Powder Technology 401 (2022) 117275

Contents lists available at ScienceDirect

Powder Technology

journal homepage: www.elsevier.com/locate/powtec

An image-processing algorithm for morphological characterisation of soot agglomerates from TEM micrographs: Development and functional description

Adrián Cabarcos, Concepción Paz *, Raquel Pérez-Orozco, Jesús Vence

CINTECX, Universidade de Vigo, Campus Universitario Lagoas-Marcosende, Vigo 36310, Spain

ARTICLE INFO

Article history:
Received 4 October 2021
Received in revised form 7 March 2022
Accepted 8 March 2022
Available online 12 March 2022

Keywords:
Soot agglomerates
Transmission electron microscopy
Image processing

ABSTRACT

The inspection of soot agglomerates from microscopy images usually relies on manual human measurements, whereas image processing tools for faster analysis are highly demanded. In this study, an automated algorithm for the extraction of morphological soot descriptors from transmission electron micrographs is presented. The proposed algorithm involves the detection of the image scale (the conversion from pixels to nanometres) using a Hough transform and an optical character recognition process. Primary particles are identified through a two-step circle Hough transform combining phase-coding and edge-based approaches, whereas size descriptors are obtained through spatial and frequency filtering. Finally, the fractal dimension is obtained for each agglomerate as a projected-area derived measurement due to an iterative process. Results were validated by comparison of a set of micrographs taken at three different magnifications with manual image processing, obtaining *p*-values greater than 0.05 and around 91.5% time saving.

1. Introduction

It is well known that soot, as a pollutant produced by the transport industry, has a negative impact on public health and nature [1]. Several studies have attempted to quantify the influence of this type of particles on soil and water pollution [2]. Other studies have tried to link the presence of soot with the impoverishment of air quality [3]. In any case, the enormous risk to health posed by the exposure to this type of emissions is also a fait accompli. Significant relationships with cancer, asthma or bronchitis among other diseases have been found [4].

The United States and the European Union are aware of this problematic situation. Transport and industry legislation is increasingly restrictive in terms of emissions allowed [5], in such a way that a deeper understanding of soot properties and formation would facilitate the development of less environmentally harmful and more efficient thermal systems.

In this sense, the soot formation process can be regarded as an extremely complex phenomenon where a conversion from hydrocarbon fuel molecules containing a few carbon atoms to carbonaceous particles containing a few millions of carbon atoms takes place [6]. In fact, in their review of soot formation from gases, Palmer and Cullis [7] point out that there are at least eight different theories of soot formation. Nowadays, the most accepted one believes that soot formation begins when fuel molecules are pyrolyzed into molecular precursors [8], which are the molecular building blocks of soot [9]. Those precursors chemically bond into singular cyclical aromatic rings which grow to form polycyclic aromatic hydrocarbons (PAHs) [10]. The evolution of PAHs chains is not well understood nowadays but eventually, PAH molecules grow large enough to become soot nuclei during a nucleation step [6]. High temperatures typically involved in flames promote those soot nuclei to growth through coagulation until they form nanometric structures known as primary particles. Furthermore, in-situ and ex-situ observations of those primary particles have shown a solid and spherical shape for these structures [11,12]. The last step of soot formation involves the physical adherence of primary particles into agglomerates. Those agglomerates have complex, irregular and curved structures where primary particles are connected through overlapping and/or carbon surface growth.

Fresh soot particulate generated through this process exhibits different characteristics depending on combustion conditions and sources. Therefore, a wide variety of methods for soot characterisation and classification have been developed over time [13]. Most of these proposed techniques are focused on particulate morphology as this knowledge provides significant information about their properties and potential health impacts. As a result of those works, numerous studies have been published on the effect of operating conditions of combustion

* Corresponding author.
E-mail address: cpaz@uvigo.es (C. Paz).

https://doi.org/10.1016/j.powtec.2022.117275

systems such as diesel engines [14,15] or flame-based burners on soot emissions [16].

Among those types of studies, transmission electron microscopy (TEM) inspection has become one of the most common techniques used to obtain morphological information about soot [17]. Those ex-situ observations allow direct measurements of geometrical descriptors from 2D projections. Otherwise, indirect methods have the disadvantage of relying on related properties such as aerodynamic mobility or light scattering [18]. In addition, the use of TEM images for soot characterisation also has its own drawbacks. Most micrograph inspection methodologies involve manual methods for both agglomerates and primary particles recognition and measurements [19–21]. Furthermore, the analysis of a large volume of micrographs is needed in order to obtain significant results. Therefore, most methodologies reported by literature require an excessive time for image analysis and results obtaining [22]. Traditional manual sizing methods can also yield on poor statistics size distributions: due to the subjective nature of manual readings, those procedures are also susceptible to human subjective errors that involve additional uncertainties to be considered on result discussion [23].

The application of computer vision tools for the inspection of TEM images has become a promising field of study for processing time saving and increasing results reliability. With the advent of both pattern recognition techniques and modern image processing methods, a significant evolution on the analysis of soot TEM images is taking place [24]. In this sense, first approaches to this new research line were mainly focused on low-level vision tools for features enhancing. E.g., a histogram equalization technique has been used to expand pixel intensity distributions and, thus, increase the image contrast between fringes and background [25]. Spatial and frequency filters were used for residual noise reductions [13,26]. Top-Hat transformations were used for TEM acquisition illumination issues correction [27], as well as threshold binarization techniques were used for micrograph segmentation among other existing image processing procedures [28].

However, the most significant leap on the integration of computer vision tools for soot micrograph inspection took place over the last decade. Recent studies were carried out in order to present new methodologies about automatic TEM images processing. Dastanpour et al. [29] proposed an automatic method for the calculation of mean diameters of primary particles based on 2D correlations from the aggregate skeleton. Grishin et al. [30] used a modified Hough transform (HT) for the automated identification and measurement of the size distribution of primary particles in individual aggregates, where their model has the limitation of only being able to detect primary particles that touch the edges of the aggregates. Mirzaei et al. [31] proposed a non-linear filtering also combined with a modified circular Hough transform for primary particle measurement. Recently, Frei et al. [32] proposed the use of a convolutional neural network for a fully automated primary particle size analysis, and Sipkens and Rogak [33] proposed the use of a k-means classifier for segmenting TEM images of soot.

In this work, an automated algorithm is proposed to perform the extraction of morphological soot descriptors from TEM micrographs. Primary particles are detected using a Hough transform where, instead of edge-based and phase-coding formulations [34,35], a two-step circular Hough transform combining both approaches has been used. Agglomerate size descriptors are obtained through a combined use of a Fourier domain filtering and an Otsu's segmentation process. From those results, instead of box counting methods or the use of double-logarithmic plots for the calculation of fractal descriptors [36], the fractal dimension of each agglomerate is calculated as a projected-area derived measurement on an iterative process. Furthermore, as a complement to previous tools for robust processing of microscopy images [36,37], the proposed algorithm also identifies the scale of the image by the combined use of an optical character recognition process and a Hough line transform.

2. Computer vision methods

The image-processing method presented here involves the analysis of TEM micrographs such as the one outlined in Fig. 1(a) where, in addition to the primary particles and agglomerates, there is a scale indicator at the bottom left of the TEM image. Different magnifications can be used in microscopy inspection (see Fig. 1(b) for some examples). Therefore, the first step will involve the recognition of the scale, which quantifies the conversion from pixels to nanometres and will allow expressing results in metric units instead of in number of pixel values. Once this conversion factor is obtained, both primary particles and agglomerates are identified, and their geometrical descriptors are extracted. Finally, the fractal dimension is calculated through the method proposed by Lapuerta et al. [38] and refined by Martos et al. [39] from the geometrical descriptors previously obtained. The mathematical and computer vision procedures required by each of those main steps schematized in Fig. 1(c), are described in detail below.

2.1. Scale recognition

The recognition of the scale implies the combination of two different labours as depicted in Fig. 2: alphanumeric characters, which indicate the unit of measure, must be interpreted; and the straight line indicative of the number of pixels, which was observed to have a maximum size of around 1200 pixels, must be measured. These elements — the straight line and the characters — are located in the lower left corner of the image, at approximately 150 pixels from both the left and the bottom margin, with a vertical distance of 30 pixels between the characters and the straight line. Therefore, those regions of the image are clipped for computational efficiency reasons.

Regarding the cropped image which contains the alphanumeric characters, an optical character recognition (OCR) method has been used to convert those typescripts to code-readable values [40]. Steps required to fulfil this aim are also shown in Fig. 2. To perform the OCR process, the greyscale input image is firstly converted to a black and white one, where a two-class segmentation is performed to separate all characters from background. The Otsu's algorithm — described in detail in Appendix A — was used for this binarization process [41].

Once the optimal threshold, T, has been found using the Otsu's method, the binarization process involves simple comparisons according to Eq. (1). Each pixel of the grey-level image $f(x,y)$ is compared to the optimal threshold obtained through the Otsu's algorithm. Based on this comparison, a new image $g(x,y)$ is created where background pixels are set to null values, whereas pixels that define the alphanumeric characters are set to unitary values.

$$g(x,y) = \begin{cases} 1, & f(x,y) > T \\ 0, & f(x,y) < T \end{cases} \tag{1}$$

Then, the algorithm searches for all connected components of the resulting image, where each of those elements will correspond to one of the characters present in the image. Each connected component, which is labelled and extracted by cropping its square contour, is then normalized to a size of 42×42 since this is the dimension of the available patterns. Finally, a matrix matching comparison procedure has been used to produce the list of characters. This pixel-by-pixel pattern comparison relies on statistical correlations between the input characters, which are correctly isolated from the rest of the image, and a set of pattern characters stored as a template. Then, the typescript characters are assumed equal to the patterns for which correlation is maximum. Code-readable values are stored on two different variables: one for the sequence of numerical digits and another one for letters, settling the nanometre (string “nm” detected by the OCR) as the unit of reference.

Once those characters have been identified, the next step involves the measure of the straight line that is part of the scale indicator. To

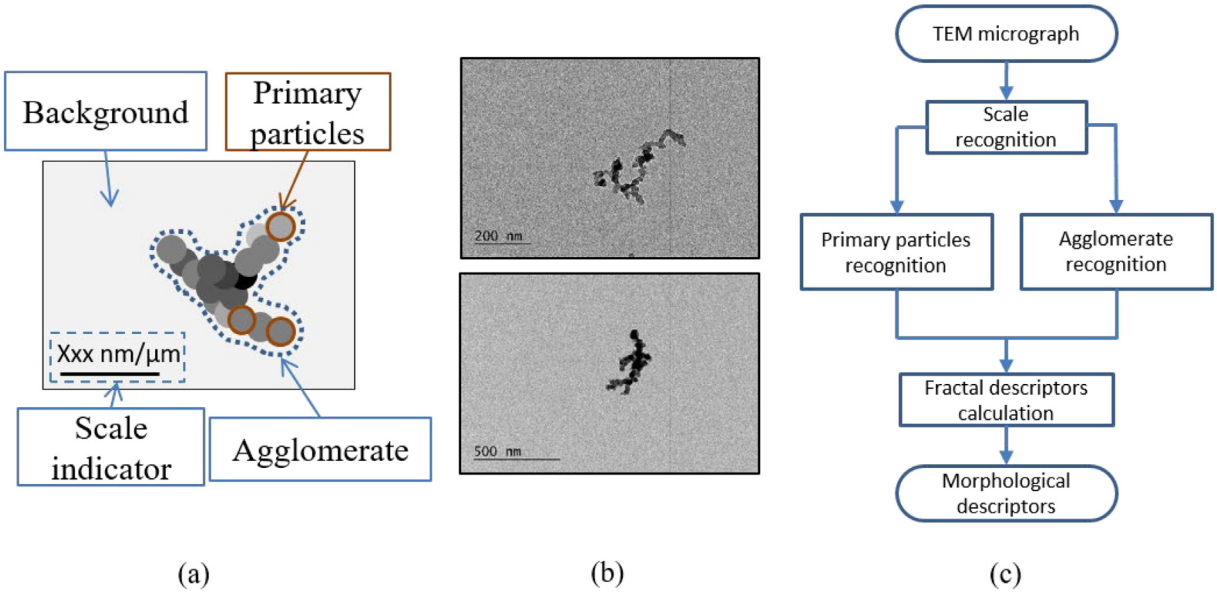

Fig. 1. (a) Schematic of a TEM image of a soot agglomerate. (b) Examples of real TEM images at different magnifications. (c) Flow chart of the proposed algorithm for the extraction of morphological descriptors of soot agglomerates from TEM micrographs.

accomplish this aim, a Sobel filter is firstly used for edge detection [42]. This method uses intensity values in the 3 × 3 region around each pixel to approximate the image gradient. Thus, regions of high spatial frequency can be identified and therefore edges can be found. To perform a successful recognition, a threshold is needed to distinguish which gradient values can be assigned as edges. A Sobel threshold equal to 0.09, which has also been successfully reported for other recognition applications [43], was applied as it was observed to provide a good balance between the preservation of the edges on the straight line and a significant filtering of edges at the noisy background.

Once the edge detection process has been performed, a traditional Hough transform algorithm has been used to recognize straight lines [44]. The implementation of this transformation is presented in detail in Appendix B and, once the scale distance in pixels has been obtained, the conversion factor from pixels to nanometres is finally obtained as the ratio between this scale distance and the typescripts previously converted to code-readable values.

2.2. Recognition of primary particles

Fig. 3 summarizes the step sequence required for the recognition of the circular projections corresponding to the primary particles. As it can be seen, a smoothing filter is firstly used for impulse noise suppression. In this sense, a median filter based on a 3 × 3 window operation was applied to the image. According to this non-linear method, each pixel is substituted by the median intensity of its neighbour. Thus, salt-and-

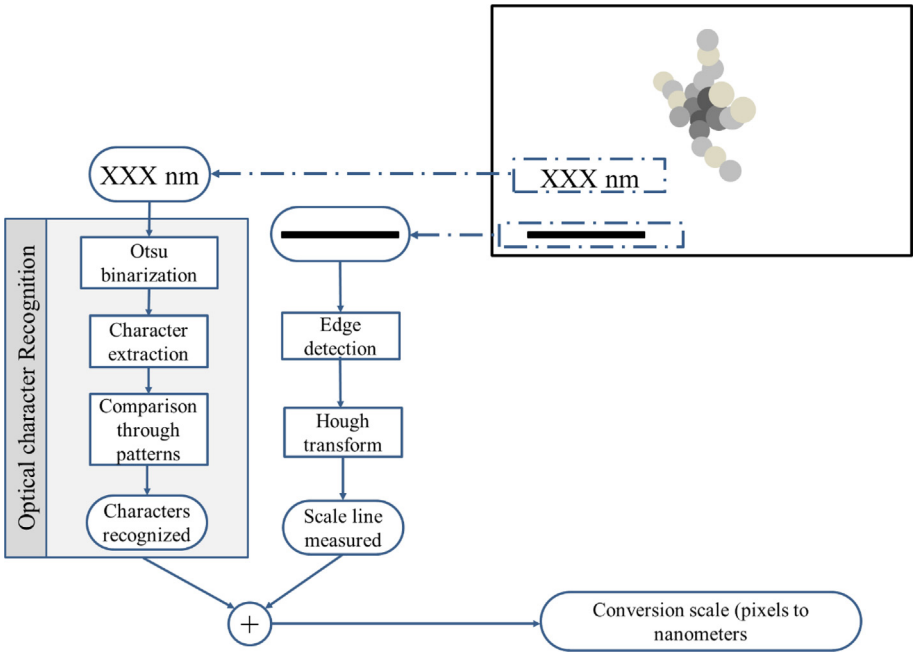

Fig. 2. Schematic of the proposed methodology for TEM scale recognition by the combined use of Hough transform and an OCR.

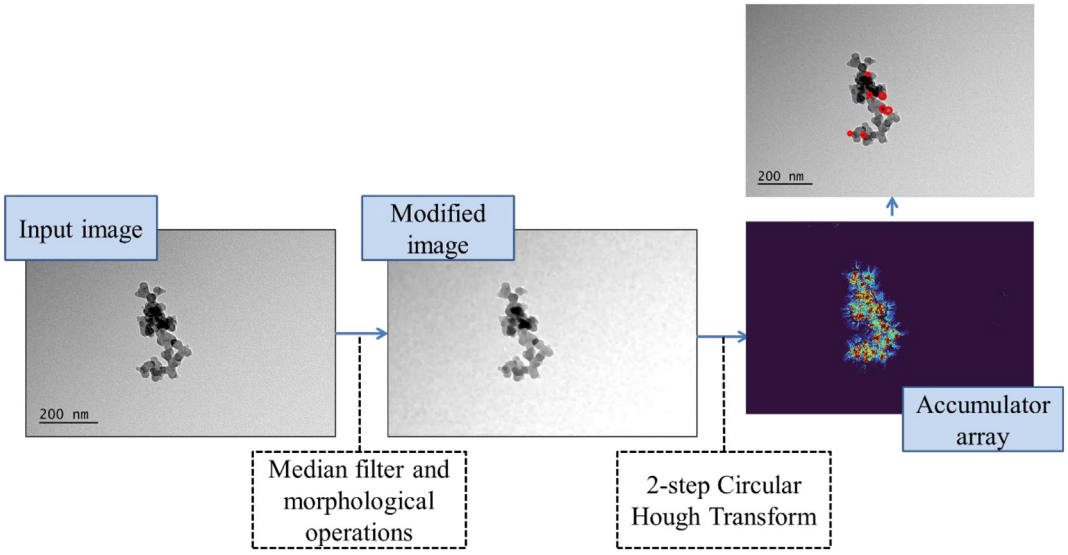

Fig. 3. Schematic of steps required for primary particle recognition from TEM images.

pepper noise can be suppressed from the greyscale image whereas soot shapes are smoothed.

After that filtering step, morphological operations are performed to simplify the agglomerates structure whereas main shapes are preserved. Those morphological processing transformations involve a set of non-linear operations related to the shape of the features present in an image, in such a way that selected morphologies can be enhanced by reordering pixel values [45].

In this sense, a closing operation was performed to the full micrograph. This transformation, which allows a selectively fill of small holes inside the foreground objects, is presented in detail in Appendix C. A disk-shaped structuring kernel was used to perform this transformation, as circular curves are the shapes to be detected by this algorithm. This kernel sets the neighbourhood on which the operations are performed and 5 nm were specifically considered for this work as it corresponds with the minimum size reported by Dastanpour et al. [46] for primary particles. It should also be pointed out that the closed operation fulfils area-preserving properties; it will enhance shapes like those defined by the structural element without introducing significant deviations in results.

Once the morphological transformation step has been done, a Sobel edge detection process as the one described in Section 2.1 is performed. Then, the Hough transform can be also used for finding circular objects. It should be noted that, unlike other alternatives such as the pair-correlation method, which directly provides a mean result, the circular HT allows finding specific circles in the agglomerate [29]. In this sense, a circle with radius R and centre (h,k) can be described parametrically according to Eq. (2).

$$(x-h)^2 + (y-k)^2 = R^2 \tag{2}$$

Therefore, as the circle has three parameters, the Hough space for circular shapes will belong to $\mathcal{R}^3$. However, a two-step algorithm for finding circles has been used instead of the conventional formulation for computational load reduction reasons [30]. The proposed method involves a combined use of a coherent Hough transform and an edge-based approach where, in the first step, a phase coding formulation has been used to compute an accumulator array of complex numbers and, thus, estimate the centres of the particles [35]. In the second step, a histogram of distances of all feature points from those centres is used to extract their radii [47].

To perform the first process, the accumulator array is obtained using a normalized annulus filter $O_{PCA}(m, n)$, given by Eq. (3), which encodes scale information in its complex phase for a discrete range of radii, r, between R_{min} and R_{max} [35].

$$O_{PCA}(m,n) = \begin{cases} \dfrac{e^{2\pi\left(\log\sqrt{m^2+n^2} - \log R_{min}\right)/(\log R_{max} - \log R_{min})i}}{2\pi r} & if\ R_{min}^2 < m^2 + n^2 < R_{max}^2 \\ 0, \text{otherwise} \end{cases} \tag{3}$$

According to this formulation, edge points are projected along a line in the direction of the edge orientation. Those projections have a phase proportional to the distance travelled and, therefore, edges lying on a same circle will be in phase and will also have projections that intersect at a common point [48]. This approach is significantly different from the edge-based method followed by Kook et al. [22], where a 2D real accumulator array is directly calculated from gradient directions of edge points, without any consideration about phase information. However, in the phase-coding approach, the contribution from each edge pixel of a circle will interfere constructively for the phase corresponding to that specific circle and destructively for other phases.

wIt should be noted that phase information was also used by Anderson et al. [49], but unlike in that previous implementation, radii in this work are estimated using a histogram of distances instead of phase decoding procedures [50]. In this sense, pixels of higher absolute values in the accumulator array provide significant indications about the presence of circles, where a sharp local maximum on a histogram bin defines the radius of the circle. Furthermore, for a better approach towards the extraction of this information, the filtering technique proposed by Ioannou et al. [47] has been used to provide an unbiased estimator of the normalized, unclouded part of the circumference of the circle, selecting the resulting maximum bin as the most likely radius.

2.3. Recognition of projected areas

In order to recognize the projected areas of the agglomerates, the original unmodified micrographs are submitted to the low-level transformation sequence shown in Fig. 4. The original image is firstly filtered on spatial domain, and subsequently, on the Fourier domain. Thus, the projected area of the agglomerate is enhanced which simplifies the

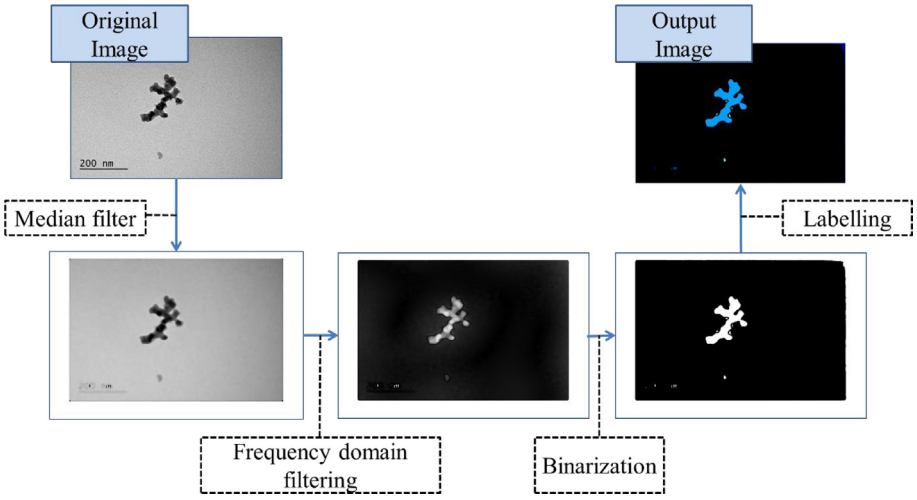

Fig. 4. Low-level vision steps required for projected area extraction.

binarization process performed by the Otsu's method. Connected components are then identified and diameters of gyration and projected areas are extracted.

In this sense, a median filter is used for a firstly image smoothing and denoising step. Fundamentals of this filtering step are identical to those explained in Section 2.2 for primary particles recognition. However, the features of the primary particles are not of interest at this point. This sub-algorithm focuses on the extraction of the position and size of the agglomerates. Therefore, a filter of greater effect is used by means of a higher pixel neighbourhood. For the purpose of this section, a 30×30 neighbourhood is considered, which will tend to smooth both the background and the interior of the agglomerate.

Then, the image is converted to its frequency domain through the Fast Fourier Transform (FFT). Due to this function transformation, which has been commonly applied for lattice fringe analysis [51], an image can be represented according to its frequency spectrum in such a way that homogeneous images have low frequency components whereas abrupt changes involve high frequency ones [52]. The frequency domain filtering process involves the step sequence shown in Fig. 5, where $H(u,v)$ represents the mask employed for the

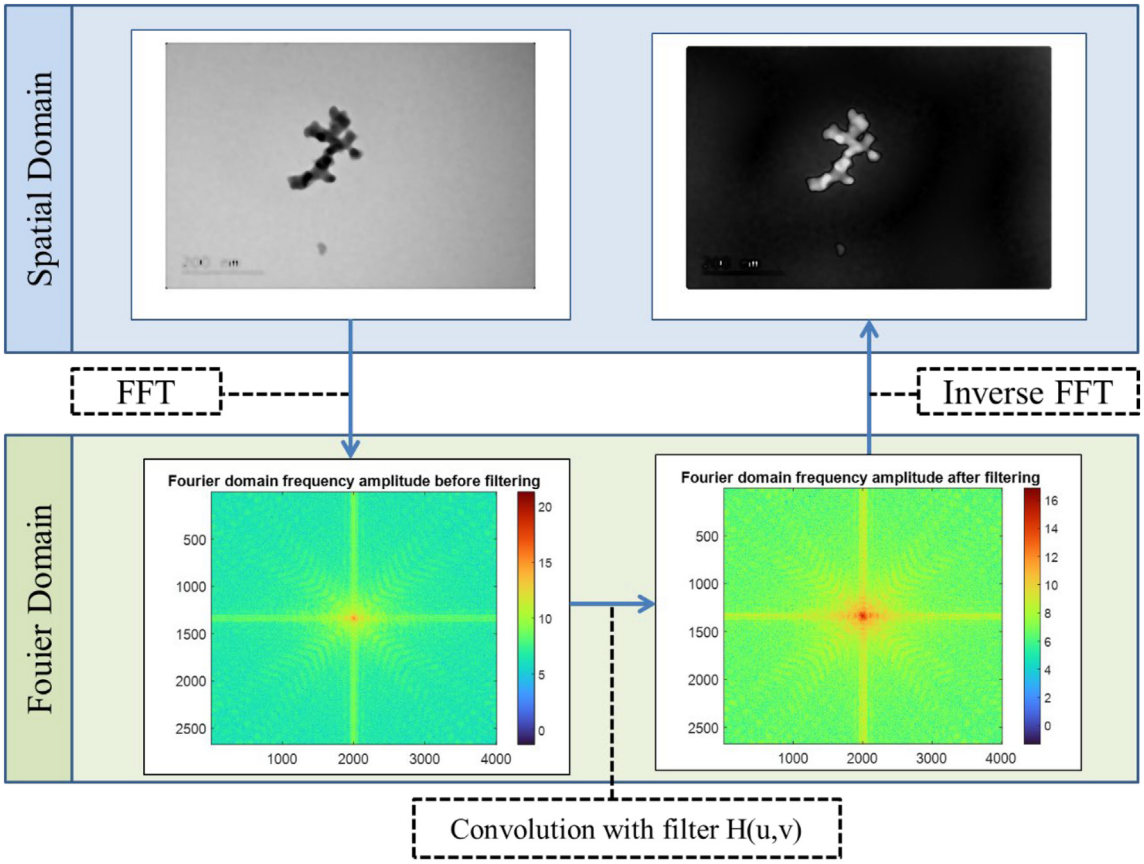

Fig. 5. Transformation from spatial domain to Fourier domain and filtering.

convolution operation. Hence, a high-pass filter is used for low frequency suppression. Thus, high frequency features of the agglomerate will be enhanced and, finally, the filtered image can be reconstructed from its Fourier coefficients, also enhancing the final contrast.

It should be noted that this combination of spatial and frequency filters is designed to exploit differences between the agglomerates and the background. This step is similar to previous image pre-processing algorithms, where bilateral and Gaussian filters were reported for denoising and smoothing purposes, whereas morphological operations were applied to remove background effects [29,33]. However, in this case, pixels in regions of large gradients are brightened due to the high-pass filtering in the frequency domain instead of the use of other alternatives such as bottom-hat and entropy filters.

Finally, once the spatial image has been reconstructed, the resulting image is binarized and connected components are labelled. Components composed of fewer pixels than the corresponding value for three primary particles are ignored. This limit depends on the magnification identified and it is based on the methodology followed by Zhang et al. [53], which suggest that agglomerates of lesser than 3 primary particles should be excluded from fractal analysis. Agglomerates in contact with some image margins are also ignored and, for remaining components, two direct descriptors of the size of the particulate are extracted [54]. Firstly, projected area, A_p, is obtained as a sum of pixels and secondly, the diameter of gyration, D_g, which can be interpreted as the diameter of a circle with the same moment of inertia with respect to their centre of gravity as the agglomerate, can be calculated according to Eq. (4).

$$D_g = 2\sqrt{\frac{1}{N_{px}}\sum_{k=1}^{N_{px}}\left|\vec{r_k}-\vec{r_c}\right|^2} \tag{4}$$

Where $\vec{r_k}$ and $\vec{r_c}$ are position vectors representing the centroid of the agglomerate and the centre of the kth pixel respectively, and N_{px} is the number of the pixels of the agglomerate. Finally, as all those results are obtained in pixel units, a conversion from pixels to nanometres is applied according to the factor scale found through the OCR procedure.

2.4. Fractal dimension calculation

Fractal dimension can be calculated according to the method proposed by Lapuerta et al. [38,55] for soot particulate. This parameter, which ranges between $1 < D_f < 3$, allows to quantify the irregularity of soot structures: The extreme value of $D_f = 1$ involves an agglomerate composed of perfectly aligned primary particles, whereas the value of $D_f = 3$ is reached when primary particles form a compact sphere. Furthermore, the method proposed by Lapuerta et al. [38,55], which has already been used for experimental studies [54], assumes that the number of primary particles, N_{pp} scales with the diameter of gyration according to Eq. (5) and it is also related to the projected area with a correction for screening or overlap as the aggregate size increases, according to Eq. (6).

$$N_{pp} = k_f\left(\frac{D_g^*}{d_{pp}}\right)^{D_f} \tag{5}$$

$$N_{pp} = \left(\frac{A_p}{a_{pp}}\right)^{z} \tag{6}$$

Where a_{pp} denotes a cross-sectional area representative of primary particles, which has been assumed for each agglomerate as equal to the mean area of its composing primary particles, and z denotes an overlapping exponent, which accounts for the number of particles hidden behind the observed ones.

It should be noted the use of a corrected diameter of gyration, D_g^*, in Eq. (5) and, thus, consider that diameters of gyration of 2D projections tend to underestimate the values of 3D structures involving an overestimation of fractal dimensions. This correction, proposed by Martos et al. [39], is given by Eq. (7)

$$D_g^* = 2\left[\frac{D_g}{2} + 3.7336\cdot10^{-5}\cdot\left(\frac{D_g}{2}\right)^{2.7935}\right] \tag{7}$$

Furthermore, Martos et al. [39] also provide a function, given by Eq. (8), to quantify the dependency between the overlapping exponent, z, and the irregularity of the agglomerate expressed in terms of D_f.

$$z = \frac{\ln N_{pp}}{\ln\left(0.8488N_{pp} + 0.1512\right)} - 1 + \left(2.5 - \frac{\ln N_{pp}}{\ln\left(0.8488N_{pp} + 0.1512\right)}\right)^{((D_f-1)/2)^{1.86}} \tag{8}$$

Finally, using Eq. (5), Eq. (6), Eq. (8) and Eq. (9), an initial guess of a fractal prefactor, k_f, can be used to refine the fractal dimension, D_f, until a convergence result which fulfils the boundary conditions described by Lapuerta et al. [38,55] is reached. Those conditions, based on theoretical arrangements of primary particles leading to extreme values of $1 < D_f < 3$, are given by Eqs. (10) and (11).

$$k_f = k_f\,(D_f = 1) - 1 + (1 + k_f(D_f = 3) - k_f(D_f = 1))^{((D_f-1)/2)^{1.86}} \tag{9}$$

$$k_f(D_f = 1) = 1/\left(\sqrt{\frac{1}{3} - \frac{1}{N_{pp}} + \frac{8}{5N_{pp}^2}}\right) \tag{10}$$

$$k_f(D_f = 3) = \frac{\pi}{\sqrt{18}}\left(\frac{5}{3}\right)^{3/2} = 1.5933 \tag{11}$$

3. Experimental set-up

An experimental test bench has been used to obtain the soot samples to be observed by the TEM microscope, and thus obtain the micrographs needed for image processing validation. Main specifications of the experimental layout have been described in detail in previous research [56,57]. However, Fig. 6 shows a schematic about the test facilities.

A combustion aerosol standard (Mini-Cast 5203 Type C) burner was used to produce the particulate matter [58], and soot agglomerates examined in this study have been generated by the combustion of 0.2 normal litres per minute (Nl/min) of 99% purity propane and 4.0 Nl/min of oxidation air. Those Mini-CAST input conditions were holding constant among the full test and were selected to produce emissions with high elemental carbon and thus obtain a graphitic refractory structure for primary particles. Furthermore, in order to obtain a suitable gradient temperature to carry out a thermophoretic sampling process, the gas stream temperature at the outlet of the heating zone was fixed at 200 °C during all test duration.

A thermophoretic sampling system, which consists of a pneumatic actuator with a TEM holder at its end where mesh grids can be housed, has been used to collect the particulate emitted by the burner [59]. The device used has been designed ad-hoc and it can be attached to the experimental setup through a hole drilled in the exhaust pipe. The coupling mechanism consists of a notch and a guide that guarantees a specific and correct positioning of the device since its orientation is essential to ensure a successful sampling [60]. Main surface of the TEM grid must always be parallel to the gas flow to avoid the risk of impact deposition instead of by a thermophoresis mechanism.

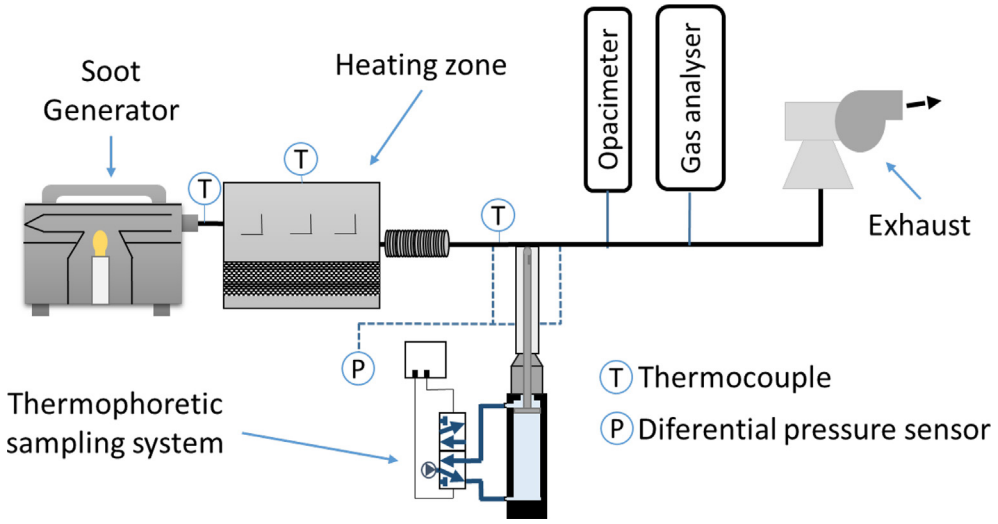

Fig. 6. Schematic of the experimental test bench used for soot particulate generation.

In order to collect the particulate matter from the exhaust gas flow, the TEM holder of the thermophoretic sampling system remained in the gas stream for 600 ms, as recommended by Lapuerta et al. [61] for similar temperature gradients. This sampling time is necessary to ensure a low grid coverage (<10%) and, therefore, to avoid the presence of overlapping agglomerates on the grid. High-magnification pictures of those TEM-grids were taken using a JEOL JEM 1010 transmission electron microscope in bright field mode, equipped with a charge-coupled device (CCD) camera. Using this TEM mode, images are formed due to the weakening of a direct beam of electrons (the illuminant source) by its interaction with the sample. Furthermore, the beam of electrons was energized at an acceleration voltage of 100 kV, with current densities between 150 and 180 pA/cm^2, allowing a maximum resolution of 0.35 nm. Spot size and tilt angle were kept constant during the microscopy stage, with values of 4 nm and 0° respectively.

4. Results and validation

Employing the TEM microscope, a set of 40 agglomerates collected from the experimental setup due to the thermophoretic sampling process was selected. High-resolution images were taken for each of those 40 agglomerates at three different magnifications — 25,000, 20,000, 15,000 x —. The diameter of the primary particles, the size of the particulate matter and its fractal dimension were evaluated following the presented procedure, which was implemented by Matlab software. The code, which involves the use of the Image Processing Toolbox and a modified version of Matlab's imfindcircles function, was executed on an AMD Ryzen 53550H (2.10 GHz) computer processor and 8GB of RAM. In order to validate the accuracy of the proposed algorithm, the same micrographs were also analysed through a manual image processing methodology, where a researcher used imageJ software (Version 1.8.0.172) for scale recognition and soot measurements. In this sense, primary particles identified by the researcher were circumscribed manually to extract their diameter, whereas a first approximation to the contour of the agglomerates was cropped by hand and then manually selecting the appropriate filter and threshold for each case to refine the result.

4.1. Recognition of the scale

In the first place, Fig. 7 shows the conversion between pixels and nanometres obtained by the combined use of the OCR and the Hough transform on the images treated with the proposed algorithm. Theoretical values should be equal to 0.261, 0.322, 0.427 nm/pixel for magnifications of 25,000×, 20,000× and 15,000× respectively. Results found showed a high accuracy in the conversion factor to be applied by the algorithm for the rest of values, where mean relative errors obtained were equal to 2% (25,000×), 1.70% (20,000×) and 2.46% (15,000×). Therefore, this first step of the proposed algorithm does not involve significant deviations in the subsequent conversion of the results. The maximum

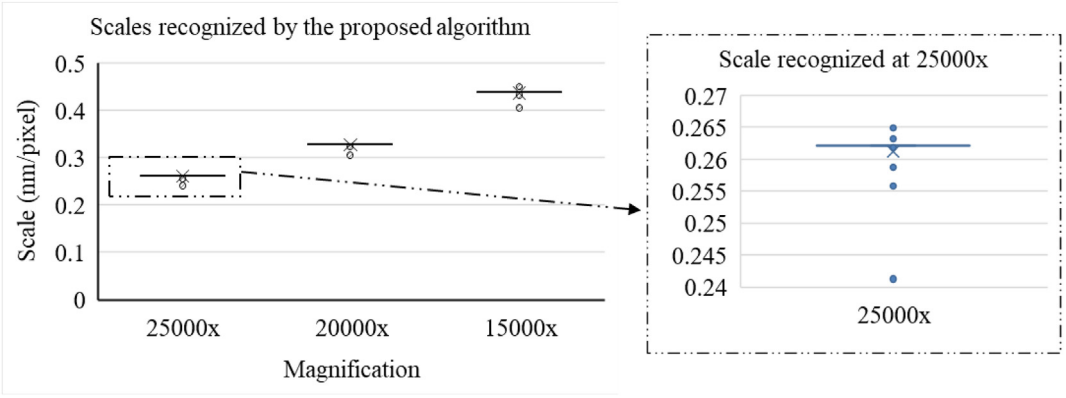

Fig. 7. Distributions of scale values recognized by the proposed algorithm. Due to the small deviations, the boxplot found for the scale recognized at 25000× is also shown in detail.

single error detected among the 120 micrographs analysed has been equal to 9.64%. This error was obtained for an image taken at a magnification of 25,000×, and it corresponds with the outline of the boxplot shown in detail in Fig. 7. Both this maximum error as well as those other errors due to the rest of outliers detected, are associated with heterogeneous illuminations in the process of image taking. Uneven backgrounds involve the detection (or even the absence) of pixels at the edges, which causes those errors in the Hough transform calculation. Note that the high accuracy in the scale recognition depends on both the identification of the correct line using the Hough transform, and a successful comparison of the image characters with the patterns available. Regarding those comparisons performed by the OCR, characters assigned show statistical correlations between 71% and 87% with patterns selected, while correlations between 46% and 62% are found for first discarded patterns. Those results are in the range of other OCR methods such as the Tesseract engine [40], which involves a maximum confidence around 92% regarding the proposed dataset, at the cost of a higher computational load as it performs operations such as a baseline fitting which are not required in this application.

4.2. Recognition of the primary particles

Regarding the analysis of the primary particles that was conducted for the 120 micrographs, Fig. 8 (a) summarizes the distribution of the complete set of diameters found for each of the three magnifications considered for this work. Results obtained through the proposed algorithm are colored as blue bars whereas results obtained by manual measurements are shown as the red ones. It was found that the proposed Hough transform tends to underestimate the diameter of the particles in comparison to manual measurements. This observation was also detected by previous approaches of the circular Hough transform [23,49], and for this dataset, the absolute error regarding the mean diameter of primary particles composing each agglomerate tends to increase to as high as a maximum around 10 nm, as the size of the agglomerates increases. In any case, the interquartile ranges of boxplots shown in Fig. 8 (a) are consistent with previous experimental results obtained in CAST-based facilities (13–33 nm), whereas outliers are also in the range of maximum sizes found for other soot sources such as nonpremixed flames (10–50 nm), or those generated by diesel engines (10–45 nm) [14,62,63].

It should be noted that several factors determine the performance of the HT, most prominently, the lower-bound diameter and the sensitivity to possible centres [49]. On the one hand, the number of primary particles detected mainly depends on the threshold applied to the accumulator matrix. As it can be seen from Fig. 8 (b), the HT exhibits a greater affinity for lower diameters (in pixels). Therefore, the algorithm tends to detect a greater number of particles at lower magnifications. Furthermore, as the threshold increases, the number of primary particles found decreases. For the range of values considered in this work, the differences between mean diameters obtained in each micrograph varies between 0.9 nm and −3 nm according to the threshold selected. On the other hand, size distributions also depend on the minimum

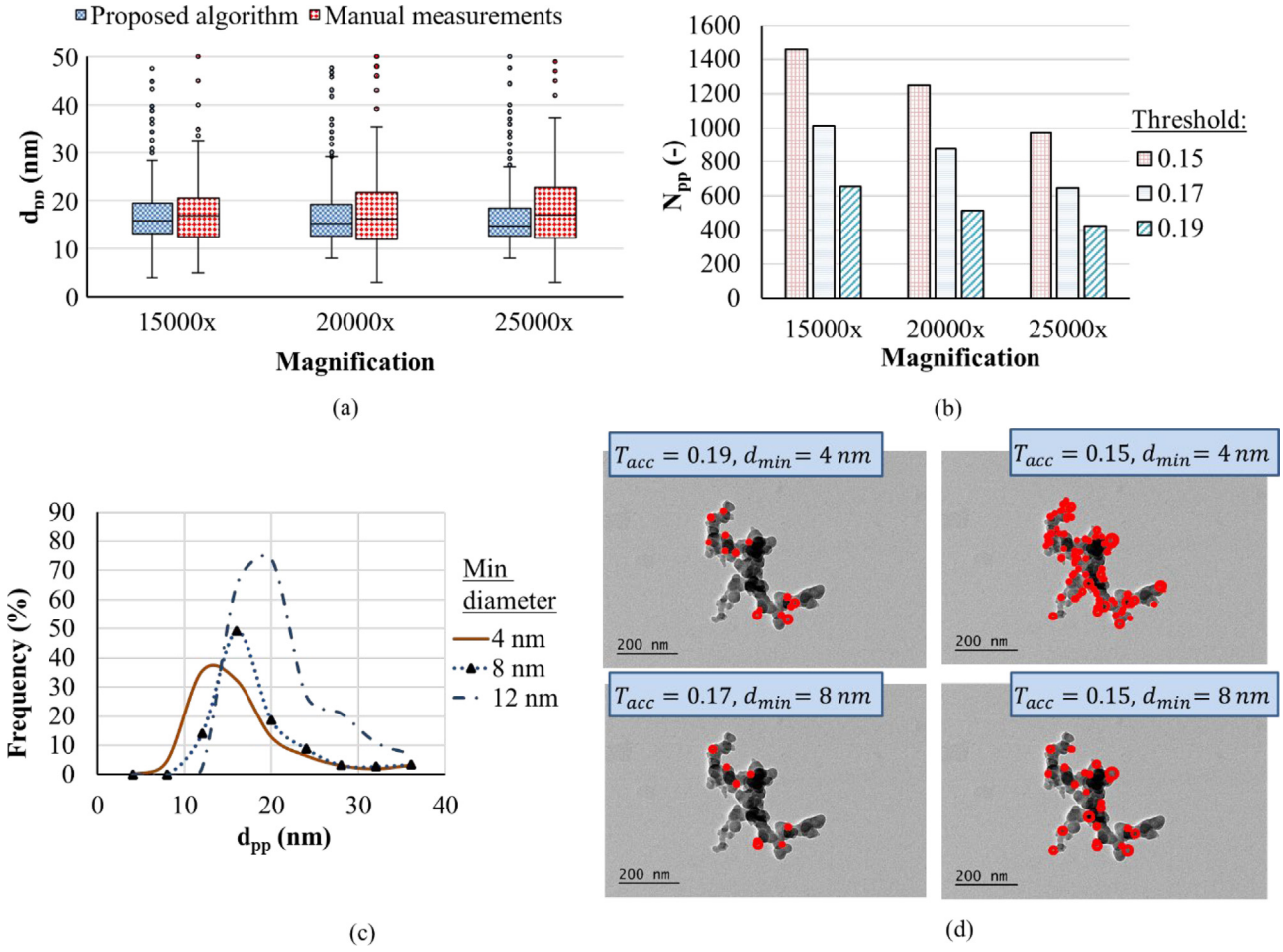

Fig. 8. (a) Comparison of diameters of primary particles detected using the proposed algorithm versus manual measurements, selecting a thresholding $T_{acc} = 0.19$ to compute possible centres. (b) Number of particles found as a function of the thresholding, T_{acc}. (c) Influence of the minimum diameter to be detected for size distribution at 25000×, considering $T_{acc} = 0.19$. (d) Examples of the combined effect of the thresholding, T_{acc} and the lower-bound diameter d_{min} on the same micrograph.

diameter to be detected. Fig. 8 (c) shows this effect for the 25,000× magnification dataset and, as it can be seen, the mean diameter tends to decrease around 5 nm as the minimum allowable diameter also decreases. Kook et al. [22] also reported the same dependency in a previous implementation of a HT approach. Furthermore, in order to illustrate these effects, micrographs shown in Fig. 8 (d) were selected to represent the effect of different thresholds and minimum diameters regarding possible centres and radii respectively. In Fig. 9 (a), a comparison of primary particles detected at the three magnifications considered is also shown.

The dependency on the filtering process carried out in the second stage of the Hough transform should also be analysed. Fig. 9 (b) relates diameters obtained considering the filter proposed by Ioannou et al. [47] with those obtained directly from the radii histogram at 25000×: For several possible radii, the filter tends to promote smaller diameters since larger values may correspond to circles that come out of the projected area of the agglomerate as exemplified in Fig. 9 (c). Furthermore, other Hough transform approaches usually involve similar steps (e.g., normalizations based on the inverse of the circumference), which could also be related with the reported tendency to reduce diameters in soot imaging of previous works [22].

Despite the differences observed, a statistical analysis of those results was conducted to quantify the similarity between distributions shown in Fig. 8 (a). Mann-Whitney tests were performed for each pair of those groups, according to the formulation described in Appendix D [64]. Table 1 summarizes the *p*-values statistically obtained through those tests. As all p-values obtained are greater than a significance level equal to 0.05 [65], it can be firstly confirmed with a high level of confidence (>95%), that the use of different magnifications on the image capturing process of the selected agglomerates does not alter the size distribution of the primary particles recognized by the proposed algorithm. Secondly, it is also confirmed that distributions obtained by manual measurements are equivalent to those distributions obtained using the algorithm presented here for this TEM dataset (*p*-values greater than 0.05).

4.3. Size and fractal descriptors

Fig. 10 (a) and (b) summarizes both projected areas and diameters of gyration found for the agglomerates captured. Analogously to the analysis of primary particles, a similar trend was observed between size distributions obtained at the three magnifications considered for the validation of the algorithm. However, it was found that the algorithm tends to reduce the size of the agglomerates regarding results from manual measurements. This difference is attributed to the suppression during the Otsu's binarization of those less intense regions resulting from the frequency filtering. This observation is consistent with previous uses of the Otsu's method, where Sipkens and Rogak [33] reported a mean reduction of area equivalent diameters between −17% and −23% for different datasets. Furthermore, this effect can even involve the split of an agglomerate in several regions as it can be clearly seen for points out of the trend in Fig. 10 (a). Sipkens et al. [37] have also reported the same constraint in an in-depth comparison of image processing tools of soot. For this specific implementation, this effect

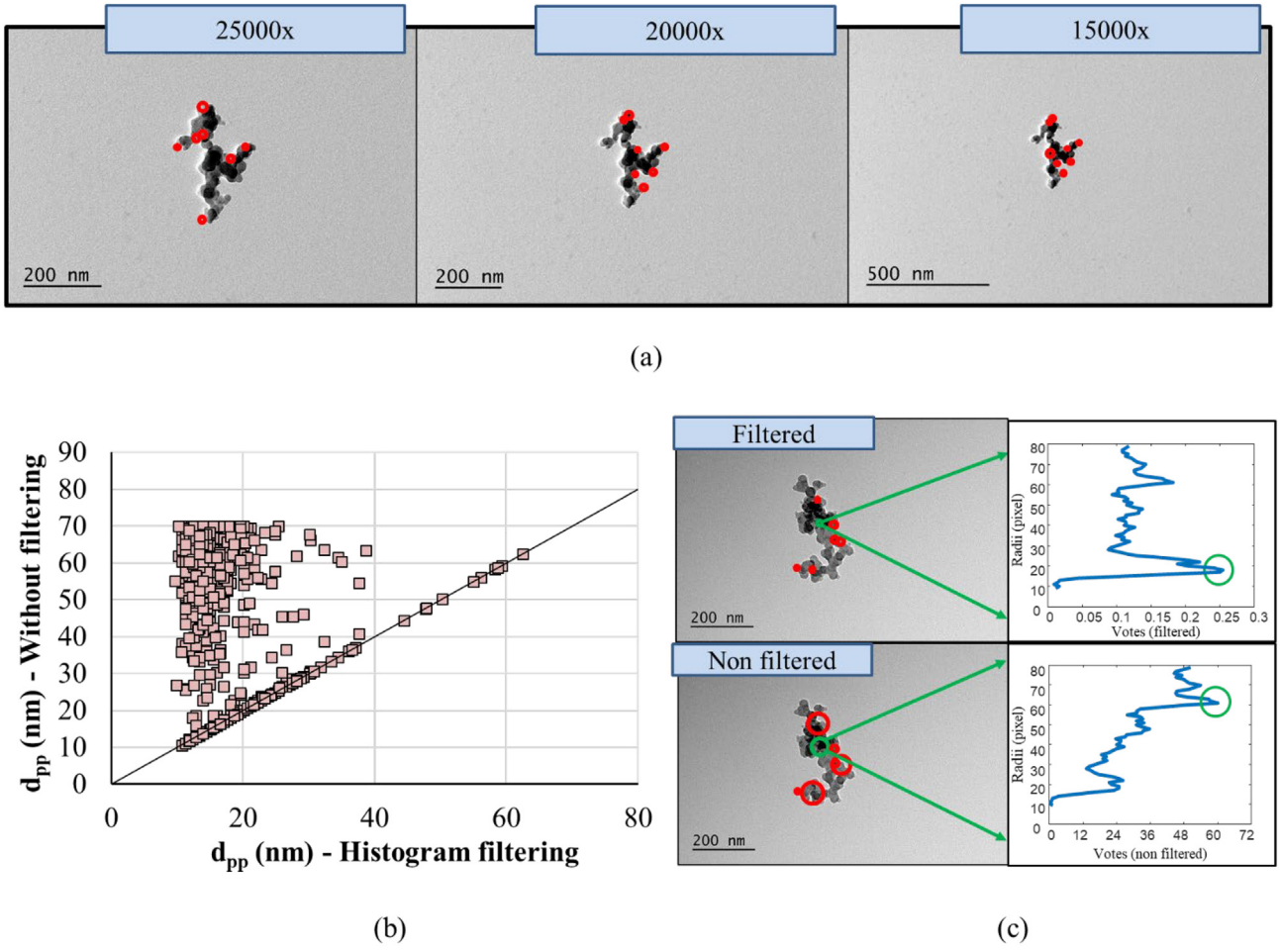

Fig. 9. (a) Examples of micrographs taken at three different magnifications. (b) Influence of the histogram filtering process on the size of the primary particles. (c) Examples of the effect of the histogram filtering on primary particle size.

Table 1
P-values obtained from Mann-Whitney tests for distributions of primary particles.

		Magnification of 25,000×		Magnification of 20,000×		Magnification of 15,000×	
		Proposed algorithm	Manual inspection	Proposed algorithm	Manual inspection	Proposed algorithm	Manual inspection
Proposed algorithm	25,000×	1	0.568	0.139	0.666	0.104	0.714
	20,000×	0.139	0.381	1	0.288	0.728	0.284
	15,000×	0.104	0.288	0.728	0.223	1	0.282

depends on a proper cut-off frequency selection and it could be aggravated for sets of images taken at a wider range of current densities or under uneven illuminations. In this sense, for this specific dataset, a cut-off frequency equal to fr = 0.25 tends to reduce the mean relative error according to Fig. 10 (c), whereas higher values involve greater errors due to a greater number of artefacts and a greater risk of area splitting.

In any case, a statistical analysis was performed to confirm that all diameter of gyration distributions obtained from the dataset are equivalent between them, as well as projected area distributions are also equivalent among magnifications. Table 2 summarizes the *p*-values statistically obtained according to the Mann-Whitney tests performed for projected areas distributions. As all *p*-values found are greater than a significance level equal to 0.05, the difference between distributions obtained by the proposed algorithm at different magnifications are statistically negligible in this dataset. Furthermore, projected area distributions obtained by a manual inspection and by the methodology proposed here are also equivalent.

Table 3 summarizes the *p*-values statistically obtained according to the Mann-Whitney tests performed for diameter of gyration distributions. All p-values obtained are greater again than the significant level of 0.05. Therefore, all distributions of diameters of gyration are also equivalent.

Making use of the results obtained for both primary particles and agglomerates, fractal dimensions have been calculated according to Section 2.4. Fig. 10 (d) summarizes those results obtained. A great similarity between distributions is found. A mean value of D_f = 1.82, which is in the range of fractal dimensions reported in experimental studies about Cast burners (1.75) or those generated by diesel engines (1.67–2.05), was obtained [63,66]. Maximum absolute errors detected for a single agglomerate between data obtained by the proposed algorithm and results of manual inspection have been quantified in 19.81%, 19.05%, 18.66% for magnifications of 25,000, 20,000 and 15,000× respectively. Furthermore, mean absolute errors found at each magnification were equal to 5.13%, 5.06% and 5.77%.

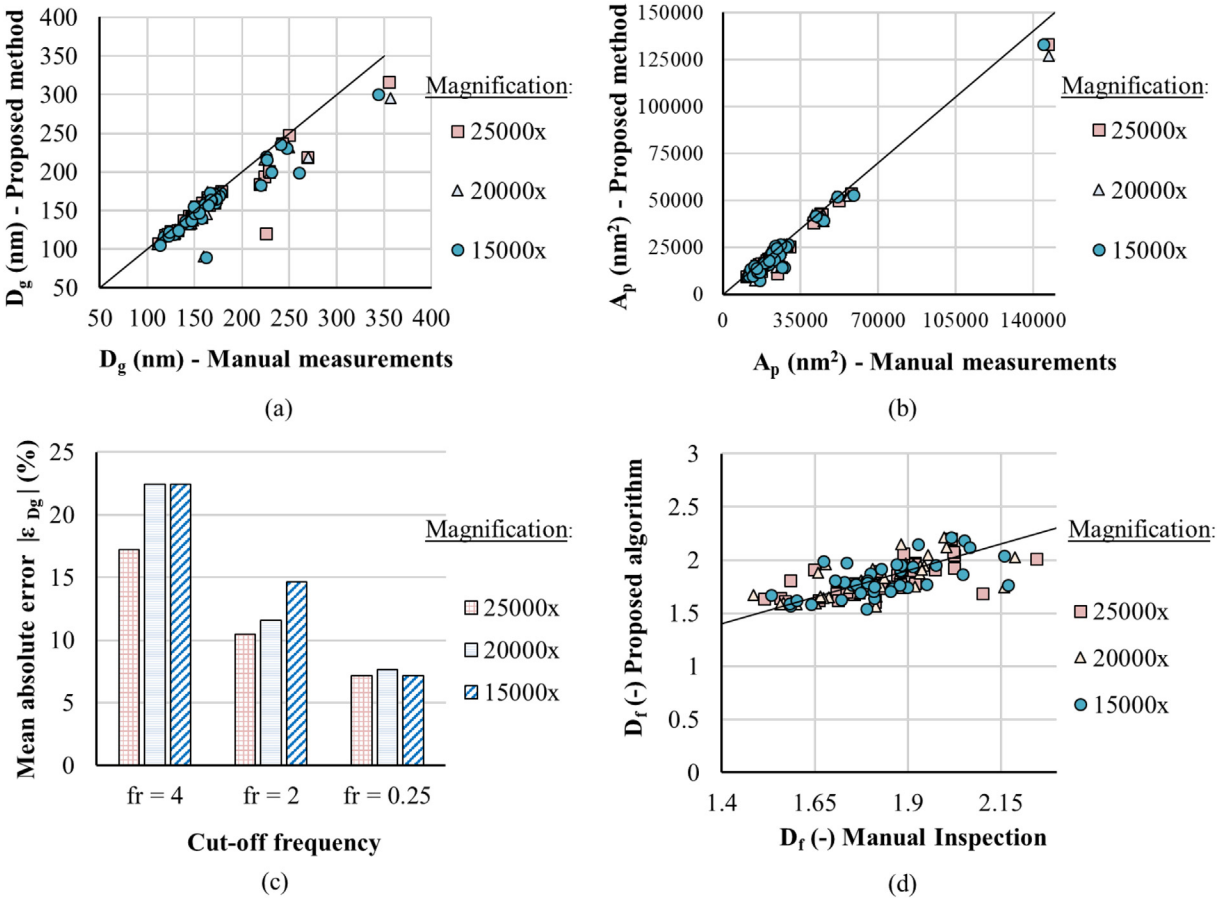

Fig. 10. (a) Comparison of diameters of gyration obtained using the proposed method versus manual measurements. (b) Comparison of projected areas obtained using the proposed method versus manual measurements. (c) Effect of the cut-off frequency on the mean absolute error of diameters of gyration for this dataset. (d) Comparison of fractal dimensions obtained using the proposed method versus manual measurements.

Table 2
P-values obtained from Mann-Whitney tests for distributions of projected areas.

		Magnification of 25,000×		Magnification of 20,000×		Magnification of 15,000×	
		Proposed algorithm	Manual inspection	Proposed algorithm	Manual inspection	Proposed algorithm	Manual inspection
Proposed algorithm	25,000×	1	0.067	0.973	0.110	0.935	0.061
	20,000×	0.973	0.118	1	0.179	0.996	0.338
	15,000×	0.935	0.116	0.996	0.173	1	0.097

Table 3
P-values obtained from Mann-Whitney tests for distributions of diameters of gyrations.

		Magnification of 25,000×		Magnification of 20,000×		Magnification of 15,000×	
		Proposed algorithm	Manual inspection	Proposed algorithm	Manual inspection	Proposed algorithm	Manual inspection
Proposed algorithm	25,000×	1	0.288	0.992	0.315	0.829	0.301
	20,000×	0.992	0.334	1	0.353	0.851	0.338
	15,000×	0.829	0.279	0.851	0.271	1	0.254

4.4. Execution time evaluation

Finally, once the accuracy of the results has been confirmed by comparison with a manual inspection process, it is necessary to estimate the saving in time that the algorithm proposed in this work involves. Execution times of main functions and elements of the code — recognition of the scale, inspection of primary particles, extraction of size descriptors of the agglomerates and fractal dimension calculation — have been monitored. Fig. 11 (a) summarizes those results, where the execution times required for the scale recognition (OCR and the Hough line transform), as well as the projected area extraction remain practically constant among micrographs (around 3 s and 9 s respectively). The timed need for the iterative process of the fractal dimension calculation is practically insignificant (around 2 ms), and a greater variability was found for primary particle measurement as a greater number of particles detected require a greater number of mathematical calculations, at an approximately rate of 0.25–1.25 s per primary particle detected according to Fig. 11 (b). The time required by the observer to inspect the micrographs has also been monitored (excluding time needed for file and results management) and a much wider range is detected. This fact is conditioned by the size of the agglomerate: the larger the size and particles, the longer the time required by the operator to perform the measurements at a rate around 8–40 s per primary particle (Fig. 11 (b)). Furthermore, the use of the proposed algorithm turns out to be clearly superior to a manual inspection process in terms of time required for each analysis. By comparing time distributions shown in Fig. 11 (a) and (b), an average saving time of 91.5% is observed due the use of the presented algorithm instead of a traditional analysis.

5. Conclusions

In this study, an automatic algorithm for morphological soot descriptors extraction from TEM micrographs was presented. The proposed methodology involves the combined use of a Hough transform and an OCR algorithm for the TEM scale recognition. Thus, the algorithm is directly usable to any TEM image regardless of the magnification used to capture the agglomerate. Using a two-step circular Hough transform, spherical projections of primary particles are identified and measured. Furthermore, the use of low-level vision tools — spatial filters, Fourier domain filters, segmentation and labelling — allow the measurement of projected areas as well as diameter of gyrations. Finally, the fractal dimension is calculated as a projected-area derived measurement due to an iterative process.

A set of micrographs taken at three different magnifications during the microscope stage were used to evaluate the accuracy of the proposed method. Those images were taken in bright field mode with current densities between 150 and 180 pA/cm^2. The error regarding the recognition of the scale is around 2%, whereas statistical comparisons between results obtained through the image processing method presented here and visual inspections were performed, finding significant equivalences in terms of *p*-values greater than 0.05. Results regarding primary particles were mainly dependent on the minimum diameter to be detected, the threshold applied to the accumulator array (involving differences between mean diameters between 0.9 nm and — 3 nm for the range considered), and the radii histogram filtering applied. Area extraction mainly depends on the cut-off frequency, decreasing the mean absolute error for this dataset from around 20% to 8% as the

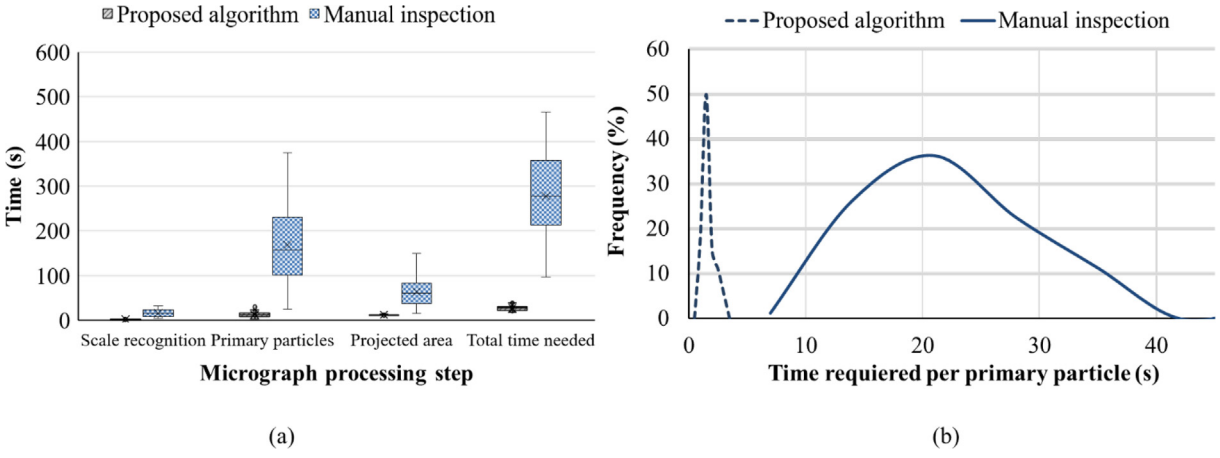

Fig. 11. (a) Time needed for both the manual inspection and the algorithm proposed in this work. (b) Time distribution required per primary particle detected.

cut-off values decrease from as high as fr = 4 to as low as fr = 0.25. Mean absolute error regarding fractal dimensions is around 5% for the dataset considered.

Based on the findings of this work, the merits of the proposed method are the ability to accurately extract a complete set of morphological descriptors of soot — diameter of primary particles, projected area, diameter of gyration and fractal dimension — with a significant saving in calculation time compared to manual inspections, where the average saving time is quantified in approximately 91.5%. As the main disadvantage, images taken in different conditions than those presented here probably need an adjustment of the parameters and computer vision processes in order to obtain satisfactory results.

Credit authorship contribution statement

Adrián Cabarcos: Methodology, Formal analysis, Data curation, Investigation, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Concepción Paz:** Conceptualization, Formal analysis, Supervision, Writing – original draft, Writing – review & editing, Resources, Project administration. **Raquel Pérez-Orozco:** Software, Validation, Writing – review & editing. **Jesús Vence:** Data curation, Validation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Otsu's method automatically finds an optimal threshold based on the observed distribution of pixel values. Assuming two classes (background and characters), the algorithm iterates until a minimum intra-class variance and therefore a maximum inter-class variance, σ_b^2, is achieved [52]. This inter-class variance is calculated according to Eq. (A.1)

$$\sigma_b^2(T) = \omega_0(T)\omega_1(T)[\mu_0(T) - \mu_1(T)]^2 \tag{A.1}$$

Where μ_0 and μ_1 are the mean grey levels of each class, and weights ω_0 and $\omega_1$are the probabilities of the two classes (background and characters), which are separated by the threshold T. Those class probabilities are computed from the 256-element histogram of the greyscale image as given by Eq. (A.2) and Eq. (A.3).

$$\omega_0(T) = \sum_{i=0}^{T-1} p(i) \tag{A.2}$$

$$\omega_1(t) = \sum_{i=T}^{256-1} p(i) \tag{A.3}$$

Where $p(i)$ denotes the histogram probabilities of the observed grey values. The Otsu's method involves an iterative process through all the possible threshold values as summarized below [41]:

1. Compute the intensity image histogram and probabilities of each intensity level.
2. Set up initial values for $\omega_i(0)$ and $\mu_i(0)$.
   2.1. Step through all possible thresholds from $T = 1$ to a T equal to the maximum intensity level.
   2.2. Update $\omega_i(T)$ and $\mu_i(T)$.
   2.3. Compute the inter-class variance, $\sigma_b^2(T)$, according to Eq. (A.1).
3. The desired threshold corresponds to the maximum inter-class variance, $\sigma_b^2(T)$.

Appendix B

The Hough line transform is a popular technique used to detect straight lines. For a given edge point (x, y) a set of pairs (ρ, θ), which correspond to candidate lines passing through the pixel, can be obtained. All those multiple pairs (ρ, θ) must verify the parametric equation shown in Eq. (B.1) and they will describe a sinusoidal curve on the Hough space. Note the use of polar coordinate representation of lines instead of the Cartesian ones to avoid infinite slope problems [52]

$$\rho = x\ cos\ \theta + y\ sin\ \theta \tag{B.1}$$

Therefore, each pixel will be characterised by a sinusoidal curve. The set of (ρ, θ) parameters for which intersection between those curves occurs will be the best candidate for straight lines. Fig. B1 outlines all this procedure in a simplified way for three points (colored in red, black and green in the figure). Each of these points defines a curve in the Hough parameter space (also colored in red, black and green). The intersection of those curves defines the polar coordinates of the line to be found in the image space. Note that, as the line sought for this specific application is the horizontal one that defines the image scale, only the horizontal solution of largest size is considered for this work.

Therefore, the Hough transform algorithm can be implemented as follows:

1. Initialize an accumulator matrix H(ρ,θ).
2. For each edge pixel detected through the Sobel filter in the image:
   a) For $\theta = \pm \pi/2$ rad.
      i) Compute the distance, ρ, from Eq. (B.1).
      ii) Increase the accumulator matrix H (ρ, θ) in that position by one.

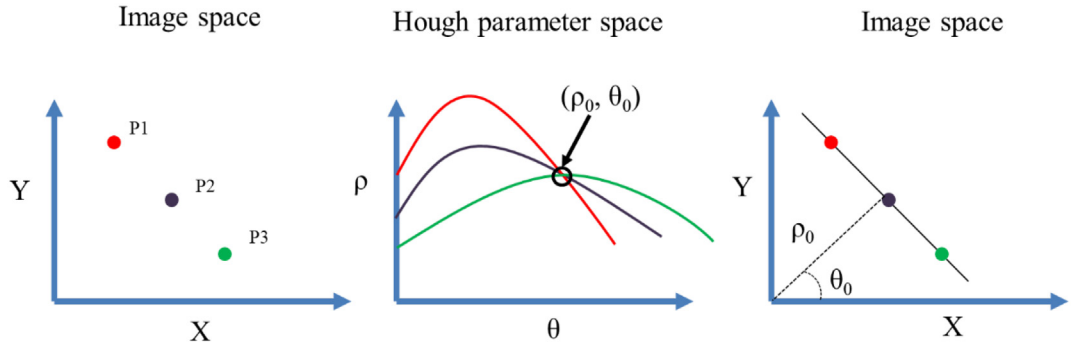

Fig. B1. Fundamentals of Hough Transform applied to line recognition.

3. Find the value of (ρ, θ) where the accumulator H (ρ, θ) reaches its maximum value.
4. Finally, Eq. (B.1) gives the selected line. The ending points of this segment are found through edge points belonging to the straight line obtained, for which Euclidean distance (in pixels) is maximum.

Appendix C

The closing operation is a mathematical morphology transformation defined as a dilation followed by an erosion [52]. The morphological dilation sets a pixel at (x,y) position to the maximum over all pixels in the neighbourhood centred at (x,y). Thus, bright regions are enlarged and dark ones are shinked. The resulting image is then eroded, which sets a pixel at (x,y) position to the minimum over all pixels in the neighbourhood. In mathematical set theory terms, the erosion of an image A by a structuring element B is given by Eq. (C.1), which states that the erosion of A by B is the set of points z such that the translation of B by z is contained in A

$$(A \ominus B) = \{z | (B)_z \subseteq A\} \tag{C.1}$$

Otherwise, the dilation, as given by Eq. (C.2), is the set composed by all elements of A which overlap with $\widehat{B}$ by at least one element.

$$(\mathrm{A} \oplus \mathrm{B}) = \{\mathrm{z} | \left[\left(\widehat{\mathrm{B}}\right)_{\mathrm{z}} \cap \mathrm{A}\right] \subseteq \mathrm{A}\} \tag{C.2}$$

The use of this sequence of basic morphological transformations involves a closing operation, given by Eq. (C.3).

$$A \cdot B = (A \oplus B) \ominus \mathrm{B} \tag{C.3}$$

Appendix D

The Mann-Whitney test involves a non-parametric comparison of two groups of paired data. It allows testing the null hypothesis of equal populations using a U statistic, which can be calculated for each group according to Eq. (D.1)

$$U_i = n_1 n_2 + \frac{n_i(n_i+1)}{2} - \sum R_i \tag{D.1}$$

Where U_i is the statistic of each population, n_1 and n_2 denotes the number of values from the first and the second groups, and $\sum R_i$ denotes the sum of the ranks from the group i of interest. The smaller of the two groups is the obtained result of U, and the z^* score of the test can be calculated according to Eq. (D.2) [64].

$$z^* = \frac{\left(U - \frac{n_1 n_2}{2}\right)}{\sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}}} \tag{D.2}$$

Finally, Eq. (D.3) specifies the general equation to calculate the p value according to a two-tailed test:

$$p\text{-}value = 2\,(1 - cdf(|z^*|)) \tag{D.3}$$

Where *cdf* denotes the cumulative distribution function of a given statistic z^*.

References

[1] D. Grantz, J.H. Garner, D. Johnson, Ecological effects of particulate matter, Environ. 29 (2003) 213–239, https://doi.org/10.1016/S0160-4120(02)00181-2.
[2] J. Schröder, K. Welsch-Pausch, M.S. McLachlan, Measurement of atmospheric deposition of polychlorinated dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs) to a soil, Atmos. Environ. 31 (1997) 2983–2989, https://doi.org/10.1016/S1352-2310(97)00107-6.
[3] C. Lachenmyer, Urban measurements of outdoor-indoor PM2.5 concentrations and personal exposure in the deep south. Part I. pilot study of mass concentrations for nonsmoking subjects, Aerosol Sci. Technol. 32 (2000) 34–51, https://doi.org/10.1080/027868200303911.
[4] C.A. Pope III, R.T. Burnett, M.J. Thun, E.E. Calle, D. Krewski, K. Ito, G.D. Thurston, Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution, JAMA. 287 (2002) 1132–1141, https://doi.org/10.1001/jama.287.9.1132.
[5] N. Hooftman, M. Messagie, J.V. Mierlo, T. Coosemans, A review of the European passenger car regulations – real driving emissions vs local air quality, Renew. Sust. Energ. Rev. 86 (2018) 1–21, https://doi.org/10.1016/j.rser.2018.01.012.
[6] Z. Mansurov, Soot formation in combustion processes (review), combust explos, Shock Waves 41 (2005) 727–744, https://doi.org/10.1007/s10573-005-0083-2.
[7] H.B. Palmer, C.F. Cullis, The formation of carbon from gases, Chem. Phys. Carbon. 1 (1965) 265–325.
[8] B.S. Haynes, H.G. Wagner, Soot formation, Prog. Energy Combust. Sci. 7 (1981) 229–273, https://doi.org/10.1016/0360-1285(81)90001-0.
[9] S.V. Shurupov, P.A. Tesner, Soot formation during isothermal pyrolysis of carbon tetrachloride and methane-carbon tetrachloride mixture, Symp. Combust. 27 (1998) 1581–1588, https://doi.org/10.1016/S0082-0784(98)80567-5.
[10] M. Frenklach, Reaction mechanism of soot formation in flames, Phys. Chem. Chem. Phys. 4 (2002) 2028–2037, https://doi.org/10.1039/B110045A.
[11] P. Verma, E. Pickering, N. Savic, A. Zare, R. Brown, Z. Ristovski, Comparison of manual and automatic approaches for characterisation of morphology and nanostructure of soot particles, Aerosol Sci. Technol. 136 (2019) 91–105, https://doi.org/10.1016/j.jaerosci.2019.07.001.
[12] P. Toth, D. Jacobsson, M. Ek, H. Wiinikka, Real-time, in situ, atomic scale observation of soot oxidation, Carbon. 145 (2019) 149–160, https://doi.org/10.1016/j.carbon.2019.01.007.
[13] A.B. Palotás, L.C. Rainey, C.J. Feldermann, A.F. Sarofim, J.B. Vander Sande, Soot morphology: an application of image analysis in high-resolution transmission electron microscopy, Microsc. Res. Tech. 33 (1996) 266–278, https://doi.org/10.1002/(SICI)1097-0029(19960215)33:3<266::AID-JEMT4>3.0.CO;2-O.
[14] M. Lapuerta, F.J. Martos, J.M. Herreros, Effect of engine operating conditions on the size of primary particles composing diesel soot agglomerates, J. Aerosol Sci. 38 (2007) 455–466, https://doi.org/10.1016/j.jaerosci.2007.02.001.
[15] J. Wei, Y. Zeng, M. Pan, Y. Zhuang, L. Qiu, T. Zhou, Y. Liu, Morphology analysis of soot particles from a modern diesel engine fueled with different types of oxygenated fuels, Fuel. 267 (2020), 117248 https://doi.org/10.1016/j.fuel.2020.117248.
[16] M.M. Maricq, Examining the relationship between black carbon and soot in flames and engine exhaust, Aerosol Sci. Technol. 48 (2014) 620–629, https://doi.org/10.1080/02786826.2014.904961.
[17] M.L. Botero, J. Akroyd, D. Chen, M. Kraft, J.R. Agudelo, On the thermophoretic sampling and TEM-based characterisation of soot particles in flames, Carbon. 171 (2021) 711–722, https://doi.org/10.1016/j.carbon.2020.09.074.
[18] F.X. Ouf, J. Yon, P. Ausset, A. Coppalle, M. Maillé, Influence of sampling and storage protocol on fractal morphology of soot studied by transmission electron microscopy, Aerosol Sci. Technol. 44 (2010) 1005–1017, https://doi.org/10.1080/02786826.2010.507228.
[19] M. Sakai, H. Iguma, K. Kondo, T. Aizawa, Nanostructure analysis of primary soot particles directly sampled in diesel spray flame via HRTEM, SAE Tech. Pap. 9 (2012) https://doi.org/10.4271/2012-01-1722.
[20] A.M. Brasil, T.L. Farias, M.G. Carvalho, A recipe for image characterization of fractal-like aggregates, J. Aerosol Sci. 30 (1999) 1379–1389, https://doi.org/10.1016/S0021-8502(99)00026-9.
[21] P.A. Marsh, A. Voet, T.J. Mullens, L.D. Price, Quantitative micrography of carbon black microstructure, Carbon. 9 (1971) 797–805, https://doi.org/10.1016/0008-6223(71)90013-3.
[22] S. Kook, Q. Renlin Zhang, T. Nian Chan, K. Aizawa, L.M. Kondo, E. Pickett, G. Cenker, O. Bruneaux, J. Andersson, E. Pagels, Z. Nordin, Automated detection of primary particles from transmission electron microscope (TEM) images of soot aggregates in diesel engine environments, SAE Int. J. Engines 9 (2015) 279–296, https://doi.org/10.4271/2015-01-1991.
[23] K. Kondo, T. Aizawa, S. Kook, L. Pickett, Uncertainty in sampling and TEM analysis of soot particles in diesel spray flame, SAE Tech. Pap. (2013) https://doi.org/10.4271/2013-01-0908.
[24] K. Yehliu, R.L.V. Wal, A.L. Boehman, Development of an HRTEM image analysis method to quantify carbon nanostructure, Combust Flame. 158 (2011) 1837–1851, https://doi.org/10.1016/j.combustflame.2011.01.009.
[25] S.A. Pfau, A.L. Rocca, M.W. Fay, Quantifying soot nanostructures: importance of image processing parameters for lattice fringe analysis, Combust Flame. 211 (2020) 430–444, https://doi.org/10.1016/j.combustflame.2019.10.020.
[26] A. Sharma, T. Kyotani, A. Tomita, A new quantitative approach for microstructural analysis of coal char using HRTEM images, Fuel. 78 (1999) 1203–1212, https://doi.org/10.1016/S0016-2361(99)00046-0.
[27] A. Galvez, N. Herlin-Boime, C. Reynaud, C. Clinard, J.-N. Rouzaud, Carbon nanoparticles from laser pyrolysis, Carbon. 40 (2002) 2775–2789, https://doi.org/10.1016/S0008-6223(02)00195-1.
[28] M. Lapuerta, O. Armas, A. Gómez, Diesel particle size distribution estimation from digital image analysis, Aerosol Sci. Technol. 37 (2003) 369–381, https://doi.org/10.1080/02786820300970.
[29] R. Dastanpour, J.M. Boone, S.N. Rogak, Automated primary particle sizing of nanoparticle aggregates by TEM image analysis, Powder Technol. 295 (2016) 218–224, https://doi.org/10.1016/j.powtec.2016.03.027.
[30] I. Grishin, K. Thomson, F. Migliorini, J. Sloan, Application of the Hough transform for the automatic determination of soot aggregate morphology, Appl. Opt. 51 (2012) 610–620, https://doi.org/10.1364/AO.51.000610.

[31] M. Mirzaei, H.K. Rafsanjani, An automatic algorithm for determination of the nanoparticles from TEM images using circular hough transform, Micron. 96 (2017) 86–95, https://doi.org/10.1016/j.micron.2017.02.008.
[32] M. Frei, F.E. Kruis, Fully automated primary particle size analysis of agglomerates on transmission electron microscopy images via artificial neural networks, Powder Technol. 332 (2018) 120–130, https://doi.org/10.1016/j.powtec.2018.03.032.
[33] T.A. Sipkens, S.N. Rogak, Technical note: using k-means to identify soot aggregates in transmission electron microscopy images, J. Aerosol Sci. 152 (2021), 105699 https://doi.org/10.1016/j.jaerosci.2020.105699.
[34] H. Yuen, J. Princen, J. Illingworth, J. Kittler, Comparative study of Hough transform methods for circle finding, Image Vis. Comput. 8 (1990) 71–77, https://doi.org/10.1016/0262-8856(90)90059-E.
[35] T.J. Atherton, D.J. Kerbyson, Size invariant circle detection, Image Vis. Comput. 17 (1999) 795–803, https://doi.org/10.1016/S0262-8856(98)00160-7.
[36] M. Altenhoff, S. Aßmann, C. Teige, F.J.T. Huber, S. Will, An optimized evaluation strategy for a comprehensive morphological soot nanoparticle aggregate characterization by electron microscopy, J. Aerosol Sci. 139 (2020), 105470 https://doi.org/10.1016/j.jaerosci.2019.105470.
[37] T.A. Sipkens, M. Frei, A. Baldelli, P. Kirchen, F.E. Kruis, S.N. Rogak, Characterizing soot in TEM images using a convolutional neural network, Powder Technol. 387 (2021) 313–324, https://doi.org/10.1016/j.powtec.2021.04.026.
[38] M. Lapuerta, R. Ballesteros, F.J. Martos, A method to determine the fractal dimension of diesel soot agglomerates, J. Colloid Interface Sci. 303 (2006) 149–158, https://doi.org/10.1016/j.jcis.2006.07.066.
[39] F.J. Martos, M. Lapuerta, J.J. Expósito, E. Sanmiguel-Rojas, Overestimation of the fractal dimension from projections of soot agglomerates, Powder Technol. 311 (2017) 528–536, https://doi.org/10.1016/j.powtec.2017.02.011.
[40] R. Smith, An overview of the tesseract OCR engine, Ninth International Conference on Document Analysis and Recognition 2007, pp. 629–633, https://doi.org/10.1109/ICDAR.2007.4376991.
[41] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Sys. Man. Cyber. 9 (1979) 62–66, https://doi.org/10.1109/TSMC.1979.4310076.
[42] H.-W. Yoo, H.-J. Ryoo, D.-S. Jang, Gradual shot boundary detection using localized edge blocks, Multimed. Tools Appl. 28 (2006) 283–300, https://doi.org/10.1007/s11042-006-7715-8.
[43] P. Suau, Adapting hausdorff metrics to face detection systems: A scale-normalized hausdorff distance approach, in: C. Bento, A. Cardoso, G. Dias (Eds.), Progress in Artificial Intelligence, 12th Portuguese Conference on Artificial Intelligence, Springer 2005, pp. 76–86.
[44] H. Ye, G. Shang, L. Wang, M. Zheng, A new method based on hough transform for quick line and circle detection, 8th International Conference on Biomedical Engineering and Informatics, IEEE 2015, pp. 52–56, https://doi.org/10.1109/BMEI.2015.7401472.
[45] R. Srisha, A. Khan, Morphological operations for image processing: understanding and its applications, National Conference on VLSI, Signal processing & Communications 2013, pp. 17–19.
[46] R. Dastanpour, S. Rogak, Observations of a correlation between primary particle and aggregate size for soot particles, Aerosol Sci. Technol. 48 (2014) 1043–1049, https://doi.org/10.1080/02786826.2014.955565.
[47] D. Ioannou, W. Huda, A.F. Laine, Circle recognition through a 2D Hough transform and radius histogramming, Image Vis. Comput. 17 (1999) 15–26, https://doi.org/10.1016/S0262-8856(98)00090-0.
[48] T.J. Atherton, D.J. Kerbyson, Using phase to represent radius in the coherent circle Hough transform, IEE Colloquium on Hough Transform, 1993, pp. 5/1–5/4.
[49] P.M. Anderson, H. Guo, P.B. Sunderland, Repeatability and reproducibility of TEM soot primary particle size measurements and comparison of automated methods, J. Aerosol Sci. 114 (2017) 317–326, https://doi.org/10.1016/j.jaerosci.2017.10.002.
[50] H.A. Elsalamony, Detecting distorted and benign blood cells using the Hough transform based on neural networks and decision trees, in: L. Deligiannidis, H.R. Arabnia (Eds.), Emerging Trends in Image Processing, Computer Vision and Pattern Recognition, Morgan Kaufmann, Boston 2015, pp. 457–473, https://doi.org/10.1016/B978-0-12-802045-6.00030-2.
[51] H.-S. Shim, R.H. Hurt, N.Y.C. Yang, A methodology for analysis of 002 lattice fringe images and its application to combustion-derived carbons, Carbon. 38 (2000) 29–45, https://doi.org/10.1016/S0008-6223(99)00096-2.
[52] E.R. Davies, Machine vision, Machine Vision, Third editionMorgan Kaufmann, Burlington, 2005 https://doi.org/10.1016/B978-0-12-206093-9.X5000-X.
[53] R. Zhang, Y. Zhang, S. Kook, Comparison between in-flame and exhaust soot particles in a single-cylinder, light-duty diesel engine, 19 th Australasian Fluid Mechanics Conference2014.
[54] M. Lapuerta, J. Barba, A.D. Sediako, M.R. Kholghy, M.J. Thomson, Morphological analysis of soot agglomerates from biodiesel surrogates in a coflow burner, J. Aerosol Sci. 111 (2017) 65–74, https://doi.org/10.1016/j.jaerosci.2017.06.004.
[55] M. Lapuerta, F. Martos, G. Martín-González, Geometrical determination of the lacunarity of agglomerates with integer fractal dimension, J. Colloid Interface Sci. 346 (2010) 23–31, https://doi.org/10.1016/j.jcis.2010.02.016.
[56] C. Paz, E. Suárez, J. Vence, J. Hoard, Evolution of EGR cooler deposits under hydrocarbon condensation: analysis of local thickness, roughness, and fouling layer density, Int. J. Therm. Sci. 161 (2021), 106744 https://doi.org/10.1016/j.ijthermalsci.2020.106744.
[57] C. Paz, E. Suárez, J. Vence, A. Cabarcos, Fouling evolution on ribbed surfaces under EGR dry soot conditions: experimental measurements and 3D model validation, Int. J. Therm. Sci. 151 (2020), 106271 https://doi.org/10.1016/j.ijthermalsci.2020.106271.
[58] L. Jing, Standard combustion aerosol generator (SCAG) for calibration purposes, 3rd ETH Workshop Nanoparticle Measurement, ETH Hönggerberg Zürich, 1999.
[59] R.A. Dobbins, C.M. Megaridis, Morphology of flame-generated soot as determined by thermophoretic sampling, Langmuir. 3 (1987) 254–259, https://doi.org/10.1021/la00074a019.
[60] C. Paz, A. Cabarcos, J. Vence, M. Conde, Experimental study of the effect of hexadecane exposure of soot agglomerates: Analysis of themorphology and fractal dimension, 15th International Conference on Heat Transfer, Fluid Mechanics and Thermodynamics 2021, pp. 964–969.
[61] M. Lapuerta, F.J. Martos, J.J. Expósito, Morphological characterization of diesel soot agglomerates based on the Beer–Lambert law, Meas. Sci. Technol. 24 (2013), 035405 https://doi.org/10.1088/0957-0233/24/3/035405.
[62] B. Hu, U. Koylu, Size and morphology of soot particulates sampled from a turbulent nonpremixed acetylene flame, Aerosol Sci. Technol. 38 (2004) 1009–1018, https://doi.org/10.1080/027868290519111.
[63] L.A.A. Yahia, E. Gehin, B. Sagot, Application of the TheRmophoretic Annular Precipitator (TRAP) for the study of soot aggregates morphological influence on their thermophoretic behaviour, J. Aerosol Sci. 113 (2017) 40–51, https://doi.org/10.1016/j.jaerosci.2017.07.018.
[64] P.H. Kvam, B. Vidakovic, P.H. Kvam, B. Vidakovic, Rank tests, Nonparametric Statistics with Applications to Science and Engineering 2007, pp. 115–139, https://doi.org/10.1002/9780470168707.ch7.
[65] D.T. Mauger, G.L. Kauffman, Statistical considerations, in: W.W. Souba, D.W. Wilmore (Eds.),Surgical Research 2001, pp. 71–80, https://doi.org/10.1016/B978-012655330-7/50011-3.
[66] M. Lapuerta, R. Ballesteros, F. Martos, The effect of diesel engine conditions on the size and morphology of soot particles, Int. J. Veh. Des. 50 (2009) 91–106, https://doi.org/10.1504/IJVD.2009.024972.