Granulometric Analysis of Maltodextrin Particles Observed by Scanning Electron Microscopy

Antoine Bottenmuller, Léo Théodon, Johan Debayle, Daniel Tobón Vélez, Mallorie Tourbin, Christine Frances, Yann Gavet

Granulometric analysis of maltodextrin particles observed by scanning electron microscopy

Antoine Bottenmuller*[1], Léo Théodon[1], Johan Debayle[1], Daniel Tobón Vélez[2], Mallorie Tourbin[2], Christine Frances[2], and Yann Gavet †[1]

[1]CNRS, UMR 5307 LGF, Centre SPIN, Mines Saint-Etienne, Univ. Lyon, Saint-Etienne, France
[2]Laboratoire de Génie Chimique, Université de Toulouse, Toulouse, France

September 20, 2023

Abstract

Maltodextrin is a substance that is being increasingly used to preserve the physicochemical and biological properties of many active compounds. Therefore, determining the particle size distribution (PSD) of the particles having maltodextrin matrices is a crucial issue to control their end-use properties. This can be done directly by laser diffraction (LD) in a dry way or via SEM image analysis. In this paper, a new method of segmentation of quasi-circular particles from grayscale images called curvature analysis method (CAM) is proposed. This method is compared to two other widely used methods: Circular Hough Transform (CHT) and Stochastic Watershed (SW). It aims in particular to reduce the drawbacks of these two methods: a large number of false detections and an inaccuracy with respect to the mean particle size in case of a large number of overlapping particles. The method is validated using synthetic images generated with a model allowing to simulate gray level images similar to the maltodextrin particle images from the SEM. The CAM method is then applied to real images and the resulting PSD is compared to the one provided by the LD technique. Overall, the results obtained by the CAM method are much better than those proposed by CHT and SW on synthetic images and than the LD method on real images.

Keywords: granulometry, SEM, numerical simulation, overlapping particles.

1 Introduction

1.1 Context

Nowadays, an increasing number of active compounds coming out of discovery (food, pharmaceutics, cosmetics, etc.) have limited applications mainly due to problems of poor stability, water insolubility or high toxicity for example. In order to increase the bioavailability and improve the stability of these compounds, different nanotechnology-based delivery systems have been developed. Among these, nanoencapsulation represents an effective approach to improve solubility, minimize degradation process, reduce toxicity and control the active absorption and biological response of molecules like polyphenols [1]. Polyphenolic compounds among which the Chlorogenic acid (CGA) group, and particularly the 5-O-caffeoylquinic acid (5-CQA), exert important therapeutic roles such as antioxidant, antibacterial, anti-inflammatory or cardioprotective [2]. Encapsulation could be performed by spray-drying technique, using the Nano Spray-Dryer B-90 from Büchi Labortechnik AG (Fig. 1), which allows to coat different active ingredients (AI), producing dry nanoparticles [3].

The wall materials in which the AI is confined provide protection from the surrounding environment, preserving their chemical, physical and biological properties, as well as enable to control its release [6]. In this study, maltodextrin has been used as a wall material to coat the 5-CQA. When encapsulating, the release of the active molecule depends on several properties of the particles, especially their size distribution, which itself depends on the process parameters. It is consequently very important to determine the size distribution of the produced particles to optimize the delivery properties of the active molecule.

*antoine.bottenmuller@gmail.com
†gavet@emse.fr

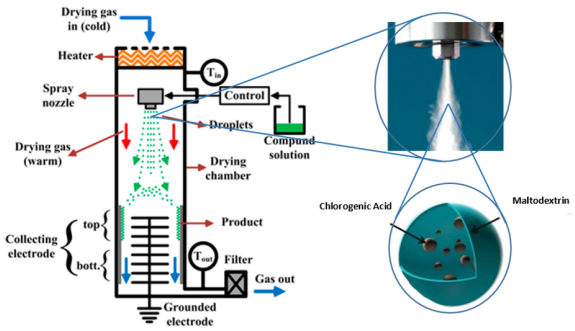


Figure 1: Schematic Nano Spray-Dryer set-up for the encapsulation of CGA by maltodextrin. Adapted from [4] and [5].

1.2 Images acquisition

A sample of the dry powder collected at the end of the spray-drying process has been analyzed by Scanning Electron Microscopy (SEM), in order to characterize their size and morphology. The equipment used is a Field-Emission Scanning Electron Microscope JEOL JSM 7100F. The dry particles are placed on a carbon film and metallized by coating with a thin layer of gold using a plasma coater, to avoid excessive electron absorption. The surface of the sample is scanned by an electron beam produced from a tungsten filament and the observation of the particle surface is done by detecting secondary electrons (with a voltage of 10 kV). Electron micrographs were taken at 20000X of magnification. The sample was scanned in order to take about twenty images representative of the entire population of particles constituting the powder (Fig. 6a).

2 Segmentation methods

To perform a granulometric analysis in such images, the use of a circle detection method is needed. Two of the most efficient and used methods from the literature for circle detection have been studied and tested for this problem: the Stochastic Watershed (SW) [7] and the Circular Hough Transform (CHT) [8]. As these two methods have some major problems in this situation, a new method has been developed to deal with them: the Curvature Analysis Method (CAM).

2.1 Stochastic watershed

Although the watershed segmentation is a strong method for contours extraction in smooth images for which the contours of the objects are clear enough and well defined, it is not robust for objects with major light effects and heterogeneous 3D aspects, and too many false contours may be detected [9], [10]. The Stochastic Watershed (SW) enhances the results by accumulating multiple watershed realizations with random markers [11]. A distance transform is then applied on these extracted contours [12], and the local maxima are calculated and defined as the centres of the different circles on the map, and the value of the distance gives their radius.

2.2 Circular Hough transform

The Circular Hough Transform (CHT) is one of the most widely used methods for circle detection [13], [14], [15]. After a low-pass filter on the gradient magnitude image and a binarization of the contours, a discrete three-dimensional CHT space is built and the circles are extracted from it [16].

2.3 Proposed method: curvature analysis

The two previous methods produce however either under or over-segmentation, especially in the presence of overlapping grains, and provide a grain size distribution far from reality. The Curvature Analysis Method (CAM) has been thought and developed to give a solution to this problem. This is a procedural algorithm divided into three steps:

1. From the original image, the so-called "linear minimum MSE map" is built; it represents the probability for each pixel to be on the contour of a grain, by minimizing a weighted Mean Squared Error (MSE) function in a pre-defined window for each pixel of the image (Fig. 2).

2. Then, a clean binary image of the thin arcs (portion of the contours) of the grains is extracted from the linear minimum MSE map (Fig. 3).

3. The associated set of circles best fitting with these arcs is finally estimated (Fig. 4).

These steps will now be detailed.

2.3.1 Building of the linear minimum MSE map

This map is constructed from a circular window scanning the grayscale gradient magnitude image ∇. The line going by the centre p_c of this window, and minimizing the weighted MSE regarding all pixels in the window, is obtained by regression. Its slope is denoted by α.

Thus, the error function E_{p_c} which must be minimized for each pixel p_c from the gradient magnitude ∇, is defined as follows:

$$E_{p_c}(\alpha) = \frac{1}{\sum_{i=1}^{n} \nabla(p_i)} \sum_{i=1}^{n} d(p_i, L_{p_c,\alpha})^2 \nabla(p_i) \tag{1}$$

with:

- $(p_i)_{i\in[\![1;n]\!]}$ the family of the n pixels surrounding the centre p_c in the pre-defined window;
- $\nabla(p_i)$ the gray value of the pixel p_i in the gradient magnitude image;
- $L_{p_c,\alpha}$ the line of slope $\alpha \in]-\infty;+\infty]$ crossing the point p_c in the image's coordinates ;
- $d(p_i, L_{p_c,\alpha})$ the Euclidean distance between the point p_i and the line $L_{p_c,\alpha}$ in the 2D Euclidean vector space:

$$d(p_i, L_{p_c,\alpha}) = \frac{(y_i - y_c) - \alpha(x_i - x_c)}{\sqrt{1+\alpha^2}}$$

By associating to each pixel p_c the corresponding value of the minimum error $E_{p_c}(\alpha)$, the "linear minimum MSE map" (Fig. 2e) is finally built and linearly normalized (its values go from 0 for the highest MSE values to 1 for the lowest ones).

2.3.2 Extraction of the thin arcs of the grains' contours

The second step's is extracting the binary image of the thin arcs of the grains' contours from the linear minimum MSE map previously built. It is divided into five sub-steps:

1. First, the linear minimum MSE map is binarized through a K-means clustering on the gray values of the map's pixels, with $K = 2$ in our case. (Fig. 3a).

2. Then the topological skeletonization [17] is applied on this binary image of the grains' contours, to get the thin topology-preserving version of these contours (Fig. 3b).

3. This skeleton is then cleaned by removing the isolated binary objects for which the area and the maximum Feret diameter are smaller than defined thresholds (Fig. 3c).

4. The remaining arcs are then split by applying an intersection detector: each positive pixel surrounded by three positive pixels or more is considered as an intersection, and is then removed (green parts in Fig. 3d).

5. Again, the arcs are split, according to a curvature criterion (red parts in Fig. 3d). Only arcs with a quite constant curvature will be remaining (Fig. 3e).

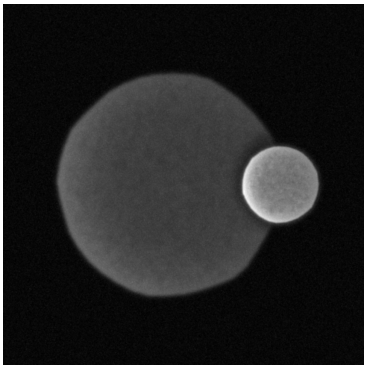
2.3.3 Circles association and rearrangement

The last step aims to get the final detected circles from the binary thin arcs of the grains' contours. It is also divided into four sub-steps:

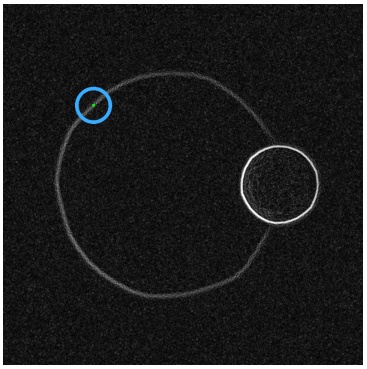
1. First, to each distinct arc is associated one unique circle, defined by the position of its centre $p_c = (x_c, y_c)$ and its radius r, computed by considering the arc A as a 2D point cloud ($A = (p_i)_{i\in[\![1;n]\!]}$) and minimizing the error function $E_{\text{circle}}(p_c, r)$, which has a well-defined solution to its minimum, defined as follows:

$$E_{\text{circle}}(p_c, r) = \frac{1}{n} \sum_{i=1}^{n} ((x_i - x_c)^2 + (y_i - y_c)^2 - r^2)^2 \tag{2}$$

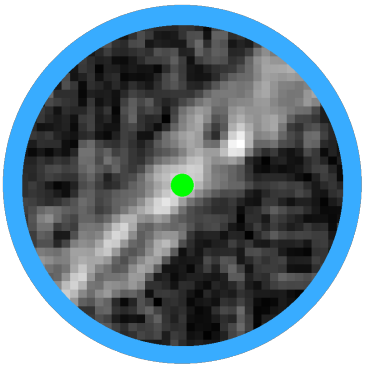
   A threshold on E_{circle} removes the less probable circles. The first remaining circles can be seen on Fig. 4a.



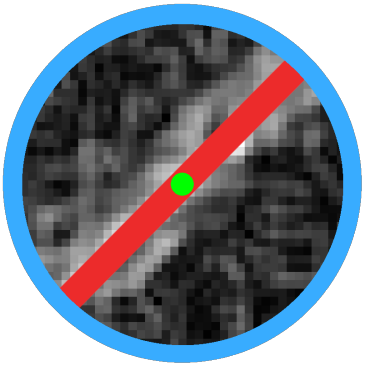
(a) Original image.



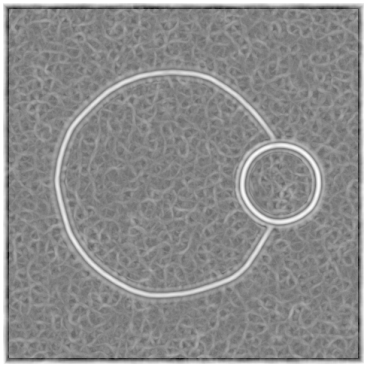
(b) Gradient magnitude with the circular window (blue) centred on a pixel p_c.



(c) View on the one window with the centre p_c (green).

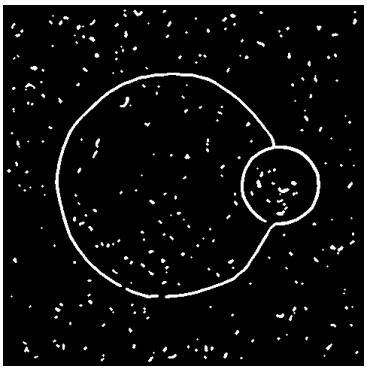


(d) Line $L_{p_c,\alpha}$ (red) going by the center p_c and minimizing the weighted MSE function.

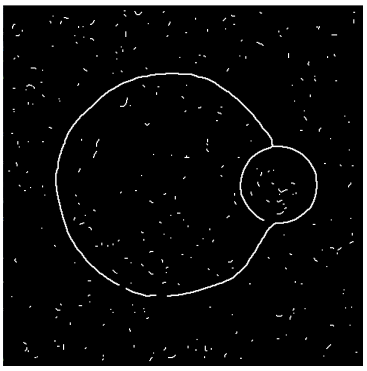


(e) Final computed linear minimum MSE map.

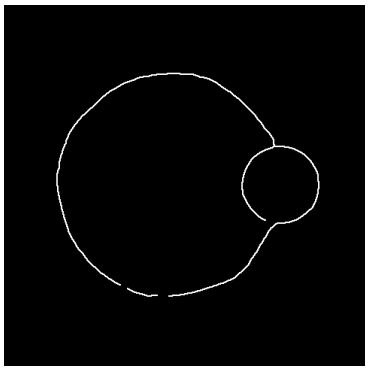
Figure 2: Construction steps of the linear minimum MSE map.



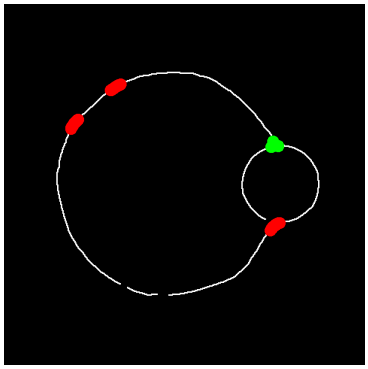
(a) Binarized image of the linear minimum MSE map from Fig. 2e.



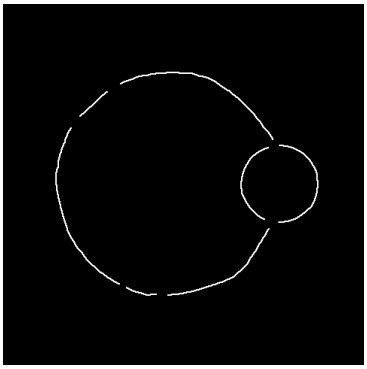
(b) Binary image skeleton.



(c) Clean skeleton.

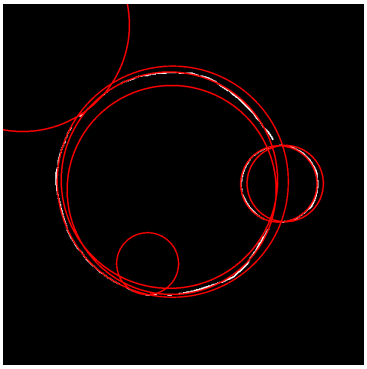


(d) Intersection area (green) and curvature irregularities (red) on clean skeleton.

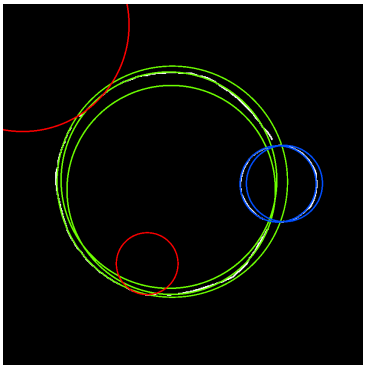


(e) Final split arcs.

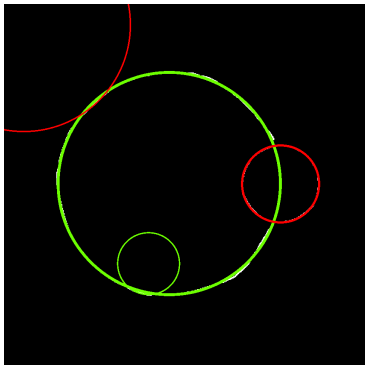
Figure 3: Steps of the extraction of the thin arcs.



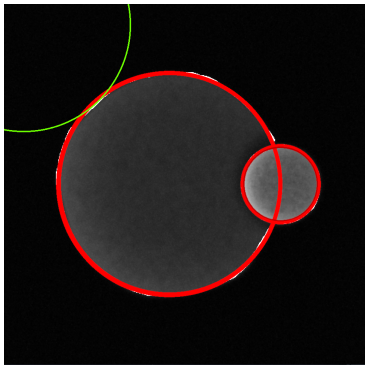
(a) All circles from arcs with an error below a threshold.



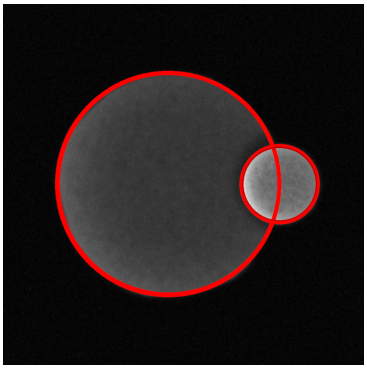
(b) Circles close enough in the space (x_c, y_c, r) to be merged, green and blue being two clusters of merging circles.



(c) Circles sharing arcs close enough from the circles edges to be merged, green being one cluster of two merging circles.

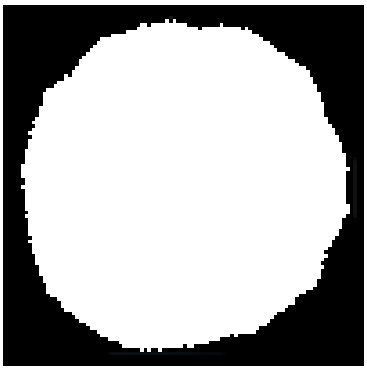


(d) Circles being brighter inside their arc than outside (green) are removed.

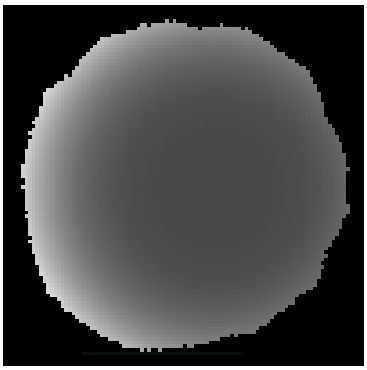


(e) Final detected circles.

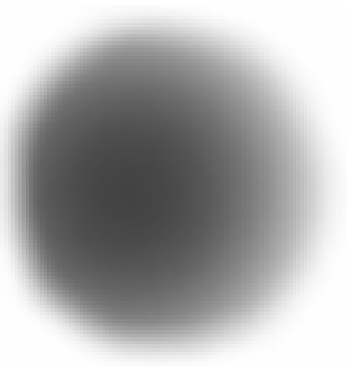
Figure 4: Steps of circles association and rearrangement.



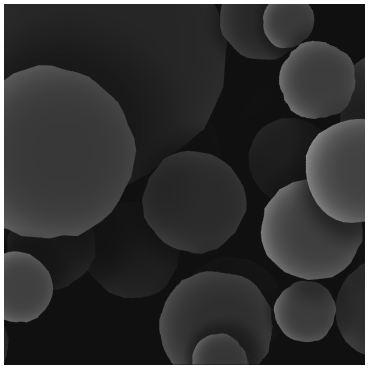
(a) Binary shape of a grain with random contours.



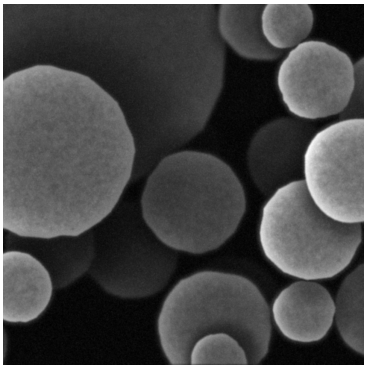
(b) Grayscale light effects added to the shape, where light comes from the left.



(c) Shadow of the grain computed from the shape and light direction.



(d) An image with several grains added one on top of the other.



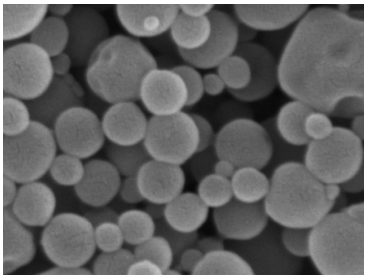
(e) Final image after being blurred and noised.

Figure 5: Example of the generation of a simulated image.

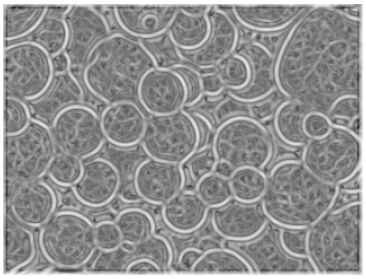
2. Close circles are then detected and merged with the same previous formula. The proximity of the circles is simply computed by the Euclidean distance in the space (x_c, y_c, r) (Fig. 4b).

3. Circles for which the arcs are fully on another circle are merged with the one circle (Fig. 4c).

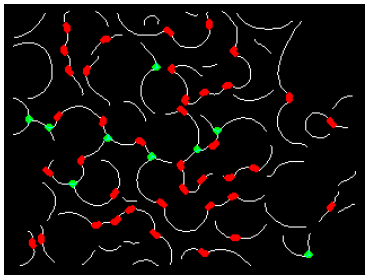
4. Finally, knowing the arcs associated to each computed circle, the circles for which the average gray value on the original image near their arcs is higher outside the circle than inside are removed (Fig. 4d). The remaining circles are then the definitive detected circles (Fig. 4e).



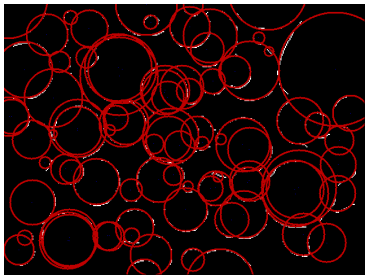
(a) Original image.



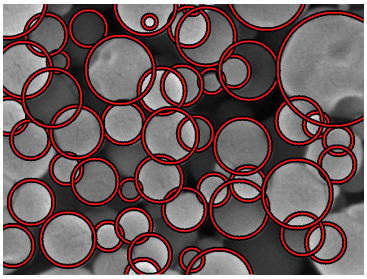
(b) Grayscale normalized linear minimum MSE map.



(c) Clean binary skeleton from MSE map with curvature splits (red) and intersections (green).



(d) Circles computed for all arcs from the skeleton.



(e) Final circles segmentation.

Figure 6: Example of application of the CAM on a real image: view of the different steps.

3 Validation

To test the three different segmentation methods in images of grains for which the ground truth (i.e. the true centre position and the true radius of the equivalent disk for each visible grain) is known, a grains simulation model has been developed, based on several properties of the real images. This model will be used to validate the results on the particle size distribution obtained with the three segmentation methods from the ground truth, and compare the accuracy of the three methods between each other.

3.1 Grains simulation model

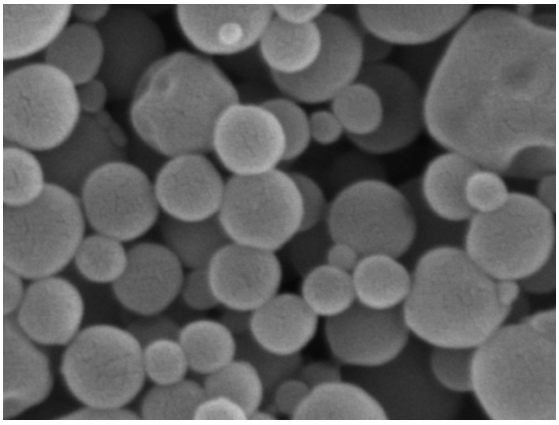
The grains simulation model is inspired by the dead leaves model [18], [19]. It takes three input parameters: the average number of grains n_{avg} on the image, the distribution law L_{radii} of the grains' radii, and the vector of the direction d_{light} of the light source in the half-sphere in the 3D space above the image's plan.

It starts by defining a definitive number of grains by choosing it through a uniform law between $n_{\text{avg}}/2$ and $3n_{\text{avg}}/2$.

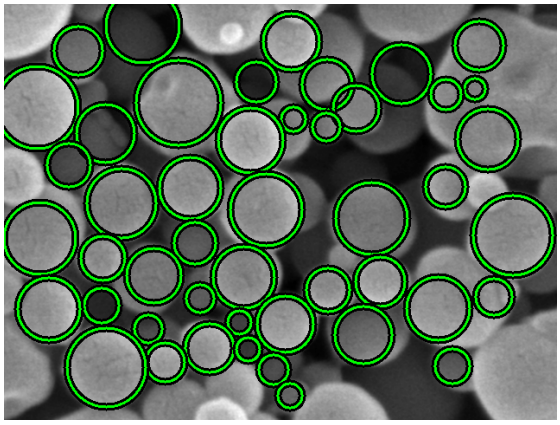
The grains are added iteratively to a grayscale image. Each grain is built and added over all the previous grains as follows (see also Fig. 5):

1. The random parameters of the disk, its radius and position, are first generated. The centre is here randomly chosen on the maxima of the distance map computed on the empty space in the image.

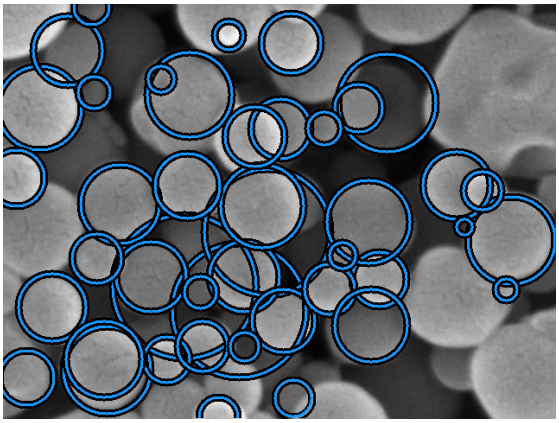
2. A random deformation of the shape is then computed, and the properties of the disk (radius, position) are updated according to this new shape. This allows to move the disk to avoid collisions with previous ones.



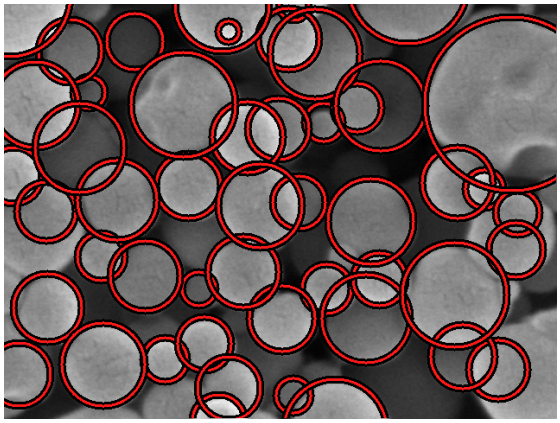
(a) Original image.



(b) Segmentation with SW ($N = 40$).



(c) Segmentation with CHT.



(d) Segmentation with CAM.

Figure 7: Example of results from the three segmentation methods on real image.

3. A gray value is generated as a function of the iterative step, and lighting effects (for volume simulation) and a shadow (based on lighting direction) are finally added.

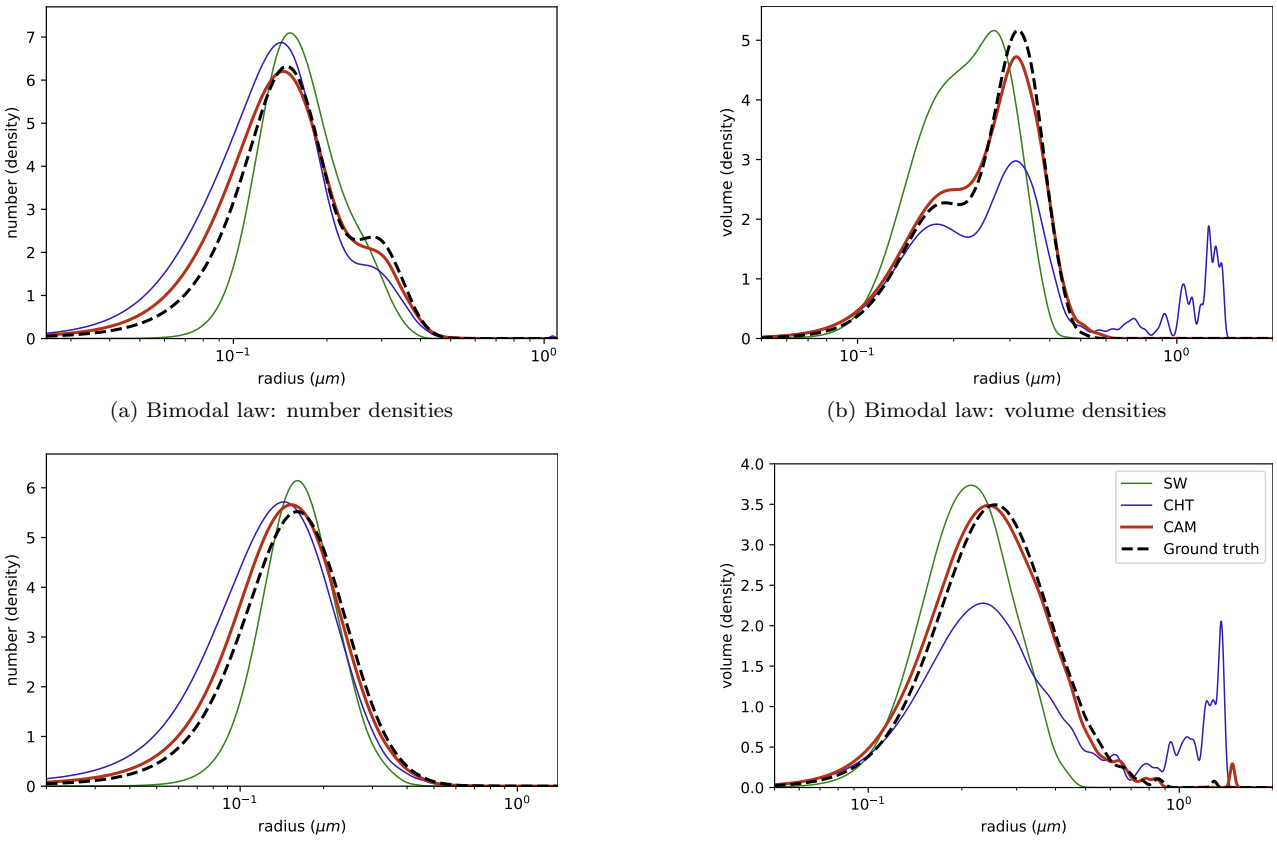
After all the grains are placed and drawn (Fig. 5d), the generated output image is blurred at different depth layers, and little blurred Gaussian noises are finally added to it (Fig. 5e).

3.2 Evaluation of the segmentation methods through the model

In this study, we consider two kinds of laws for the Particle Size Distribution (PSD): a bimodal law, and a log-normal law. For each law, 100 images have been generated with the grains simulation model. The ground truth (labels) for all the images is well known and concerns only the visible grains.

The three segmentation methods are then applied on all generated images, with the best inner parameters possible set for each method, resulting in a list of radii of detected circles.

The PSD densities, in number and in volume, of the three segmentation methods as well as the ground truth, are retrieved for the two laws. The four resulting diagrams are presented Fig. 8, and the values of the mean and the standard deviation (STD) of each distribution are given in the Tables 1 and 2.



(a) Bimodal law: number densities

(b) Bimodal law: volume densities

(c) Log-normal law: number densities

(d) Log-normal law: volume densities

Figure 8: Diagrams of densities (PSD) retrieved from the simulated images with SW, CHT and CAM methods.

The volume densities aim to highlight the largest radii, which are mostly false detections in our simulations when getting over 1μm, even if they aren't abundant in number. They show that, on the one hand, a noticeable proportion (in volume) of false large grains over 1μm of radius have been detected by the CHT, where the ground truth indicates almost no true circles. The SW and CAM, on the other hand, have almost no false detection over this threshold.

The three methods seem all to underestimate the size of the grains in average from the ground truth (see Tables). The graphs of densities in number show that CHT and CAM have both more false little grains than they should have, even if this is less noticeable for the CAM. But the global shape of the true distribution law can be identified on both methods' curves in the four graphs, where this one shape fades with the SW.

More generally, the SW's density curves seem to be quite concentrated around their mean value: their standard deviation is much smaller than the ground truth compared to the two other methods' ones, and their minimum and maximum radii never get far from the mean, preventing detection of much larger or much smaller true grains. On the contrary, the CHT seems to detect a noticeable amount of false grains in extreme radius

Properties	Ground truth	SW	CHT	CAM
Mean	0.199 μm	0.190 μm	0.280 μm	0.191 μm
STD	0.082 μm	0.058 μm	0.101 μm	0.082 μm

Table 1: Mean and STD of densities on bimodal law.

Properties	Ground truth	SW	CHT	CAM
Mean	0.199 μm	0.287 μm	0.282 μm	0.190 μm
STD	0.082 μm	0.054 μm	0.104 μm	0.081 μm

Table 2: Mean and STD of densities on log-normal law.

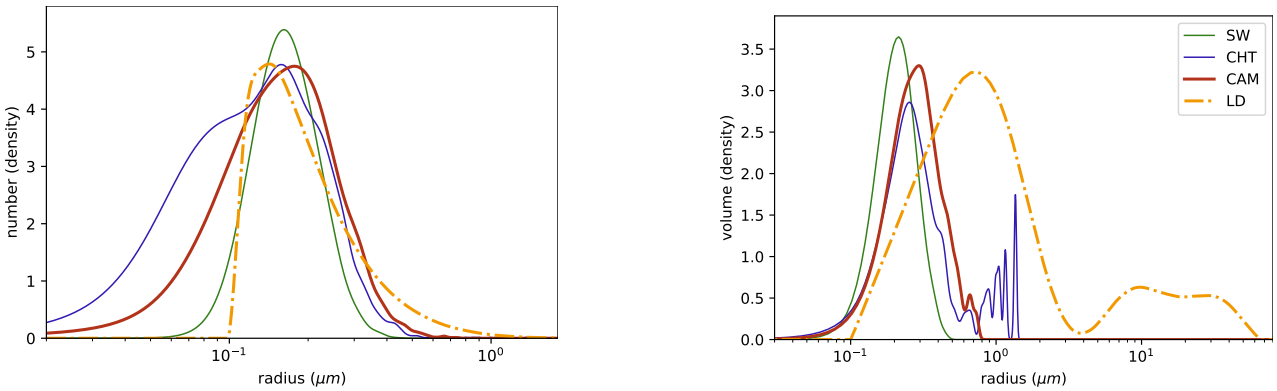
values despite the thresholds put in its parameters, and its large standard deviations illustrate well this. The CAM, on the other hand, seems to be more stable for both number and volume densities, despite mean values still a little smaller than the truth's ones, and its density curves, as well as its mean and standard deviation values, seem to be the closest ones from the ground truth's in every cases.

4 Application on real images

4.1 Results

A total of 20 real images of particles of maltodextrin observed by SEM is provided for this study. They all have the same scale. The three segmentation methods are applied (see Fig. 6 for the main processing steps of the CAM, and Fig. 7 for an example of segmentation comparison between the three methods on a real image), and just like for the validation part, the diagrams of the PSD from the lists of the detected circles' radii are built, for densities in both number and volume.

Ground data have been collected by Laser Diffraction (LD) in a dry way with the Mastersizer MS 3000 particle size analyser [20]. But as this method of particle size measurement suffers from a few defaults [21], it cannot be considered as ground truth. The resulting PSD is illustrated Fig. 9, and the corresponding means and STDs are given in Table 3.



(a) Real images: number densities

(b) Real images: volume densities

Figure 9: Diagrams of densities (PSD) on real images.

Like in the validation process, the graphs and the standard deviations show that the SW's density remains quite concentrated around its mean, where the CHT's one appears very large, with a high proportion of small (Fig. 9a) and large (Fig. 9b) radius values at the same time, and has even once again in its volume density what seems to be a significant proportion of false grains radii over 1μm, limited by the threshold given in its parameters. The CAM seems once again to be more moderated, extended in both small and large radii zones, but in reasonable proportions.

Although the LD's curve seems to be quite close to the ones from the segmentation methods in the graph in number, it is way far from these ones in the graph in volume, showing extremely large values of grains radii up to 60μm. When looking at real images, no grain seems to have a radius over 1μm, and, in reality, no grain could be as large as the LD's curve indicates. It is likely due to grains clusters which are considered as only one single grain when physically analysed by this technique to get its corresponding radius.

Properties	LD	SW	CHT	CAM
Mean	0.217 μm	0.285 μm	0.284 μm	0.207 μm
STD	0.133 μm	0.052 μm	0.105 μm	0.093 μm

Table 3: Mean and STD of densities on real images.

Based on the CAM results, it can then be concluded that the real PSD on the particles of maltodextrin must be close to a log-normal distribution law, for which the mean is around 0.207μm and the standard deviation around 0.093μm.

4.2 Discussion

When applying the SW, a major problem appears: if the number of markers is small, only the main contours of the largest grains are extracted, the tiniest ones having just a little chance to get targeted by a marker, and when increasing this number N, false contours begin to appear very quickly inside the largest grains. Because the size of the grains is far from being constant and because there are many overlapping cases in such images, it seems that there is no proper value for N to detect small and large grains at the same time. These observations could explain the fact that the SW's density curve seems to be quite concentrated around its mean.

Although the CHT seems to be robust on images with well-defined and separated objects, it seems to be quite defective when the grains are condensed and overlapping, and when heterogeneous 3D aspects and major light effects affect the grains appearance. The 3D CHT space building process doesn't consider contours' properties (such as the curvature) and does not make the difference between the pixels of the contours from two different objects. It can therefore easily mix them and detect false circles between two or more different grains if they are close enough or overlapping. This could explain the relatively high proportion of very small and very large false grains in the granulometric analysis.

As the CAM analyses only parts (the arcs) of the contours of the grains, and as the grains are far from being perfectly circular, the detected circle completely depends on the curvature of the one part of contour: if the arc has a too small radius of curvature compared to the radius of the true grain to which it belongs, the associated circle will be smaller than the true corresponding grain, and, on the contrary, if the arc is almost flat, the associated circle will be bigger. Moreover, when the arc is only a little piece of the grain's contours and when its curvature is not steady, the eq. 2 tends to give a smaller radius than the ground truth. This could explain the underestimates of the mean radius of the circles detected by the CAM.

This application finally shows that the LD technique provides a PSD way far from the ones obtained by any of the three segmentation methods and even from the observations by eye. It can not then be considered as trustworthy for such types of particles and configuration.

5 Conclusion

The particles of maltodextrin from this study are facing particular features: the irregularity of the contours of the grains, the rough aspect of their surface combined with strong light effects, and the way the particles are condensed and mostly overlapping, make the granulometric analysis on such images quite challenging.

A new segmentation approach - the Curvature Analysis Method (CAM) - has been developed to address the problems raised by the existing methods from the literature such as the Circular Hough Transform (CHT) and the Stochastic Watershed (SW). A 2D grains simulation model has also been created to assess and compare the accuracy of the three segmentation methods with a known ground truth.

The analysis of the PSD obtained from each of the three methods on simulated images shows that the CAM's density curves are more accurate than the ones obtained from the SW and the CHT, with less false detections and a better estimate of the mean radius and of the standard deviation. The same results can be observed on real images.

As this study focuses only on topology-based segmentation methods, and to go further, deep learning approaches, such as CNN or U-Nets for objects localisation, could be used as segmentation tools and could give efficient and accurate results to this PSD evaluation problem.

References

[1] Conte, Valentino, D. Cristo, and D. Salle, "Polyphenols nanoencapsulation for therapeutic applications," *Journal of Biomolecular Research and Therapeutics*, 2016.

[2] M. Naveed, M. Naveed, V. Hejazi, M. Abbas, M. Abbas, M. Abbas, M. Abbas, A. A. Kamboh, G. J. Khan, M. Shumzaid, F. Ahmad, D. Babazadeh, X. FangFang, F. Modarresi-Ghazani, L. Wenhua, and Z. Xiaohui, "Chlorogenic acid (cga): A pharmacological review and call for further research." *Biomedicine & Pharmacotherapy*, 2018.

[3] C. Arpagaus, C. Arpagaus, A. Collenberg, D. Rütti, E. Assadpour, and S. M. Jafari, "Nano spray drying for encapsulation of pharmaceuticals." *International Journal of Pharmaceutics*, 2018.

[4] C. I. Piñón-Balderrama, C. Leyva-Porras, Y. Terán-Figueroa, V. Espinosa-Solis, C. Alvarez-Salas, and M. Z. Saavedra-Leos, "Encapsulation of active ingredients in food industry by spray-drying and nano spray-drying technologies," *Processes*, 2020.

[5] M. Cano-Sarabia and D. Maspoch, *Nanoencapsulation.* Dordrecht: Springer Netherlands, 2014, pp. 1–16. [Online]. Available: https://doi.org/10.1007/978-94-007-6178-0_50-2

[6] Z. Fang and B. Bhandari, "Encapsulation of polyphenols – a review," *Trends in Food Science and Technology*, 2010.

[7] J. Angulo and D. Jeulin, "Stochastic watershed segmentation." *International Symposium on Mathematical Morphology and Its Application to Signal and Image Processing*, 2007.

[8] D. Ioannou, W. Huda, and A. F. Laine, "Circle recognition through a 2d hough transform and radius histogramming," *Image and Vision Computing*, 1999.

[9] H. Zhang, Z. Tang, Y. Xie, X. Gao, Q. Chen, Q. Chen, and Q. Chen, "A watershed segmentation algorithm based on an optimal marker for bubble size measurement," *Measurement*, 2019.

[10] S. Lou, L. Pagani, W. Zeng, X. Jiang, X. J. Jiang, and P. J. Scott, "Watershed segmentation of topographical features on freeform surfaces and its application to additively manufactured surfaces," *Precision Engineering-journal of The International Societies for Precision Engineering and Nanotechnology*, 2020.

[11] F. Malmberg, C. L. L. Hendriks, and R. Strand, "Exact evaluation of targeted stochastic watershed cuts," *Discrete Applied Mathematics*, 2017.

[12] R. Fabbri, L. da Fontoura Costa, J. C. Torelli, and O. M. Bruno, "2d euclidean distance transform algorithms: A comparative survey," *ACM Computing Surveys*, 2008.

[13] M. Smereka and I. Duleba, "Circular object detection using a modified hough transform," *International Journal of Applied Mathematics and Computer Science*, 2008.

[14] R. Yuan, J. Zhang, and J.-M. Zuo, "Lattice strain mapping using circular hough transform for electron diffraction disk detection." *Ultramicroscopy*, 2019.

[15] K. Adem, "Impact of activation functions and number of layers on detection of exudates using circular hough transform and convolutional neural networks," *Expert Systems With Applications*, 2022.

[16] M. Mirzaei and H. K. Rafsanjani, "An automatic algorithm for determination of the nanoparticles from tem images using circular hough transform," *Micron*, 2017.

[17] J. K. Lakshmi, M. Punithavalli, and M. Punithavalli, "A survey on skeletons in digital image processing," *International Conference on Digital Image Processing*, 2009.

[18] C. Bordenave, Y. Gousseau, and F. Roueff, "The dead leaves model: a general tessellation modeling occlusion," *Advances in Applied Probability*, 2006.

[19] D. Jeulin, "Dead leaves models: From space tessellations to random functions," *null*, 2021.

[20] G. S. Faé, F. Montes, F. Montes, E. Bazilevskaya, R. M. Añó, and A. R. Kemanian, "Making soil particle size analysis by laser diffraction compatible with standard soil texture determination methods," *Soil Science Society of America Journal*, 2019.

[21] G. Varga, F. Gresina, G. Újvári, J. Kovács, and Z. Szalai, "On the reliability and comparability of laser diffraction grain size measurements of paleosols in loess records," *Sedimentary Geology*, 2019.