., made up of two blended brushstrokes) can be detected by roughly
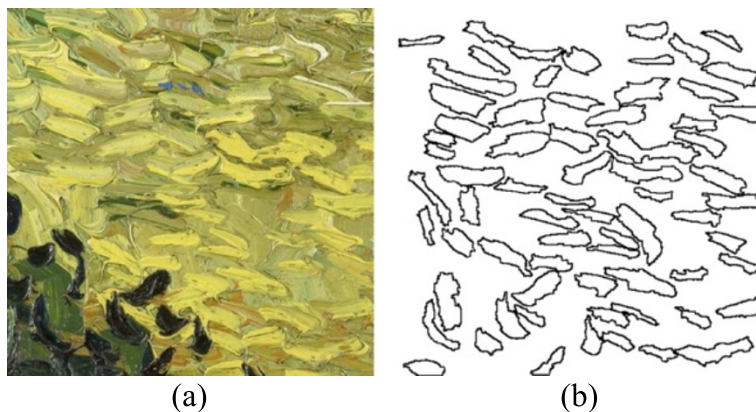
Figure 3 Experimental results. (a) Patch of 512 × 512 pixels extracted from the van Gogh painting *Landscape at twilight*, June 1890, used with permission of the Van Gogh Museum Amsterdam (Vincent van Gogh Foundation). **(b)** Valid brushstrokes obtained by running the proposed technique on 150 randomly selected seeds.
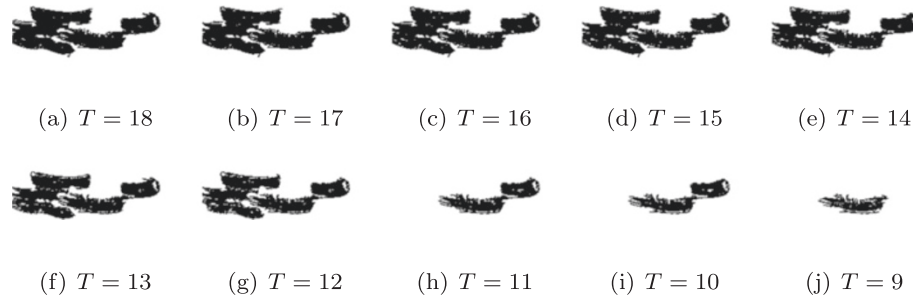
Figure 4 Segmentation phase. (a-j) Intermediate results (before the application of morphological transformations) produced by the iterative application of the segmentation step on the patch (and seed) in Figure 2.

comparing its covered area with respect to that of the ellipse approximating the brushstroke itself (as identified in the characterization phase). Thus, the first shape-based validation rule can be written as

$$(1 - r_e) \times \sigma_e < \sigma_r < (1 + r_e) \times \sigma_e \qquad (2)$$

where r_e is a fractional coefficient and σ_e represents the area of the ellipse.

The second validation rule takes into account the ratio between the length and the maximum width of the region. In order to be classified as valid, the shape of the candidate brushstroke should be comprised within a range of aspect ratios. In mathematical terms, the region should comply with the following constraint

$$\phi_r < r_w \times \lambda_r \qquad (3)$$

where r_w is again a fractional coefficient, whereas λ_r and ϕ_r represent the length and width of the region, respectively.

As for area-based segmentation constraints in (1), the effectiveness of shape constraints used in the validation phase and defined by (2) and (3) depends on two specific variables, namely r_e and r_w. These parameters need to be set up before running the algorithm. Since they are not known *a priori*, a calibration phase had to be designed. Details of the calibration phase as well as about the influence of the above parameters on algorithm performances will be discussed in Section 6.

4 Manual brushstroke extraction

Since a ground-truth does not exist for the considered dataset, the evaluation of performances of the proposed methodology needs to be carried out against human judgements or against the results produced by other algorithms.

Concerning human judgments, in this work, both expert and non-expert subjects were considered by working with two different models for manual brushstroke extraction. In the first model, subjects were asked to approximate a brushstroke by indicating its orientation, length, and width. In the second model, brushstrokes were described by their manually extracted contour. The first model was adopted in the experimental tests reported in this work by a set of subjects encompassing students and teachers from a computer graphics course in the MS degree in Computer Engineering. The second model was used by the art experts identified by Li et al. to generate the results reported in [9].

This section presents the procedure for gathering data according to the first model. Details concerning the second model can be found directly in [9].

In order to enable visual inspection and let non-expert human subjects approximate brushstrokes according to the first model, an *ad hoc* application was developed as a plug-in for the ImageJ image processing program [27]. Through the plug-in, images under analysis could be loaded and presented on the computer screen, together with seeds to be considered.

During manual analysis, selected human subjects were first requested to use the plug-in to specify the points corresponding to the beginning and the end of the line segment better characterizing the orientation of the brushstroke identified by the particular seed. When the two points had been specified, the line connecting the points was superimposed on the image in order to allow the subjects to possibly adjust their judgements. Data gathered in this step were used to obtain orientation and length measures. Subjects were then requested to configure line thickness in order to define a rectangular region roughly approximating the brushstroke shape. Finally, subjects were allowed to refine the region width through a number of control handles located on the boundaries of the rectangular shape. In this way, a better approximation of the actual brushstroke width was obtained.

As it will be shown in the following, measures obtained during the manual brushstroke extraction phase could be directly correlated with information extracted in the characterization phase. The steps of a typical judgement experiment are illustrated in Figure 5.
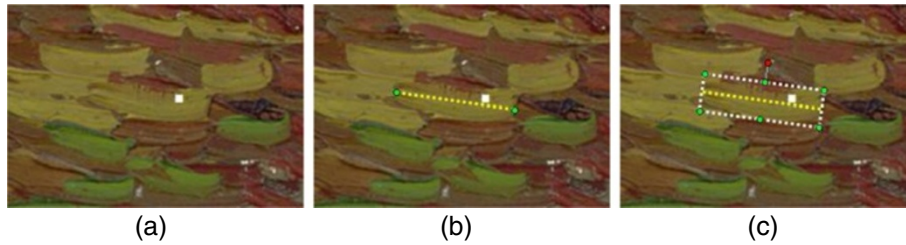
Figure 5 Gathering of human judgements. (a) Part of an image patch presented to the human subject (including seed location, which actually identifies the brushstroke to be analyzed). **(b)** Line segment positioned by the human subject. **(c)** Adjustment of shape orientation and width.

5 Criteria for evaluating performances

This section provides detailed information about the set of metrics that have been used to perform a quantitative comparison of the results obtained by the devised method with observations by non-expert subjects, as well as with judgments by expert subjects and with the outcomes of the automatic tool proposed by Li et al. [9].

A first set of metrics was derived from the methodology proposed in [10] (hence, in the following, the same notation will be used). Specifically, to compare the accuracy of the orientation measure computed by the particular algorithm considered on brushstroke i with respect to the corresponding observation by the human subject j, the angular distance $D_\alpha(i,j)$ has been defined as the smallest angle between the two orientations. As a consequence, this metric produces an angular distance confined in the interval 0° to 90°.

In order to assess algorithm performances in obtaining orientation estimates with respect to a given subject j, all the N brushstrokes have been considered and the mean square angular distance between automatic results and human judgements was computed as

$$\text{MSD}_\alpha(j) = \frac{1}{N}\sum_{i=1}^{N} D_\alpha(i,j)^2 \qquad (4)$$

Furthermore, to obtain an indication of the overall algorithm performances with respect to orientation measurements, the average MSD_α was computed by considering the contribution of the observations by all the m human subjects.

For assessing algorithm accuracy with respect to length measures, the mean square percentage distance between automatically and manually estimated values for the subject j has been computed as

$$\text{MSPD}_\lambda(j) = \frac{1}{N}\sum_{i=1}^{N} D_\lambda(i,j)^2 \qquad (5)$$

In (5), $D_\lambda(i,j)$ takes into account the relative distance between measured lengths for the i-th brushstroke and it is defined as

$$D_\lambda(i,j) = \frac{\lambda_A(i) - \lambda(i,j)}{\lambda_A(i)} \qquad (6)$$

where $\lambda_A(i)$ represents the length of the i-th brushstroke computed by the algorithm, whereas $\lambda(i,j)$ is the manual judgement by the j-th subject.

Similarly, accuracy in estimating width for subject j has been computed by making reference to the mean square percentage distance computed as

$$\text{MSPD}_\phi(j) = \frac{1}{N}\sum_{i=1}^{N} D_\phi(i,j)^2 \qquad (7)$$

with

$$D_\phi(i,j) = \frac{\phi_A(i) - \phi(i,j)}{\phi_A(i)} \qquad (8)$$

In (8), $\phi_A(i)$ is the automatically-computed width of the i-th brushstroke, whereas $\phi(i,j)$ represents the corresponding observation provided by the human subject j.

As for orientation measures, average MSPD_λ and MSPD_ϕ were calculated in order to get an indication of the overall algorithm performances concerning length and width judgements, independently of the specific subject (and brushstroke) being considered.

It is worth remarking that by computing the square root of expressions in (4), (5), and (7), $\text{RMSD}_\alpha(i)$, $\text{RMSPD}_\lambda(i)$, and $\text{RMSPD}_\phi(i)$ could be obtained, which convey an immediate indication of the error (in degrees) and of the relative error (as a percentage) in estimating brushstroke orientation, length, and width, respectively.

For the sake of comparability with the technique by Li et al., a further set of metrics was defined with the aim of quantifying the ability of a given automatic technique to produce brushstrokes adequately 'covering' and 'covered by' manually extracted ones. The new metrics were designed by extending those defined in [9] by using the same notation.

In particular, B_i, with $i = 1,\ldots,n$, is assumed to identify the n brushstrokes that are found by the automatic algorithm for each considered patch, whereas B_j^*, with $j = 1,\ldots,m$, are the m brushstrokes that have been

marked by the human subjects. B_i and B_j^* are the sets of pixel coordinates in the brushstroke (extracted automatically and manually, respectively). $B_i \cap B_j^*$ is the set of overlapped pixels in the considered brushstrokes. If the overlap between the brushstrokes is greater than 80%, i.e., $\left| B_i \cap B_j^* \right| / |B_i| > 80\%$, then B_i is defined as *covered*. When B_i is covered by B_j^*, $C_{i,j}$ is set to 1 (otherwise, to 0). $C_{i,.} = \sum_{j=1}^{m} C_{i,j}$, with $C_{i,.} \in \{0, 1\}$, indicates whether B_i is covered by any manual brushstroke. $C_{.,j} = \sum_{i=1}^{n} C_{i,j}$ is the number of automatically extracted brushstrokes that are covered by manual brushstroke B_j^*.

Based on the above notation, in order to compare automatically and manually obtained results, the authors of [9] defined the following metrics.

- *Valid rate*: $\sum_{i=1}^{n} C_{i,.}/n : C_{i,.} = 1$, the percentage of automatically extracted brushstrokes that are covered by one manual brushstroke.
- *Detection rate*: $\sum_{j=1}^{m} C_{.,j}/m : C_{.,j} \geq 1$, the percentage of manual brushstrokes that cover at least one automatic brushstroke.

A possible limit of the above metrics is that they could label a manual brushstroke as detected even if only a very small portion of it is covered by an automatically extracted one. Moreover, the validity criterion for automatic brushstrokes does not account for their representativeness. Hence, in this work, two other metrics were defined by simply inverting the above notation.

- *Fill rate*: $\sum_{j=1}^{m} C_{.,j}/m : C_{.,j} = 1$, the percentage of manually extracted brushstrokes that are covered by one automatic brushstroke.
- *Representativeness rate*: $\sum_{i=1}^{n} C_{i,.}/n : C_{i,.} \geq 1$, the percentage of automatic brushstrokes that cover at least one manual brushstroke.

The newly introduced metrics are expected to provide information that is complementary to that in [9] by characterizing the capability of the automatic technique to effectively cover manually extracted brushstrokes.
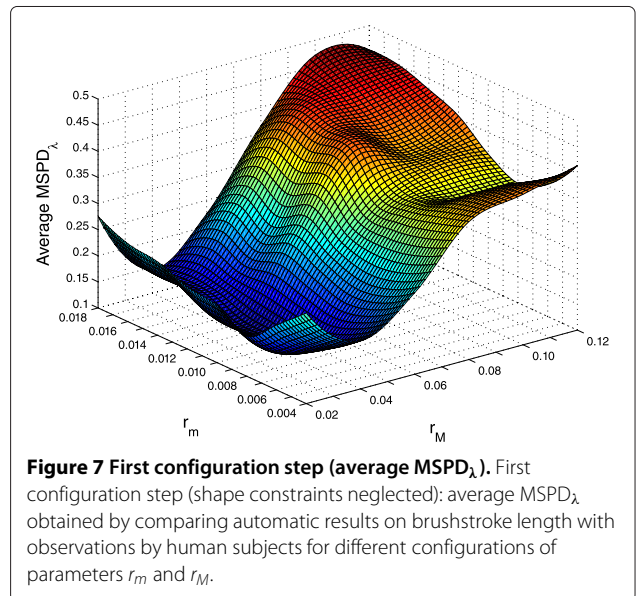
6 Configuration

The proposed characterization method is semi-automatic, meaning that a few parameters should be set before running the algorithm to extract individual brushstrokes. To this aim, a kind of calibration step has been added to the described methodology. In particular, the calibration concerns two parameters for describing the area constraints in the segmentation phase (Section 3.2) and two parameters for describing the shape constraints in the validation phase (Section 3.4). The goal of the calibration step is to find out a configuration of such parameters that minimizes the differences between the results gathered by the
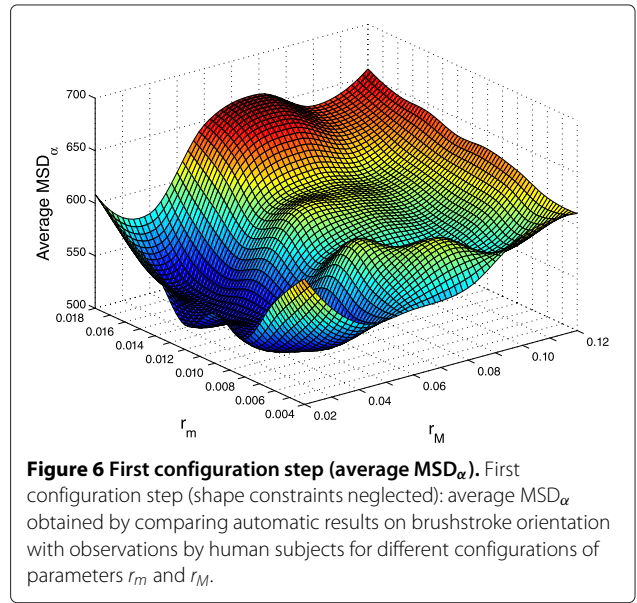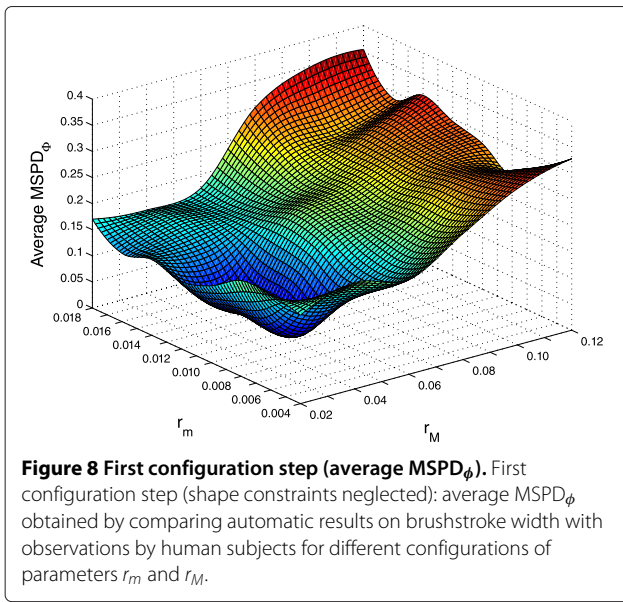

Figure 6 First configuration step (average MSD$_\alpha$). First configuration step (shape constraints neglected): average MSD$_\alpha$ obtained by comparing automatic results on brushstroke orientation with observations by human subjects for different configurations of parameters r_m and r_M.

automatic procedure and those obtained with the human judgements in terms of orientation, length, and width for the considered brushstrokes. In this way, the outcomes of the calibration step drive the choice of the constraints and influence the brushstrokes extraction procedure.

The parameter space represented by the area and shape constraints was systematically explored in order to find the configuration providing the highest accuracy. For this, a set composed by five image patches of 512×512 pixels each was randomly extracted from several digital reproductions of van Gogh's works provided by the van Gogh Museum, Amsterdam (Vincent van Gogh Foundation). From each patch, 30 seeds were randomly selected and

Figure 7 First configuration step (average MSPD$_\lambda$). First configuration step (shape constraints neglected): average MSPD$_\lambda$ obtained by comparing automatic results on brushstroke length with observations by human subjects for different configurations of parameters r_m and r_M.

Figure 8 First configuration step (average MSPD$_\phi$). First
configuration step (shape constraints neglected): average MSPD$_\phi$
obtained by comparing automatic results on brushstroke width with
observations by human subjects for different configurations of
parameters r_m and r_M.

presented to three human subjects for visual examina-
tion. Measures collected through the plug-in illustrated in
Section 4 were stored for later use. Selected points were
then analyzed through the proposed automatic technique.

At first, the parameters controlling, respectively, the
minimum and maximum area constraints r_m and r_M in
(1) were set as the ones minimizing the distances with
respect to human observations (metrics on brushstroke
orientation, length, and width presented in Section 5).
Fractional parameters r_m and r_M were varied in steps of

0.001 in the range]0, 0.02] and in steps of 0.01 in the range
]0, 0.2], respectively. The couple of parameters providing
the minimum mean square (percentage) distances were
selected.

Numerical results in a representative range for mini-
mum and maximum area thresholds ($0.004 \leq r_m \leq 0.018$
and $0.02 \leq r_M \leq 0.12$) are plotted in Figures 6, 7 and 8.
In particular, the average MSD$_\alpha$ is illustrated in Figure 6,
whereas average MSPD$_\lambda$ and MSPD$_\phi$ are reported in
Figures 7 and 8, respectively.

The configuration providing the best trade-off for all the
observations was the one represented by $r_m = 0.01$ and
$r_M = 0.04$, as shown in Figures 6, 7 and 8. Under the above
conditions, the computed metrics were MSD$_\alpha = 508.41$,
MSPD$_\lambda = 0.129$, and MSPD$_\phi = 0.074$.

After having identified the best configuration for the
parameters describing the area constraints, the parame-
ter space represented by shape constraints was explored.
In this case, the parameters controlling the brushstroke
aspect ratio r_w in (2) and the brushstroke area with respect
to the area of a reference brushstroke shape r_e in (3)
were set. Like before, the aim was to minimize the dis-
tances with respect to human observations. The value
of r_w was varied in the range]0, 1] (in steps of 0.10),
whereas r_e was allowed to assume values in the range
]0, 0.50] (in steps of 0.05). Parameters r_m and r_M were
fixed to the values identified in the previous configuration
step.

Numerical results in the ranges $0.15 \leq r_e \leq 0.40$
and $0.20 \leq r_w \leq 0.80$ are plotted in Figures 9, 10 and
11. In particular, average MSD$_\alpha$ is shown in Figure 9,
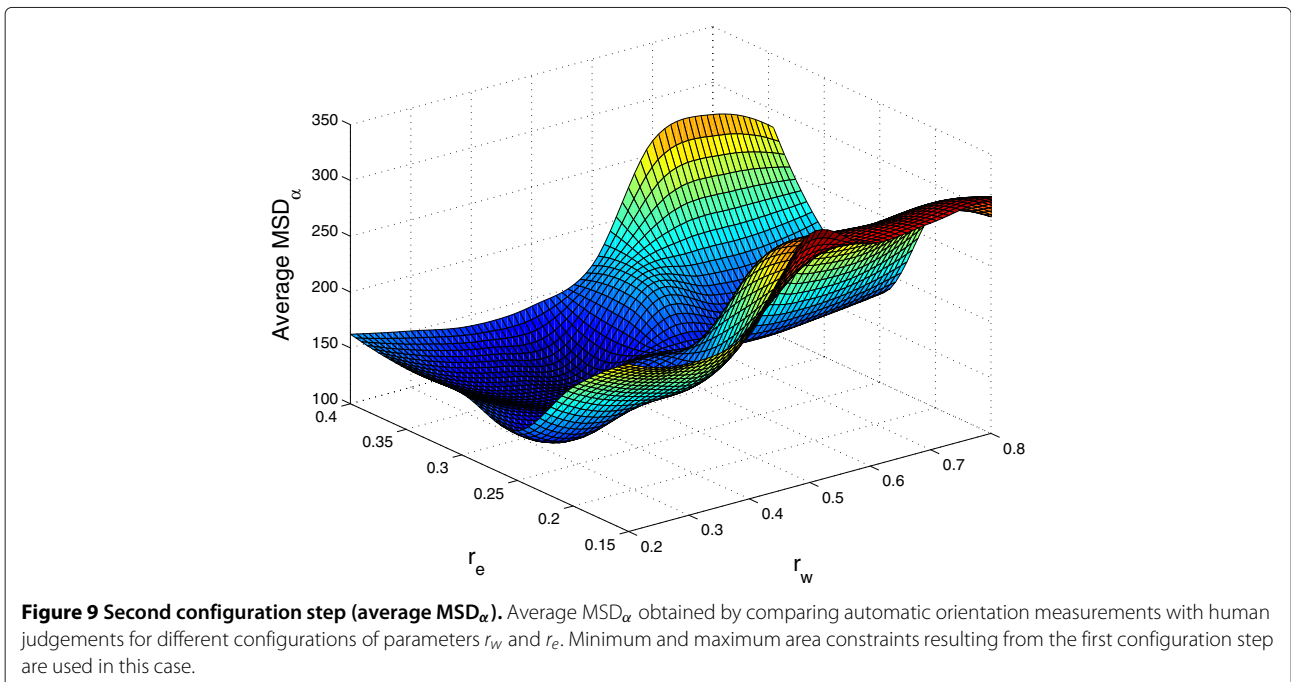
Figure 9 Second configuration step (average MSD$_\alpha$). Average MSD$_\alpha$ obtained by comparing automatic orientation measurements with human
judgements for different configurations of parameters r_w and r_e. Minimum and maximum area constraints resulting from the first configuration step
are used in this case.
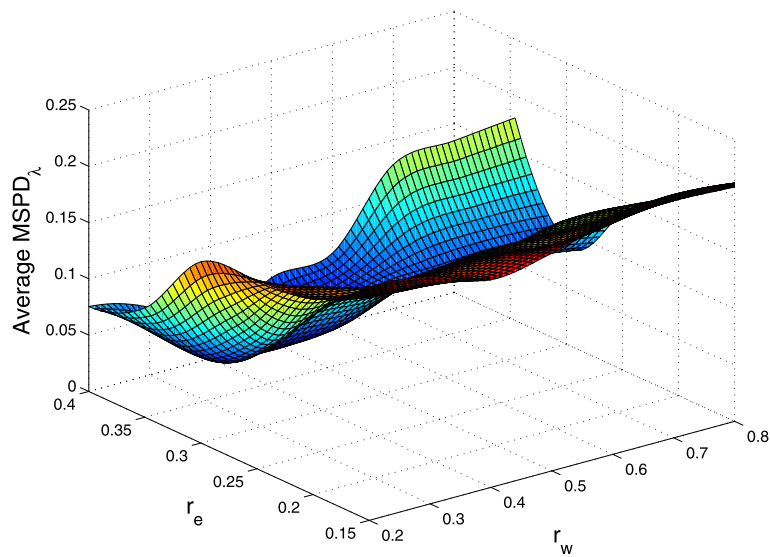
Figure 10 Second configuration step (average MSPD$_\lambda$). Average MSPD$_\lambda$ obtained by comparing automatic length measurements with human judgements for different configurations of parameters r_w and r_e. Minimum and maximum area constraints resulting from the first configuration step are used in this case.

whereas average MSPD$_\lambda$ and MSPD$_\phi$ are illustrated in Figures 10 and 11, respectively. The analysis of data shown in Figures 9, 10 and 11 provided again the best trade-off for fixing the shape constraint parameters. In particular, the choice of $r_e = 0.35$ and $r_w = 0.40$ resulted in the highest accuracy for all the observations.

As expected, the introduction of shape constraints had the additional effect of generally improving the accuracy

of automatic measurements. In fact, under such conditions, the computed metrics were MSD$_\alpha$ = 108.76, MSPD$_\lambda$ = 0.025, and MSPD$_\phi$ = 0.047 (i.e., lower than those obtained by enforcing only area constraints).

7 Experimental results

In this section, quantitative results obtained by the proposed technique are analyzed and compared against
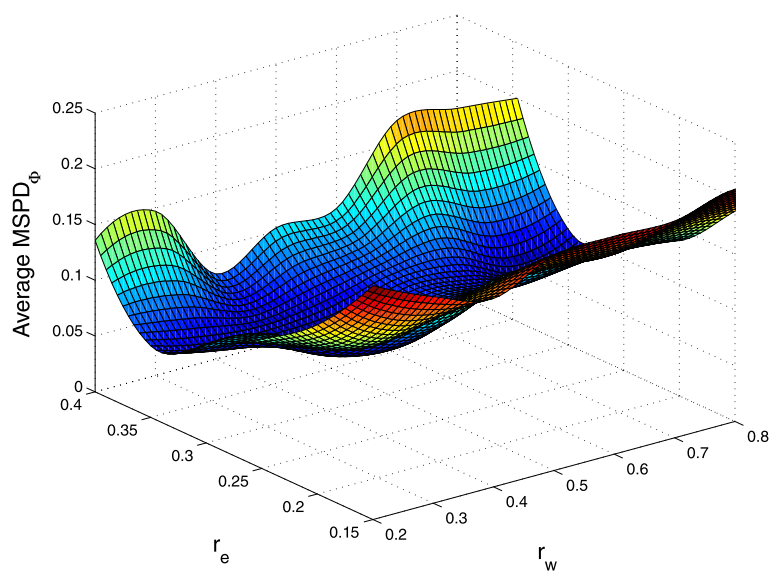
Figure 11 Second configuration step (average MSPD$_\phi$). Average MSPD$_\phi$ obtained by comparing automatic width measurements with human judgements for different configurations of parameters r_w and r_e. Minimum and maximum area constraints resulting from the first configuration step are used in this case.

numerical judgements by human subjects as well as with respect to the achievements of the automatic method in [9].

Firstly, an evaluation was carried out by considering the results produced by using the proposed technique and observations by non-expert human subjects by working on 20 patches extracted from 10 works of the considered dataset. For each patch, automatic processing was performed on 25 randomly selected seeds. The application of the designed method resulted in the extraction of 375 valid brushstrokes. On average, at the considered resolution of 86.1 dpi, the automatic extraction process of one brushstroke took about 460 ms on a standard PC equipped with an Intel Pentium dual-core CPU @ 1.73 GHz and 1 GB of DDR2 RAM.

Corresponding seeds were presented to ten human subjects for visual analysis. Judgements concerning brushstroke characteristics were collected using the tool illustrated in Section 4.

Numerical results obtained by human subjects were then used as a reference for assessing the accuracy of automatic measurements. To this aim, data obtained through the proposed brushstroke extraction and characterization method were cross-compared to visual observations by using the metrics on orientation, length, and width derived from [10] and reported in Section 5. Results for each of the three metrics above are illustrated in Tables 1, 2 and 3, respectively.

In Table 1, the second column reports the mean square angular distance $\text{MSD}_\alpha(j)$ for any given subject j (indicated in the first column). Similarly, in Tables 2 and 3, the second column presents the mean square percentage distance for length $\text{MSPD}_\lambda(j)$ and width $\text{MSPD}_\phi(j)$ for each subject. The third column of each table lists

the Pearson product-moment correlation coefficients r_{ah} between measures produced by the proposed automatic technique and judgements by the selected human subjects. The fourth column gives a rough indication of the agreement between human subjects. Since no ground-truth was available, the correlation coefficient $r_{\text{hs\_mean}}$ has been computed with respect to the mean of all human judgements for each valid brushstroke extracted. Finally, the last row of each table shows the performances of the proposed algorithm with respect to the mean of human judgments.

By analyzing the results in Tables 1, 2, and 3, it can be observed that the proposed algorithm performs at a level that is rather close to that of human subjects. The

Table 2 Comparison of length measures

Subject (*j*)	MSPD$_\lambda$ (*j*)	r_{ah}	$r_{\text{hs\_mean}}$
01	0.01	0.96	0.97
02	0.01	0.96	0.96
03	0.01	0.96	0.96
04	0.07	0.96	0.96
05	0.01	0.97	0.97
06	0.07	0.97	0.97
07	0.08	0.97	0.96
08	0.07	0.97	0.97
09	0.07	0.97	0.97
10	0.06	0.97	0.97
Proposed algorithm	-	-	0.99

Agreement between automatic results and visual observations in terms of mean square percentage distance (second column) and correlation between automatic judgements and visual observations for any given human subject *j* (third column) and among subjects (fourth column).

Table 1 Comparison of orientation measures

Subject (*j*)	MSD$_\alpha$ (*j*)	r_{ah}	$r_{\text{hs\_mean}}$
01	81.25	0.94	0.93
02	103.77	0.94	0.94
03	142.95	0.89	0.88
04	91.09	0.92	0.92
05	115.67	0.91	0.88
06	107.28	0.92	0.95
07	128.39	0.91	0.93
08	113.74	0.93	0.95
09	106.84	0.90	0.91
10	98.23	0.91	0.95
Proposed algorithm	-	-	0.89

Agreement between automatic results and visual observations in terms of mean square distance (second column) and correlation between automatic results and visual observations for any given human subject *j* (third column) and among human subjects (fourth column).

Table 3 Comparison of width measures

Subject (*j*)	MSPD$_\phi$ (*j*)	r_{ah}	$r_{\text{hs\_mean}}$
01	0.14	0.86	0.85
02	0.12	0.87	0.88
03	0.16	0.85	0.89
04	0.14	0.85	0.89
05	0.13	0.87	0.88
06	0.19	0.83	0.89
07	0.14	0.85	0.90
08	0.16	0.85	0.90
09	0.14	0.88	0.91
10	0.16	0.83	0.89
Proposed algorithm	-	-	0.95

Agreement between automatic results and visual observations in terms of mean square percentage distance (second column) and correlation between automatic results and visual observations for any given human subject *j* (third column) and among subjects (fourth column).
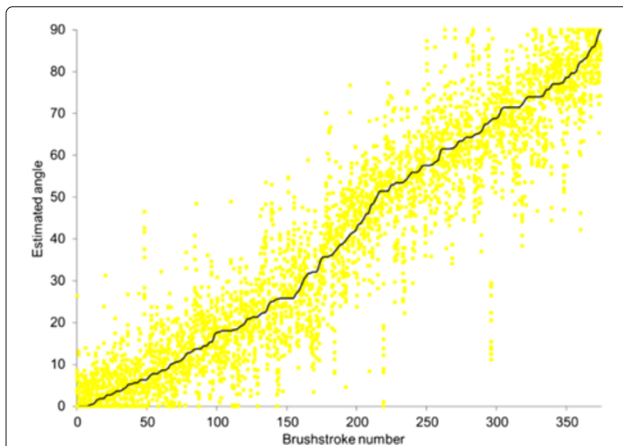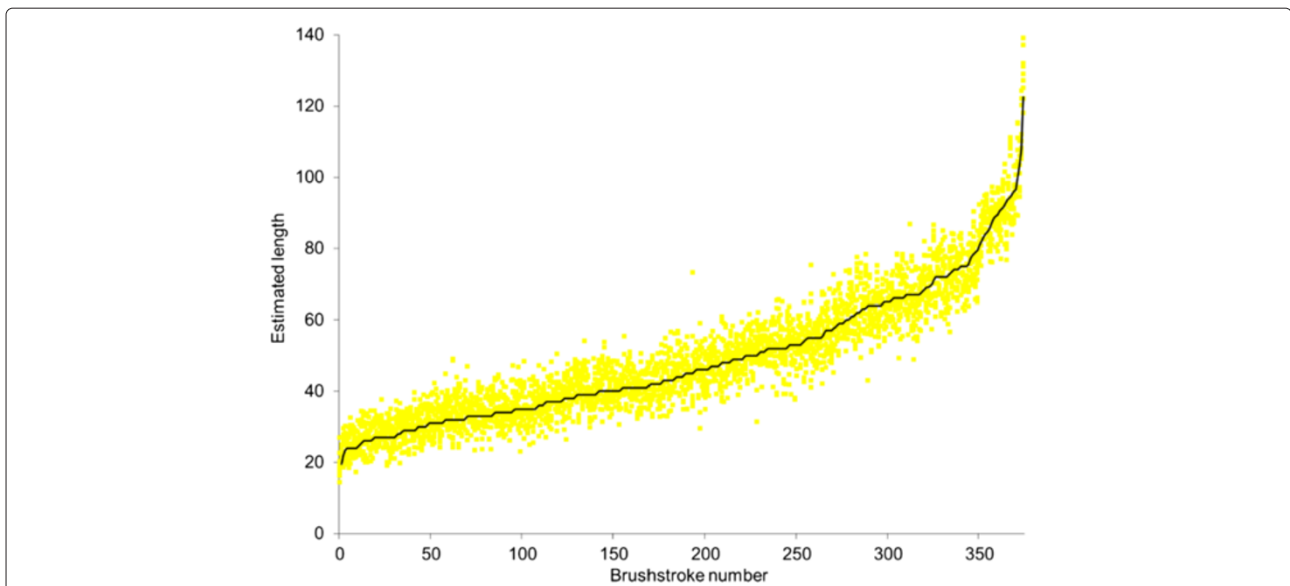
Figure 12 Plot of distances between orientation computed by the algorithm and human judgement for individual brushstrokes. Solid line represents results obtained by the proposed algorithm, whereas dots depict individual human judgements. Values are sorted in order of ascending orientation according to automatically obtained results.

correlation among human subjects is 0.92, whereas the correlation of the proposed algorithm to human subjects is only 0.89. This result may be due to possible irregularities in the shape of the automatically extracted brushstrokes.

A complete overview of all the results obtained in the above evaluation phase is available in Figures 12, 13 and 14. In these figures, the differences between judgements provided by human subjects (dots) with respect to measurements computed by the proposed technique (solid line) are depicted. Depending on the particular figure considered, brushstrokes are sorted in order of ascending orientation, length, or width. These figures complement the results presented in Tables 1, 2 and 3, respectively, by showing the complete distribution of brushstrokes considered in the random selection process. Moreover, besides showing the deviation of the observations, they confirm that the results of the proposed algorithm are quite well aligned with the mean of human evaluations. Lastly, they also show that the manual evaluation process is not always an easy task. In fact, large differences in human evaluations may be observed in the plots (e.g., on the angles identified for brushstroke 295). These differences are reasonably due to the uncertainty in analyzing a brushstroke identified by a randomly selected seed, as the evaluation may become challenging for human subjects if the starting point is located on the boundaries of the brushstroke (uncertainty in identifying which brushstroke should be evaluated), if the brushstroke is occluded by other brushstrokes (uncertainty in prevailing orientation), etc.

algorithm shows a good correlation for length and width values. In fact, for length values, the average correlation among human subjects is 0.97, whereas the correlation of the proposed algorithm to human subjects is 0.99. Similarly, for width values, the average correlation among human subjects is 0.89, whereas the correlation of the proposed algorithm to human subjects is 0.95. In the case of orientation measures, the correlation is not as good as for the other characterization parameters. In fact, the average

Figure 13 Plot of distances between length computed by the algorithm and human judgement for individual brushstrokes. Solid line represents results obtained by the proposed algorithm, whereas dots depict individual human judgements. Values are sorted in order of ascending length according to automatically obtained results.
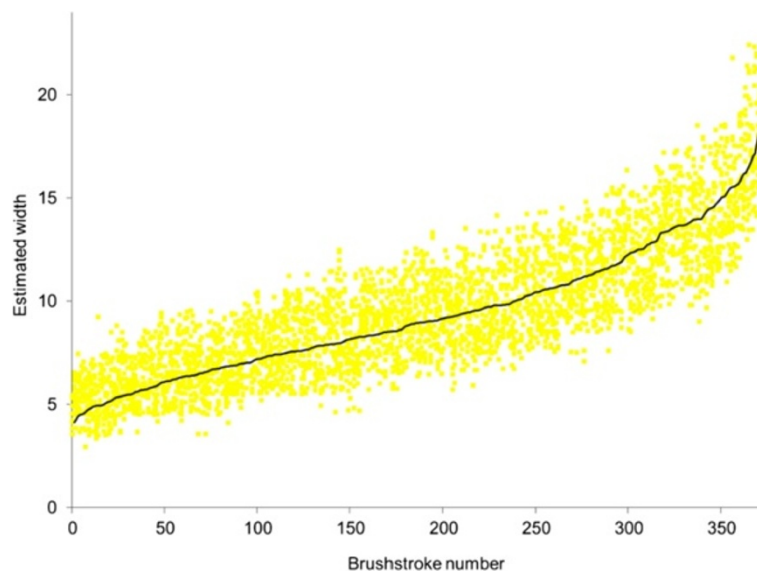
Figure 14 Plot of distances between width computed by the algorithm and human judgement for individual brushstrokes. Solid line represents results obtained by the proposed algorithm, whereas dots depict individual human judgements. Values are sorted in order of ascending width according to automatically obtained results.

In summary, results above suggest that the proposed technique represents a good approximation of the choices made by human subjects, with an average distance on orientation measures that is about 11° and a percentage distance that is less than 9% and 5% on length and width judgements, respectively.

Furthermore, an *a posteriori* visual analysis carried out on the 375 valid brushstrokes extracted in this evaluation phase confirmed the ability of the designed approach to limit the impact of false positives (i.e., brushstrokes validated by the algorithm whose features are clearly different from those obtained by visual examination). On the other hand, concerning irregular shapes discarded by the algorithm, the visual analysis revealed that, in 76 cases out of 125, the cause was that overlapped brushstrokes could not be separated by the algorithm (but they were still detectable by human evaluation). In the remaining cases, they were due to the positioning of the initial seed into areas where the brushstroke was almost covered by other layers (and they were hard to detect even by human subjects).

In order to check the ability of the designed approach to adapt to application scenarios possibly characterized by digital reproductions at different resolutions, specific tests were performed working with image patches of different sizes. The parameters identified during the calibration phase were maintained fixed. Results concerning average MSD_α, $MSPD_\lambda$, and $MSPD_\phi$ obtained by three human subjects on 30 seeds from three different image patches are reported in Table 4. It could be easily observed that the

impact of image resolution on algorithm performances is almost unnoticeable, thus making the proposed method applicable in heterogenous situations.

It is worth observing that, although the metrics adopted in the above evaluation phase are derived from [10], a direct comparison of the results obtained by the proposed technique with achievements reported in the reference paper is not feasible since in [10] the prevailing orientation of a whole patch is extracted, rather than the orientation of each single brushstroke. Moreover, in the above work, length and width characteristics are not considered. Lastly, different datasets are used.

Based on all such considerations, after having analyzed the behavior of the algorithm in approximating how non-expert human subjects perceive some representative features of a brushstroke, it was decided to extend the scope of the evaluation and compare the performances of the proposed technique with both manually and automatically generated results reported in [9]. To this aim, the algorithm presented in this paper was run on the patches from [9] illustrated in Figure 15.

Table 4 Performances of the proposed algorithm in terms of brushstroke orientation, length, and width for different image resolutions

Resolution (dpi)	MSD_α	$MSPD_\lambda$	$MSPD_\phi$
86.1	104.95	0.010	0.14
129.1	117.24	0.012	0.14
172.2	112.58	0.012	0.11
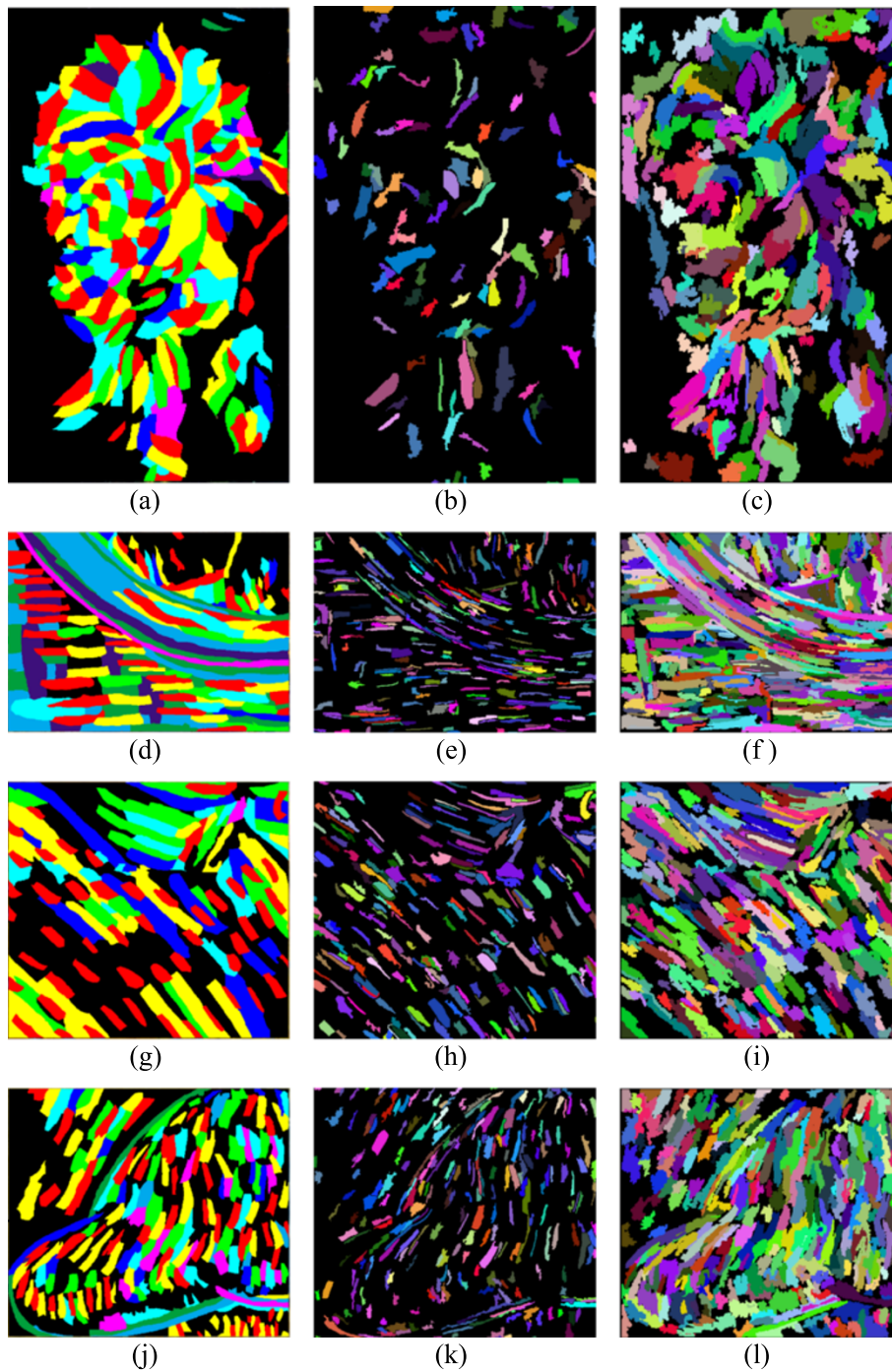
Figure 15 Comparison of brushstrokes marked manually and extracted automatically by the algorithm in [9] and the proposed algorithm. Comparison of brushstrokes marked manually by art experts **(a,d,g,j)** with brushstrokes extracted automatically by the algorithm in [9] **(b,e,h,k)** and by the proposed one **(c,f,i,l)**. Manually marked and automatically extracted brushstrokes have been provided by and used with the permission of the authors of [9]. Painting IDs from top to bottom: F218, F386, F518, F538.

For each patch, the figure shows the brushstrokes identified by art experts by manually marking, i.e., tracing, their irregular boundaries (first column), as well as the results obtained with the algorithm in [9] (second column) and the proposed one (third column). Original patches from van Gogh's works used as input for the brushstroke extraction processes can be found in the reference paper.

It is worth observing that the added value of this further comparison is twofold. On the one hand, it allows the proposed technique to be tested in a context where manual measures have been gathered by different subjects with a different methodology. On the other hand, it allows the performances of the devised algorithm to be assessed on a wider dataset, thus evaluating the ability of the approach presented in this paper to be exploited in more general scenarios.

As pointed out by the authors of [9], their automatic technique might represent a physical brushstroke by approximating it with (i.e., by splitting it on) several separated brushstrokes. By qualitatively (i.e., visually) comparing results in Figure 15, it can be easily observed that brushstrokes extracted by the automatic algorithm in [9] are generally thinner and shorter than those extracted by hand. Hence, this approach might perform well in extracting some brushstroke characteristics considered in the reference work, like orientation or straightness, but it could have difficulties in approximating other features like, for instance, length and width. At the same time, the visual comparison also highlights that the algorithm proposed in this paper extracts a larger number of brushstrokes, which sometimes extends outside the boundaries of the manually extracted ones.

In order to analyze the above behavior from a quantitative point of view, the Valid rate, Detection rate, Fill rate, and Representativeness rate presented in Section 5 were computed on the patches in Figure 15. Results obtained are summarized in Tables 5 and 6. As expected, the algorithm in [9] achieves higher values for the first two metrics, since automatically extracted brushstrokes are generally very small so that they often remain within the boundaries of manual brushstrokes. However, the proposed method achieves better performances for the last two metrics, meaning that it is able to approximate a single manually marked brushstroke with a smaller number of automatically extracted brushstrokes that are larger than those of [9].

Since the above metrics considering brushstrokes coverage are complementary, none of the algorithms considered can be regarded as superior unless a specific application scenario is defined. Hence, it was decided to adopt again the approach in [10] and compare the results of the two automatic approaches against human observations by art experts in terms of brushstrokes orientation, length and width. Results obtained on the patches in Figure 15 are reported in Table 7.

By considering tabulated data, it can be easily observed that, with the exception of one orientation measure (for which the algorithm in [9] achieves a higher accuracy), for the considered metrics and patches, the technique illustrated in this paper achieves results which are closer to observations made by art experts.

8 Conclusion

In this work, a novel technique for automatic extraction of individual brushstrokes from van Gogh's paintings has been presented. In order to evaluate the effectiveness of the devised solution, automatically obtained measures about brushstroke orientation, length, and width were considered against visual observations by human subjects. The comparison showed that results obtained by the proposed technique are quite well aligned with human

Table 5 Comparison between the algorithm in [9] and the proposed algorithm (Valid rate and Detection rate) for Figure 15

Painting ID	Valid rate Li et al.	Valid rate proposed	Detection rate Li et al.	Detection rate proposed
F218	*42.7%*	20.2%	*21.6%*	20.5%
F386	*73.7%*	41.6%	*68.4%*	59.7%
F518	*60.7%*	35.3%	*75.2%*	60.1%
F538	*49.1%*	23.5%	*44.9%*	32.2%

The comparison of results obtained by using the algorithm in [9] and the one proposed in this paper is in terms of Valid rate and Detection rate for the patches in Figure 15. Higher values marked in italics are better.

Table 6 Comparison between the algorithm in [9] and the proposed algorithm (Fill rate and Representativeness rate) for Figure 15

Painting ID	Fill rate Li et al.	Fill rate proposed	Representativeness rate Li et al.	Representativeness rate proposed
F218	29.1%	*55.2%*	2.10%	*22.3%*
F386	15.0%	*24.6%*	0.71%	*13.2%*
F518	24.2%	*40.3%*	8.43%	*29.8%*
F538	12.9%	*31.8%*	3.22%	*23.7%*

The comparison of results obtained by using the algorithm in [9] and the one proposed in this paper is in terms of Fill rate and Representativeness rate for the patches in Figure 15. Higher values marked in italics are better.

Table 7 Comparison between orientation, length, and width results for Figure 15

Painting ID	MSD_α Li et al.	MSD_α proposed	$MSPD_\lambda$ Li et al.	$MSPD_\lambda$ proposed	$MSPD_\phi$ Li et al.	$MSPD_\phi$ proposed
F218	144	*137*	1.29	*1.06*	2.40	*0.86*
F386	112	*94*	3.89	*3.66*	2.75	*0.74*
F518	*47*	68	1.44	*0.79*	2.76	*0.45*
F538	97	*83*	1.12	*0.38*	2.50	*0.43*

Comparison between orientation, length and width results computed over brushstrokes extracted automatically (using the algorithm in [9] and the one proposed in this paper) and manually extracted ones (from the work by Li et al.) for the patches in Figure 15. Lower values marked in italics are better.

judgments. Moreover, a comparison with a recent algorithm for the analysis of paintings by van Gogh and his contemporaries indicated that the proposed technique is capable to cover a wider area of the artist's brushstrokes, thus getting improved performances for some of the evaluation metrics considered. Thus, it can be concluded that the proposed solution could represent an interesting aid in tasks requesting unsupervised brushstroke identification and description.

Given the accuracy that could be achieved, future works will be aimed at analyzing the applicability of the designed algorithm in practical scenarios relying on the automatic analysis of painting features. For instance, it could be interesting to integrate the proposed approach into an artwork authentication or painting/artist classification framework. Such activities would additionally contribute at (indirectly) assessing the performances and effectiveness of the proposed technique.

Lastly, since in this work the devised algorithm relies only on brushstroke shape features, future research activities will be aimed at investigating the integration of complementary attributes like occlusion patterns, color distribution, viscosity, etc. This choice might further improve the characterization process as well as the overall algorithm performances.

Endnote
[a]In the tests reported in the following, T has been initially set to 50.

References
1. M Hours, *Conservation and Scientific Analysis of Painting*. (Van Nostrand Reinhold, New York, 1976)
2. A Kirsh, RS Levenson, *Seeing Through Paintings: Physical Examination in Art Historical Studies*. (Yale University Press, London, 2000)
3. A Callen, *The Art of Impressionism: Painting Technique and the Making of Modernity*. (Yale University Press, London, 2000)
4. CR Johnson, E Hendriks, in *Proc. 1st IP4AI Workshop*. Brushwork in the Paintings of Vincent Van Gogh (Amsterdam The Netherlands, 2007)
5. R Sablatnig, P Kammerer, E Zolda, in *Proc. 14th Int. Conference on Pattern Recognition*. Heirarchical classification of paintings using face and brush stroke models (Brisbane Queensland, 16–20 August 1998), pp. 172–174
6. CR Johnson, E Hendriks, IJ Berezhnoy, E Brevdo, SM Hughes, I Daubechies, J Li, E Postma, JZ Wang, Image processing for artist identification - computerized analysis of Vincent van Gogh's painting brushstrokes. IEEE Signal Process. Mag. **25**(4), 37–48 (2008)
7. H Maitre, F Schmitt, C Lahanier, in *Proc 2001 IEEE Int. Conference on Image Processing*. 15 Years of image processing and the fine arts (Thessaloniki, Greece, 7–10 October 2001), pp. 557–561
8. MM van Dantzigm, *Pictology*. (E.J. Brill, Leiden, 1973)
9. J Li, L Yao, E Hendriks, JZ Wang, Rhythmic brushstrokes distinguish van Gogh from his contemporaries: findings via automated brushstroke extraction. IEEE Trans. Pattern Anal. Mach. Intell. **34**(6), 1159–1176 (2012)
10. IE Berezhnoy, EO Postma, HJ van den Herik, Automatic extraction of brushstroke orientation from paintings. Machine Vision App. **20**(1), 1–9 (2008)
11. DG Stork, in *Proc. 13th International Conference on Computer Analysis of Images and Patterns*. Computer vision and computer graphics analysis of paintings and drawings: an introduction to the literature (Münster, Germany, 2–4 September 2009), pp. 9–24
12. MC Vill, R Sablatnig, in *Proc. of Computer Vision Winter Workshop*. Stroke ending shape features for stroke classification (Moravske Toplice, Slovenija, 4–6 February 2008), pp. 91–98
13. T Melzer, P Kammerer, E Zolda, in *Proc. 14th Int. Conference on Pattern Recognition*. Stroke detection of brushstrokes in portrait miniatures using a semi-parametric and a model based approach (Brisbane, Queensland, 16–20 Aug 1998), pp. 474–476
14. D Keren, in *Proc. 16th International Conference on Pattern Recognition*. Painter identification using local features and naive, Bayes (Quebec City, Canada, 11–15 August 2002), pp. 474–477
15. HJ van den Herik, EO Postma, Discovering the visual signature of painters. Future Directions for Intelligent Systems and Information Sciences. **45**, 129–147 (2000)
16. Y Marchenko, TS Chua, I Aristarkhova, J Ramesh, in *Proc. IEEE International Conference on Multimedia and Expo*. Representation and retrieval of paintings based on art history concepts (Taipei, 30 June 2004), pp. 1023–1026
17. IE Berezhnoy, EO Postma, J van den Herik, in *Proc. 16th Int. Conference of the Association for History and Computing*. Computerized visual analysis of paintings (Amsterdam, September 2005), pp. 28–32
18. H Wendt, P Jaffard, S Abry, *Bruegel's Drawings Under the Multifractal Microscope, 37th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto*, (Japan, 25–30 March 2012), pp. 3909–3912
19. P Abry, H Wendt, S Jaffard, When van Gogh meets Mandelbrot: multifractal classification of painting's texture. Signal Process. **93**(3), 554–572 (2013). ISSN 0165–1684
20. H Qi, A Taeb, SM Hughes, Visual stylometry using background selection and wavelet-HMT-based Fisher information distances for attribution and dating of impressionist paintings. Signal Process. **93**(3), 541–553 (2013). ISSN 0165–1684
21. J Li, JZ Wang, Studying digital imagery of ancient paintings by mixtures of stochastic models. IEEE Trans. Image Process. **13**, 340–353 (2004)
22. S Xu, Y Xu, SB Kang, DH Salesin, Y Pan, H-Y Shum, Animating Chinese paintings through stroke-based decomposition. ACM Trans Graphics. **25**(2), 239–267 (2006)
23. M Shahram, DG Stork, D Donoho, in *Proc. of SPIE, Computer Image Analysis in the Study of Art*. Recovering layers of brush strokes through statistical analysis of color and shape: an application to van, Gogh's Self Portrait with Grey Felt Hat (San Jose, CA, 19 March 2008), pp. 68100D1–8
24. Y Kuang, DG Stork, F Kahl, in *Proc. of SPIE, Computer Vision and Image Analysis of art II*. Improved curvature-based inpainting applied to fine art: recovering van, Gogh's partially hidden brush strokes (San Francisco Airport, California, USA, 23 January 2011), pp. 78690I1-15
25. I Sobol, Uniformly distributed sequences with an additional uniform property. USSR Comput. Mathematics Math. Phys. **16**, 236–242 (1977)
26. AW Fitzgibbon, RB Fisher, in *Proc. 5th British Machine Vision Conference*. A Buyer's guide to conic fitting (Birmingham, September 1995), pp. 513–522
27. MD Abramoff, PJ Magelhaes, SJ Ram, Image processing with ImageJ. Biophotonics Int. **11**(7), 36–42 (2004)