

Une approche pour l'analyse granulométrique de particules condensées sur des images en niveaux de gris

Antoine BOTTENMULLER¹ Léo THÉODON¹ Johan DEBAYLE¹
Daniel TOBÓN VÉLEZ² Mallorie TOURBIN² Christine FRANCES² Yann GAVET¹

¹ MINES Saint-Etienne, CNRS, UMR 5307 LGF, Centre SPIN, Saint-Etienne, France

² Laboratoire de Génie Chimique, Université de Toulouse, Toulouse, France

Résumé – ImageJ Particle Sizer est probablement l'un des outils les plus utilisés pour effectuer une analyse granulométrique sur des images en niveaux de gris. Bien qu'il s'agisse d'une méthode de segmentation performante pour des cercles bien distincts, elle produit une sur-segmentation lorsqu'elle est appliquée à des images dans lesquelles les grains sont condensés et se chevauchent. Dans cet article, nous proposons un nouvel algorithme de segmentation, la Méthode d'Analyse de Courbure (Curvature Analysis Method - CAM), pour faire face à ce problème. Un modèle stochastique de simulation de grains a été développé pour comparer la précision de la Distribution des Tailles des Particules (PSD) issue de la CAM à celle issue d'ImageJ, par rapport à la vérité terrain des images simulées. Ces deux méthodes sont ensuite appliquées à un ensemble d'images réelles de particules de maltodextrine synthétisées, et comparées à la PSD donnée par Diffraction Laser (DL). Les résultats montrent finalement que la CAM fournit des distributions plus justes qu'ImageJ sur les images simulées, et que la DL fournit une PSD loin de la vérité sur les images réelles.

Abstract – ImageJ Particle Sizer is probably one of the most known and used tools to perform a granulometric analysis in gray-scale images. Although it is a high-performance segmentation method for well separated circles, it produces over-segmentation when applied on images in which the grains are condensed and overlapping. In this paper, we propose a new segmentation algorithm, the Curvature Analysis Method (CAM), to face the problems raised by such images. A stochastic grains simulation model has been developed to compare the accuracy of the CAM's Particle Size Distribution (PSD) to the ImageJ's one, regarding the ground truth of simulated images. They are then applied on a provided set of real images of synthesised maltodextrin particles, and compared to the PSD given by Laser Diffraction (LD). The results finally show that the CAM provides much more accurate results than the CHT does on simulated images, and that the LD technique can not be considered as trustworthy as it gives a PSD far from the truth.

1 Introduction

En analyse d'images, la granulométrie est un domaine visant à mesurer la distribution en taille d'un ensemble de particules. Plusieurs outils ont été développés pour effectuer l'analyse granulométrique la plus précise possible dans des images en niveaux de gris, comme la Transformée de Hough Circulaire (Circular Hough Transform - CHT) [5], l'une des méthodes de segmentation les plus utilisées et efficaces pour la détection de cercles. Cette méthode est fournie dans ImageJ Particle Sizer [1], permettant de détecter directement à travers ce logiciel des cercles dans les images en niveaux de gris, et de calculer la Distribution des Tailles des Particules (Particle Size Distribution - PSD) à partir des rayons des cercles détectés.

Cependant, bien que cet outil soit performant pour des grains séparés et bien distincts [8], elle crée une sur-segmentation de l'image lorsque les particules commencent à se condenser et à se chevaucher. La diminution du facteur de sensibilité de cette méthode de détection de cercles entraîne rapidement, à l'inverse, une sous-segmentation de l'image.

Pour répondre à ce problème de sur-segmentation des images représentant des particules condensées en chevauchement, une nouvelle méthode de détection de cercles en niveaux de gris a été mise au point : la Méthode d'Analyse de Courbure (Curvature Analysis Method - CAM). (Partie 2)

Les performances de la CAM doivent alors être comparées à celles de la CHT à travers le logiciel ImageJ, afin de valider cette nouvelle méthode de segmentation. Pour ce faire, un

modèle de simulation stochastique de grains a été développé, générant des images réalistes de grains aléatoires condensés et superposés en niveaux de gris, et pour lesquelles la vérité terrain (i.e. le centre et le rayon de chacune des particules dessinées) est connue. En appliquant les deux méthodes de segmentation - la CAM et ImageJ -, aux images simulées, les PSD correspondantes sont construites et comparées à celle de la vérité terrain connue. (Partie 3)

Un ensemble d'images réelles de particules de maltodextrine synthétisées observées par Microscopie Electronique à Balayage (MEB) est fourni pour cette étude afin de tester cette nouvelle méthode sur un cas pratique réel. Les deux méthodes de segmentation lui sont alors appliquées, et leurs PSD sont comparées à celle fournie par la technique de Diffraction Laser (DL) [2]. Les résultats sont discutés en fin d'article. (Partie 4)

2 Une autre méthode de segmentation proposée : la CAM

La Méthode d'Analyse de Courbure (CAM) a été développée pour éviter les problèmes de sur-segmentation induites par ImageJ, qui semble associer les bords de différents grains à un seul et même cercle. Le but de la CAM est alors d'extraire les contours de chaque grain de l'image pour leur associer un unique cercle à partir d'une analyse de la courbure des arcs formant ces contours. Ceci en trois grandes étapes :

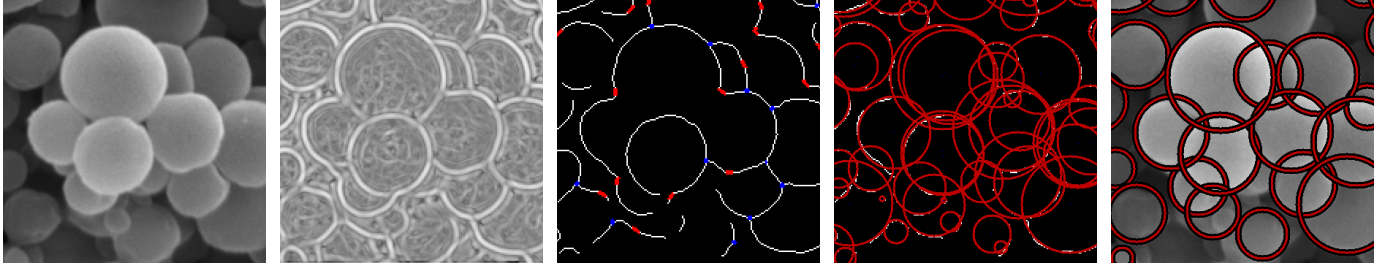


FIGURE 1 : Processus de détection de cercles par la méthode CAM. De gauche à droite : Image originale (gauche), carte des probabilités d'appartenance des pixels à un contour de grain (centre-gauche, en niveaux de gris inversés), analyse de la courbure des arcs du squelette des contours des grains (centre), ensemble des cercles associés aux arcs (centre-droite), et cercles finaux (droite).

1. D'abord, la "carte d'erreur minimale linéaire" est créée :

Elle représente, pour chaque pixel de l'image, une certaine probabilité d'appartenir au contour d'un grain. De l'image d'origine (Fig. 1, gauche), l'image amplitude du gradient ∇ est créée. Une fenêtre circulaire W est préalablement définie. Pour chaque pixel p_c de l'image, on cherche alors la pente α de la droite $L_{p_c, \alpha}$ traversant p_c , qui minimise la fonction d'erreur E_{p_c} définie dans la fenêtre circulaire W centrée en p_c par :

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

avec :

- $(p_i)_{i \in [1; n]}$ la famille des n pixels de l'image contenus dans la fenêtre W centrée en p_c ;
- $\nabla(p_i)$ la valeur de gris du pixel p_i dans l'image amplitude du gradient ∇ ;
- $L_{p_c, \alpha}$ la droite de pente $\alpha \in] -\infty; +\infty[$ dans les coordonnées de l'image, traversant le pixel p_c ;
- $d(p_i, L_{p_c, \alpha})$ la distance Euclidienne entre le pixel p_i et la droite $L_{p_c, \alpha}$ dans l'espace vectoriel 2D :

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

En associant à chaque pixel p_c son erreur $E_{p_c}(\alpha_{\min})$ minimale correspondante, et en prenant l'image inverse de celle résultante, la "carte d'erreur minimale linéaire" est ainsi construite (Fig. 1, centre-gauche).

2. Ensuite, les arcs des contours des grains sont extraits :

La "carte d'erreur minimale linéaire" précédemment créée est d'abord binarisée par un K-means [6] sur les valeurs de gris de l'ensemble des pixels de la carte. Dans notre cas, $K=2$, et seuls les pixels du cluster à la plus haute valeur de gris (i.e. faible erreur $E_{p_c}(\alpha_{\min})$) sont conservés. Une opération de skeletonisation [4] est alors appliquée sur l'image binaire, donnant les contours binaires affinés des grains. Après un nettoyage des petits éléments, un algorithme de détection local d'irrégularités de la courbure des contours binaires fins est finalement appliqué sur les éléments restants, permettant de séparer les contours, au niveau des "irrégularités de courbure" (marqueurs rouges, Fig. 1, centre), en un ou plusieurs arcs à courbure locale quasi-constante.

3. Enfin, les cercles associés sont calculés et réarrangés :

Pour chacun des arcs fins précédemment extraits, un unique cercle est calculé par minimisation d'une fonction d'erreur, définie par la distance moyenne entre les pixels de l'arc considéré et le bord du cercle associé (Fig. 1, centre-droite). Connaissant les arcs liés à chacun des cercles, des opérations d'association des cercles sont alors effectuées. D'abord, les cercles considérés comme proches dans l'espace 3D $\langle x_{\text{centre}}, y_{\text{centre}}, r \rangle$ sont fusionnés. Puis, ceux dont les arcs sont confondus avec le bord d'un autre cercle sont fusionnés avec ce dernier. Enfin, les cercles dont l'intérieur, localement au niveau de leurs arcs, est plus sombre que l'extérieur, sont supprimés. Les cercles restants forment alors l'ensemble final des cercles détectés (Fig. 1, droite).

3 Validation de la méthode

3.1 Modèle de simulation de grains

L'objectif du développement d'un modèle de simulation stochastique est de générer des images réalistiques de grains condensés aléatoires, pour lesquelles la vérité terrain est connue. L'application des deux méthodes de segmentation - la CAM et ImageJ - sur des images simulées, permet alors de valider la méthode CAM en comparant la précision de sa PSD à celle d'ImageJ, par rapport à celle de la vérité terrain.

Après avoir initialisé une image vide, le modèle génère les grains et les ajoute un par un de la manière suivante :

1. Une forme binaire aléatoire est d'abord générée (Fig. 2, gauche) : considérant un réel r issu d'une loi de probabilité donnée, un disque de rayon r est créé, dont les contours sont déformés aléatoirement.
2. Des effets de lumière sont ensuite ajoutés (Fig. 2, centre-gauche) : avec une direction globale de la lumière environnante donnée, une ombre propre est générée et ajoutée à la forme binaire selon la direction considérée.
3. Une ombre portée par le grain est enfin créée en considérant la forme binaire et la direction globale de la lumière (Fig. 2, centre). Le grain et son ombre sont alors ajoutés sur l'image à une position suivant une loi uniforme sur les maxima de la carte des distances [3] du fond.

L'image est bruitée et floutée par couches (Fig. 2, droite).

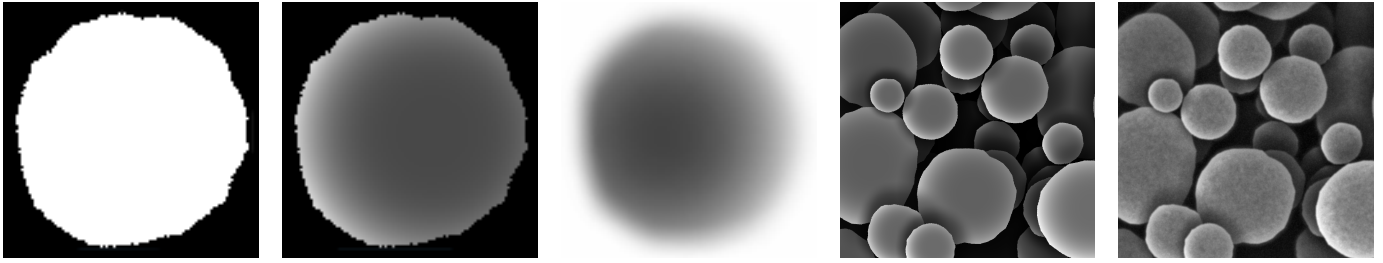


FIGURE 2 : Processus de génération aléatoire de grains. De gauche à droite : Forme binaire aléatoire du grain (gauche), ajout d'effets de lumière à tons de gris sur le grain (centre-gauche), ombre générée par le grain et dirigée (centre), exemple d'image aléatoire formée à partir d'un ensemble de grains (centre-droite), et la même image bruitée et floutée par couches (droite).

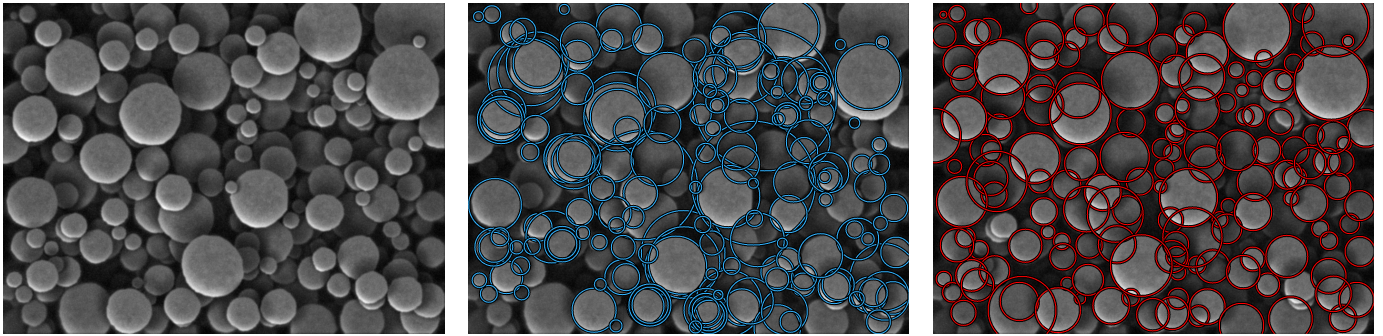


FIGURE 3 : Comparaison visuelle des résultats des deux méthodes de segmentation sur un exemple d'image synthétisée à l'aide du modèle de simulation aléatoire développé. De gauche à droite : Image aléatoire simulée (gauche), segmentation de l'image par la méthode basée sur la CHT d'ImageJ Particle Sizer (centre), et segmentation par la méthode CAM proposée (droite).

3.2 Comparaison de la méthode avec ImageJ sur des images simulées

Afin de comparer l'efficacité de la méthode CAM à celle fournie par ImageJ Particle Sizer, un total de 100 images aléatoires ont été générées à travers le modèle, sur lesquelles les grains simulés ont un rayon suivant une loi log-normale.

Les deux méthodes de segmentation, la CAM et ImageJ, sont alors appliquées sur l'ensemble des images simulées. Les distributions respectives des tailles des particules (PSD) issues de ces deux méthodes sont construites à partir des rayons de l'ensemble des cercles détectés, et sont ensuite comparées à la distribution de la vérité terrain (ground truth). Les résultats sont illustrés dans la figure suivante (Fig. 4).

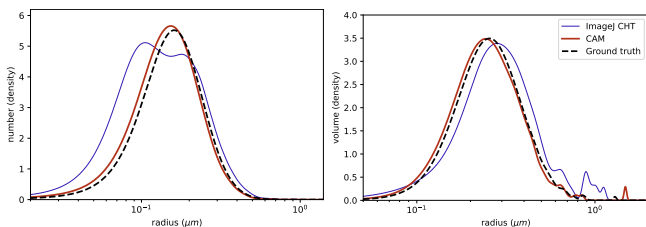


FIGURE 4 : Graphiques des distributions des tailles des particules sur des images simulées avec une loi log-normale, issus des résultats de segmentation avec la CAM et ImageJ, et comparés à la vérité terrain (ground truth). Distributions en nombre (gauche) et en volume (droite).

TABLE 1 : Moyenne et écart-type des distributions sur images simulées suivant une loi log-normale.

Propriétés	Vérité terrain	ImageJ	CAM
Moyenne	0.199 μm	0.179 μm	0.190 μm
Ecart-type	0.082 μm	0.103 μm	0.081 μm

Les courbes de densité illustrées Figure 4 permettent, d'une part, de mettre en avant les nombreux faux petits cercles détectés par ImageJ sur le graphique en nombre, ainsi que la proportion volumique non-négligeable de faux grands cercles mise en avant sur le graphique en volume, faisant ainsi éloigner la courbe de densité issue d'ImageJ (en bleu) de celle de la vérité terrain (en noir). Là où, d'autre part, la CAM (en rouge) semble faire peu de fausses détections dans les valeurs extrêmes, et semble fournir une distribution (PSD) bien plus stable et proche de la vérité terrain, avec une forme globale de courbe similaire à cette dernière.

Le tableau 1 permet de confirmer l'étalement de la PSD produite par ImageJ par rapport à la vérité terrain, avec un écart-type de la distribution bien plus important, là où l'écart-type de la CAM est bien plus proche de la vérité. Bien que la moyenne issue de la CAM soit également plus proche de la vérité terrain par rapport à celle d'ImageJ, les deux méthodes semblent sous-estimer le rayon moyen des grains.

Un exemple visuel des résultats d'application des deux méthodes de segmentation sur une image simulée (Fig. 3) permet de constater les nombreuses fausses détections de petits et de grands grains issues d'ImageJ, là où la segmentation faite par la CAM semble bien plus conforme à la vérité.

4 Application aux images réelles

4.1 Résultats d'application de la méthode

L'objectif du développement de la CAM est l'évaluation de la distribution des tailles de particules réelles de maltodextrine synthétisées en laboratoire et observées par Microscopie Electronique à Balayage (MEB). Un ensemble de 20 images de ces particules en niveaux de gris est produit. La CAM et ImageJ sont appliquées aux images, et les PSD produites sont comparées à celle donnée par Diffraction Laser (DL), une méthode de mesure physique de grains en suspension par diffraction de la lumière. Les résultats sont donnés ci-après.

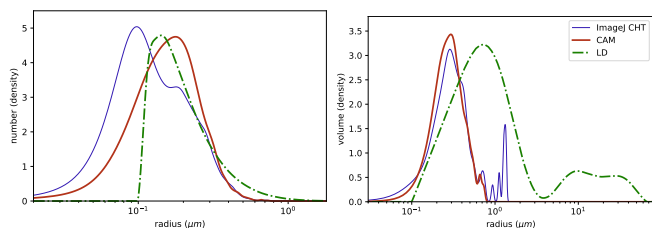


FIGURE 5 : Graphiques des distributions des tailles des particules sur des images réelles, issus des résultats de segmentation avec la CAM, ImageJ et avec diffraction laser (DL). Distributions en nombre (gauche) et en volume (droite).

TABLE 2 : Moyenne et écart-type des distributions sur images réelles.

Propriétés	DL	ImageJ	CAM
Moyenne	0.217 μm	0.191 μm	0.207 μm
Ecart-type	0.133 μm	0.113 μm	0.093 μm

En considérant la CAM comme suffisamment digne de confiance d'après les résultats de validation de la méthode, l'étalement de la courbe de densité issue de la segmentation faite par ImageJ, avec de faux petits et de faux grands cercles détectés en nombre, peut à nouveau être constatée ici sur les graphiques en nombre (Fig. 5, gauche) et en volume (droite) ainsi qu'avec l'écart-type de sa PSD donné dans le tableau 2.

La courbe de densité issue de la DL (en vert, Fig. 5) ainsi que son écart-type (tableau 2) semblent bien éloignés des résultats donnés par la CAM. De plus, la courbe de la DL ne semble pas correspondre à la vérité observable, car de nombreux grains de rayon inférieur à 0.1 μm sont en réalité visibles, et aucun grain de rayon supérieur à 1 μm ne l'est, ce que contredit la courbe de la DL, rendant ainsi cette méthode peu fiable [7].

4.2 Discussion

Bien qu'ImageJ soit robuste sur des images avec des grains bien définis et séparés, il semble défaillant lorsque ces derniers se condensent. Le processus de construction de l'espace de Hough ne tient pas compte des propriétés de courbure et ne fait pas la différence entre les pixels des contours de deux objets différents. Il peut donc facilement détecter de faux cercles entre plusieurs grains s'ils sont suffisamment condensés, ce qui pourrait expliquer la proportion élevée de faux grains.

La CAM, quant à elle, n'analyse que certaines parties (arcs) des contours des grains, et comme ces derniers sont loin d'être parfaitement circulaires, le cercle détecté dépend entièrement de la courbure de l'arc : s'il a un trop petit rayon de courbure, le cercle associé sera plus petit que le grain correspondant, et, au contraire, si l'arc est presque plat, le cercle associé sera plus grand. De plus, lorsque l'arc n'est qu'un petit morceau des contours du grain ou lorsque sa courbure varie beaucoup, le calcul du cercle associé a tendance à donner un rayon plus petit que la vérité terrain. Voir Fig. 3. Cela pourrait expliquer la sous-estimation de la moyenne de la PSD calculée.

5 Conclusion

Une densité élevée de particules dans une image rend son analyse granulométrique difficile. Le logiciel ImageJ Particle Sizer, dont la méthode de segmentation est basée sur la transformée d'Hough circulaire, crée dans de telles images une forte sur-segmentation, éloignant la densité calculée de la réalité.

Une nouvelle méthode de segmentation, la Méthode d'Analyse de Courbure (CAM), a été développée pour répondre aux problèmes soulevés par ImageJ. Un modèle stochastique de simulation de grains a également été créé pour évaluer et comparer la précision de la CAM avec celle d'ImageJ.

L'analyse de la distribution des particules obtenue à partir des deux méthodes de segmentation sur les images simulées montre que la courbe de densité de la CAM est plus juste que celle obtenue par ImageJ, avec moins de fausses détections et une meilleure estimation du rayon moyen. Les mêmes résultats peuvent être observés sur des images réelles.

Les méthodes étudiées ici se basent uniquement sur une analyse morphologique et topologique des images. Des méthodes d'apprentissage profond, telles que les CNN ou les U-Nets, pourraient être utilisées comme outils de segmentation pour donner des résultats bien plus efficaces à ce problème d'évaluation de la distribution des tailles de particules.

Références

- [1] Tomasz BORUCZKOWSKI, Hanna BORUCZKOWSKA, Wioletta DROZDZ, Marta MISZCZAK et Waław LESZCZYŃSKI : Use of imagej software for assessment of mechanical damage to starch granules. *Processes*, 2022.
- [2] Giovanni Stefani Faé et AL. : Making soil particle size analysis by laser diffraction compatible with standard soil texture determination methods. *Soil Science Society of America Journal*, 2019.
- [3] Ricardo FABBRI, Luciano da FONTOURA COSTA, Julio C. TORELLI et Odemir Martinez BRUNO : 2d euclidean distance transform algorithms : A comparative survey. *ACM Computing Surveys*, 2008.
- [4] J. Komala LAKSHMI, M. PUNITHAVALLI et M. PUNITHAVALLI : A survey on skeletons in digital image processing. *International Conference on Digital Image Processing*, 2009.
- [5] Mohsen MIRZAEI et Hossein Khodabakhshi RAFSANJANI : An automatic algorithm for determination of the nanoparticles from tem images using circular hough transform. *Micron*, 2017.
- [6] Kristina P. SINAGA et Miin-Shen YANG : Unsupervised k-means clustering algorithm. *IEEE Access*, 2020.
- [7] György VARGA, Fruzsina GRESINA, Gábor ÚJVÁRI, János KOVÁCS et Zoltán SZALAI : On the reliability and comparability of laser diffraction grain size measurements of paleosols in loess records. *Sedimentary Geology*, 2019.
- [8] Renliang YUAN, Jiong ZHANG et Jian-Min ZUO : Lattice strain mapping using circular hough transform for electron diffraction disk detection. *Ultramicroscopy*, 2019.