

Characterizing powder materials using keypoint-based computer vision methods



Brian L. DeCost, Elizabeth A. Holm*

Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:

Received 27 May 2016

Received in revised form 23 August 2016

Accepted 25 August 2016

Available online 26 October 2016

Keywords:

Powder materials

Computer vision

Machine learning

Image data

Feature extraction

ABSTRACT

We applied the bag of visual words model for visual texture to a dataset of realistic powder micrograph images drawn from eight closely related particle size distributions. We found that image texture based powder classification performance saturates at $89 \pm 3\%$ with 640 training images (80 images per class). This classification accuracy is comparable to classification using conventional segmentation-based particle size analysis. Furthermore, we found that particle size distributions obtained via watershed segmentation are generally not statistically equivalent to the ground truth particle size distributions, as quantified by the two-sample Kolmogorov-Smirnov test for distribution equivalence. We expect image texture classification methods to outperform particle size analysis for more challenging real-world powder classification tasks by capturing additional information about particle morphology and surface textures, which add complexity to the image segmentation task inherent in particle size distribution estimation.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The discipline of materials science has long history of quantifying microstructural images to interpret structure–properties–processing relationships. However, many of the microstructures found in technologically important materials systems are complex and do not map well onto a reductionist segment-and-measure paradigm, at least under the constraints of our current understanding of mesoscale physics. Where explicit mesoscale materials models are not available, there has been interest in applying more flexible image representations to microstructure data in an attempt to circumvent these limitations [1–5]. For example, we recently demonstrated that the ‘bag of visual words’ (BOVW) image texture representation can be used to automatically differentiate between qualitatively different microstructures [5].

As materials scientists we are ultimately interested in learning quantitative structure → properties and processing → structure mappings. We hope that image texture characterization methods can aid us in this goal. A critical outstanding question in this approach is the problem of physical size: can scale-invariant texture features (like the BOVW approach) support quantitative structure → properties mappings that depend strongly on relative

feature sizes? In this work, we move towards quantitative microstructure characterization by focusing on microstructures that vary primarily in their physical dimensions. Here we focus on a simple model system: powder materials with different particle size distributions (PSD).

The size distribution and morphology of powder materials play important roles in powderbed fusion additive manufacturing (AM) [6–8] and other powder metallurgy applications [9]. This class of AM techniques enables rapid building of parts with complex geometries in a layer-by-layer process where a powder feedstock material is locally fused using a computer-routed laser or electron beam [10,11]. The PSD and morphology can play a large role in powder rheology [6,8], and can affect the density distribution in the powder bed, which in turn can alter energy absorption and thermal conductivity [12,13]. Together these factors influence both the final microstructure of the built part, and its properties such as porosity, hardness, mechanical strength, and surface roughness [14,15].

Understanding how powder character affects the build process and final properties is especially important when recycling the unused portion of the powder bed, because the powder character has been shown to change with recycling [14–16]. Unused powder can contain distorted partially melted particles, sintered particle clusters, or solidified melt-pool ejecta [14,15]; recycled powders are typically sieved to remove these, and to ensure that the recycled powder has a maximum particle size consistent with

* Corresponding author.

E-mail addresses: bdecost@andrew.cmu.edu (B.L. DeCost), eaholm@andrew.cmu.edu (E.A. Holm).

the powder bed thickness. Recent powder rheometry studies indicate that sieving may not always mitigate the adverse effects of recycling; powder flowability may decrease [15,17] or increase [14,16] after recycling. For these reasons, it is important to characterize the morphology of AM powder feedstock material, and to couple this with an understanding of how powder morphology influences the build process [18,15,17].

A wide array of characterization techniques has been applied to AM powder feedstock materials [18], including laser PSD measurements [19], X-ray computed tomography (XCT) to measure three dimensional particle sizes and shapes, X-ray diffraction for phase identification, scanning electron microscopy to qualitatively study surface structure, and energy-dispersive spectroscopy (EDS) and X-ray photoelectron spectroscopy (XPS) to measure chemical composition. Importantly, the PSD alone does not provide information about powder morphology that may be important for understanding the behavior of the powder [15]. Consider the gas-atomized powders shown in Fig. 1. Some of the powder particles are non-spherical and appear to have non-trivial surface roughness. Furthermore, many of the particles obscure and occlude each other, and smaller particles tend to agglomerate. These morphological features may not be adequately resolved by laser particle size distribution measurements [18] or region-based image analysis techniques [17], and it may be impractical and/or cost-prohibitive to conduct XCT and powder rheometry experiments regularly in a production setting. Therefore it is of interest to develop powder characterization techniques based on image texture analysis to complement the suite of analytical tools currently in use.

In this study, we investigate the capability of image texture methods, specifically the BOVW model, to capture quantitative differences in microstructure, where the primary aspect of interest is the relative size of microstructure features. Using 3D rendering to create realistic powder micrographs, we generated a dataset of 2048 synthetic micrographs representing eight different powder PSDs. This dataset provides a realistic materials science application, which we use to exercise and evaluate the BOVW representation for microstructure characterization and to delineate the amount of microstructure information required to support a data-driven powder characterization approach. We also evaluate particle size measurements obtained through a conventional segmentation-based method by comparing to the ground truth particle sizes in the rendered images. We conclude that image texture methods, together with rheological and build-performance studies, offer a promising approach to evaluate and qualify powder feedstock for AM processes.

2. Methods

2.1. Synthetic powder micrographs

We created a dataset of realistic powder micrographs; Fig. 2 shows example renderings. This microstructure synthesis task

was performed using Blender [20], an open source computer graphics suite used for 3D modeling, rendering, animation, and scientific visualization; the scripts and texture resources used to generate this dataset are included in a data-in-brief summary [21]. In the present study, we limit ourselves to powders consisting of spherical particles with spatially uncorrelated positions. We consider eight powders differing only in their PSD, which are shown in Fig. 3. For each generating distribution we rendered 256 images, yielding a total dataset of 2048 synthetic powder micrographs. Fig. 2 shows one randomly selected example rendering for each of these eight particle size distributions. We used half of the dataset (1024 images; 128 images per class) for model training and validation, and held out the other half of the dataset as an independent testing set.

The PSDs in Fig. 3 consist of pairs of size distributions that are relatively similar to each other, as detailed in Table 1. This represents a challenging classification problem. Distributions *a* and *b* are lognormal distributions with the same mean particle size of $\mu_a = \mu_b = 0.1$, and shape parameters of $\sigma_a = 0.5$ and $\sigma_b = 0.6$, measured in arbitrary Blender units. Distributions *c* and *d* are a normal distribution and a lognormal distribution, each with a mean particle size of $\mu_c = \mu_d = 0.1$. These are the cleanest powders in the dataset, differing primarily in their upper and lower tails, with small shape parameters of $\sigma_c = 0.025$ and $\sigma_d = 0.25$. Distributions *e* and *f* are lognormal distributions with the same mean particle size of $\mu_e = \mu_f = 0.12$, and the same shape parameters as distributions *a* and *b* ($\sigma_e = \sigma_a = 0.5$ and $\sigma_f = \sigma_b = 0.6$). Distributions *g* and *h* are weibull distributions that have been parameterized to fit data sampled from the lognormal distributions *e* and *f*, respectively.

Each $11 \times 11 \times 2$ (arbitrary Blender units) render volume contains 800 particles with radii sampled from one of the eight generating distributions. The imaging plane is normal to the shorter (*z*) axis of the render volume, and intersects the centroid of the render volume. The camera is located 10 Blender units away from the imaging plane along the *z* direction, and images are rendered at a resolution of 512×512 pixels. Particle positions are randomly sampled from a uniform distribution over the render volume. Because of this, the spherical particles often occlude and/or intersect each other. Each particle is modeled by a spherical mesh, which we imbue with realistic surface texture. By wrapping an image of macroscopic zinc grains found in the passivating surface of galvanized iron. We considered several procedurally-generated image textures, but found that mapping an image of macroscopic zinc grains found in the passivating surface of galvanized iron onto the surface of each particle qualitatively yields more realistic results, as the actual metal powders are polycrystalline.

2.2. Bag of visual words image representation

We characterize the synthetic powder micrographs using a bag of visual words (BOVW) image representation [5], which is an

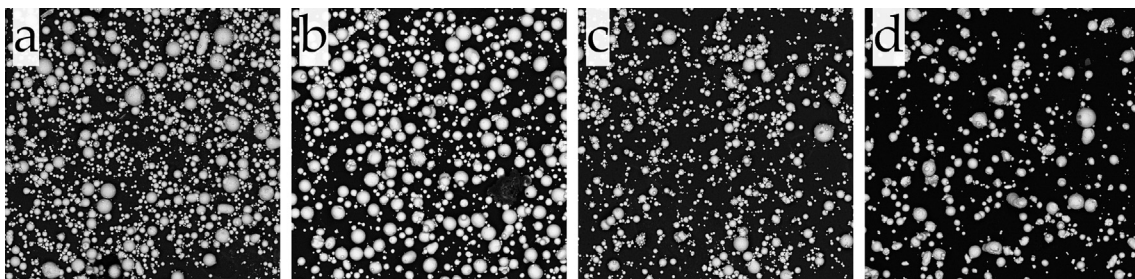


Fig. 1. Scanning electron micrographs of gas atomized powders of a: Al, b: Ti-6Al4V, c: Inconel-718, and d: a maraging steel. Courtesy of A.D. Rollett, R. Cunningham, H. Jain.

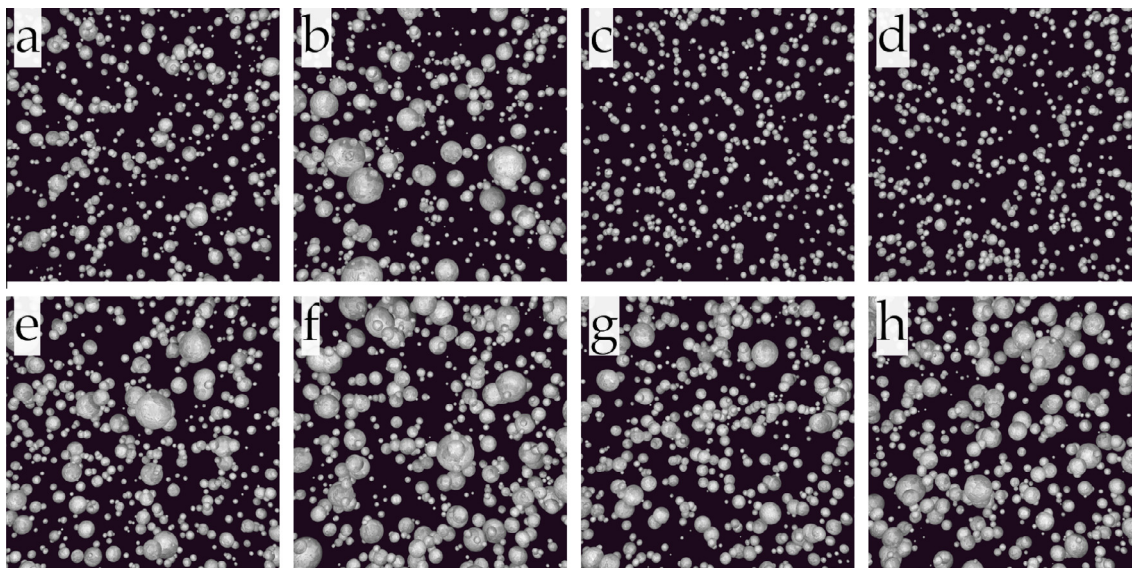


Fig. 2. a–h: Synthetic powder micrographs with particles sizes sampled from each of the eight generating distributions in Fig. 3.

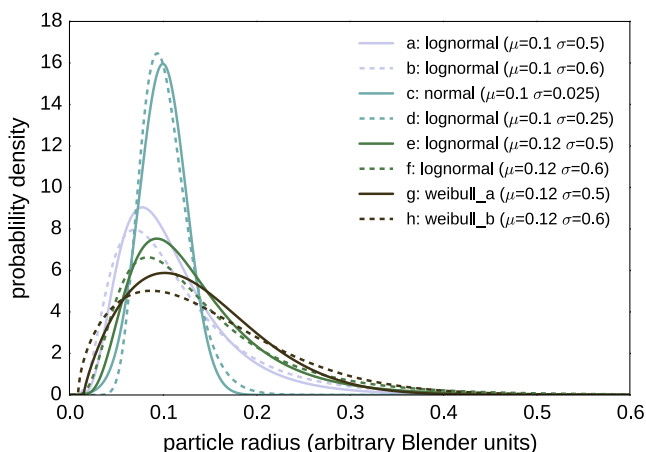


Fig. 3. The eight generating particle size distribution functions, measured in arbitrary Blender units.

Table 1

Distribution parameters for the eight generating particle size distributions. Distributions *g* and *h* are weibull distributions fit to samples from the lognormal distributions *e* and *f*; the parameters listed are for the lognormal distributions.

Class	Family	Mean (μ)	Shape (σ)
<i>a</i>	Lognormal	0.1	0.5
<i>b</i>	Lognormal	0.1	0.6
<i>c</i>	Normal	0.1	0.025
<i>d</i>	Lognormal	0.1	0.25
<i>e</i>	Lognormal	0.12	0.5
<i>f</i>	Lognormal	0.12	0.6
<i>g</i>	Weibull	0.12	0.5
<i>h</i>	Weibull	0.12	0.6

established approach in computer vision for object and texture recognition tasks [22–24]. In this model, an image is represented by a (discrete) distribution of local image pattern descriptors, often called keypoint features. Commonly used pattern descriptors include image patches [25], histograms of oriented image gradient values [26–29], and binary descriptors of spatial patterns of image intensity values (BRIEF [30], ORB [31], BRISK [32], etc). Each carry trade-offs between descriptive power, invariance to image trans-

formations, and computational expense. In the BOVW model, these keypoint features are organized into a visual dictionary (via k-means clustering [33,34], for example). The resulting BOVW representation is an occurrence frequency histogram obtained by assigning each keypoint feature in an image to the nearest cluster center (i.e. visual word) in the visual dictionary [23,24].

In this work, we select distinctive keypoint features across a range of sizes by applying the Difference of Gaussians (DoG) [27] and Harris-LaPlace (HL) [35] interest point detectors, which localize blob-like and corner-like image features, respectively. We characterize the regions surrounding these keypoint features using the Scale Invariant Feature Transform (SIFT) [27,28] computed using the VLFeat library [36], which computes a spatially resolved histogram of oriented image gradient values. Since these SIFT descriptors are computed in a reference frame that is normalized with respect to both the scale and the dominant image gradient orientation of the keypoint feature, the resulting image texture representation is robust to changes in feature scale and orientation [27,28]. We use k-means clustering [33,34] on a subset of SIFT descriptors extracted from the training set of images to compile a dictionary of *k* visual words, where the size *k* of the dictionary is chosen based on the distribution of keypoint features in the training set [37,24]. For each image, we obtain SIFT descriptors at each of the DoG and HL interest points, map each SIFT descriptor to the closest visual word in the dictionary extracted from the training set, and construct a normalized occurrence frequency histogram over the visual words [23,24]. This BOVW representation is a generic image fingerprint that is appropriate for any image drawn from the same distribution as the training set. Using the χ^2 metric [38] $D(X, Y) = \frac{1}{2} \sum_{i=1}^m \frac{(x_i - y_i)^2}{x_i + y_i}$ for comparing discrete *m*-dimensional probability distributions *X* and *Y*, we can quantitatively study relationships between different microstructures using e.g. Support Vector Machine classification (SVM) [39,40], for which we use the Scikit-Learn machine learning library [41].

2.3. Watershed segmentation

To compare the machine vision approach to conventional segment-and-measure characterization, we use ImageJ [42], an open source image analysis toolkit, to segment the synthetic powder micrographs with the watershed [43] algorithm. This scheme

first uses a simple threshold operation to distinguish the particles from the background of the image. The resulting images tend to contain large contiguous regions corresponding to multiple particles, due to agglomeration and occlusion of particles in the original image. The watershed operation partially mitigates this effect. The threshold image is used to compute a distance map from each pixel to the background region. The negated distance map is analogous to a surface with many basins. The watershed algorithm finds new region boundaries corresponding to the watershed lines on this surface by flooding the surface from the local minima.

We measure particle size distributions for each image by computing circle-equivalent radii for each of the segmented regions obtained via watershed segmentation. We constructed particle size histograms for each image using 25 bins of equal width spanning the range of particle sizes from 0.0 to 0.6 (arbitrary Blender units). To directly compare the watershed results to the BOVW results, we use these empirical particle size histograms as input to the same χ^2 SVM classification approach applied to the BOVW image representation. We also quantitatively evaluate the accuracy of the measured size distributions with the generating distributions using the Kolmogorov-Smirnov (KS) test [44]. The KS test is a simple non-parametric statistical hypothesis test based on the maximum difference between the measured and theoretical cumulative particle size distributions. The null hypothesis of the KS test states that the empirical distribution is equal to the reference distribution it is being compared with.

3. Results

3.1. Classification results

3.1.1. Bag of visual words

To quantitatively evaluate how well the BOVW representation can characterize powder micrographs, we performed χ^2 -kernel SVM classification [39,40]. We use half of our dataset (128 images from each of the eight generating distributions) to tune model parameters (the number of visual words in the dictionary k and the SVM regularization parameter C) via 5×8 -fold cross-validation, randomly partitioning the training set into 7 folds for training and one fold for testing, repeated five times. We chose a dictionary size of $k = 100$ using the commonly used ‘elbow’ heuristic [45,46], which expresses the trade-off between the number of clusters (visual words) and the compactness of the clusters. Cluster compactness is quantified by the sum of squared distances between each data point x and the nearest cluster center c in the set of k clusters C . In other words, we perform k-means clustering on the training set of keypoint descriptors for dictionary sizes, and graphically analyze (as in Fig. 4a) the objective function k-means objective function $\sum_{x \in X} \min_{c \in C} (\|x - c\|^2)$. An appropriate number of clusters occurs at the ‘elbow’ where the objective function begins to converge with respect to the number of clusters. For the BOVW application, we are not overly concerned with more sophisticated clustering techniques, as this step primarily serves to partition the keypoint descriptors into discrete categories [24].

We set the SVM regularization parameter $C = 4$ by performing cross-validation for different values of C in the range $\{2^{-5}, 2^{-4}, \dots, 2^5\}$, and selecting C so as to maximize the classification performance for the validation sets while ensuring that the classification performance for the training sets is not significantly higher, which would indicate overfitting. The cross-validation results shown in Fig. 4b indicate that validation accuracy increases slightly as the amount of regularization is decreased (i.e. as C increases), up until around our choice of $C = 4$ (indicated by the dashed line), where the training performance begins to diverge from the validation performance. For $C = 4$, the cross-validation

accuracy was 0.894 ± 0.022 and the training accuracy was 0.919 ± 0.004 .

After selecting appropriate model parameters, we re-trained the SVM classifier using the full training set, and measured the classification performance on the independent testing set. We obtained an overall classification accuracy of 0.889 on these 128 images per class (1024 total). The confusion matrix in Fig. 6a shows a detailed summary of these classification results. The SVM classifier shows the highest selectivity for distributions a and b (about 98% correctly classified), followed by distributions c and d (about 90% correctly classified). There is a higher misclassification rate between the two pairs of closely-related distributions (e, g) and (f, h), where each pair of distributions consists of a lognormal distribution and a weibull distribution with parameters selected to fit the corresponding lognormal distribution. However, despite the similarity between these PSDs, between 80% and 85% are correctly classified using the BOVW model. Overall, these results demonstrate that the image texture approach is sensitive to small quantitative differences between images and can attain high classification accuracy even for very similar microstructural families.

3.1.2. Segment-and-measure results

In order to compare the BOVW image texture method to more traditional microstructural quantification, we used the particle size histograms obtained via watershed segmentation to classify the same set of synthetic powder micrographs. We again used SVM classification with the χ^2 kernel, selecting the SVM regularization parameter C by performing cross-validation for different values of C , for C in $\{2^{-10}, 2^{-9}, \dots, 2^3\}$. For $C = 0.01$, cross-validation resulted in a training set accuracy of $0.949 \pm 0.003\%$, and a validation set accuracy of $0.912 \pm 0.014\%$. Retraining the classifier on the full training set with $C = 0.01$ yielded a training accuracy of 0.944 and a testing set accuracy of 0.902, almost the same as for the image texture method.

The confusion matrix in Fig. 6b shows the classification results in detail. Compared with the BOVW results, the particle size classifier is not as symmetric about the diagonal elements. Instead, it is biased towards particular distributions: b and f , the two lognormal distributions with shape parameter 0.6. Notably, it has somewhat less difficulty discriminating between the two strongly peaked distributions c and d , but this may be related to the bias towards distribution f .

The χ^2 kernel particle size classification results demonstrate that watershed segmentation is sufficient to build a discriminative classifier distinguishing between the eight PSDs in this study with accuracy comparable the BOVW method. However, these results do not necessarily imply that the watershed segmentation process accurately recovers the true PSD. Given the level of particle agglomeration and occlusion in both experimental and synthetic powder micrographs, an image-based approach to quantitatively measuring the PSD may be much more challenging than discriminating between different distributions. Fig. 5a compares the cumulative PSD obtained from the generating distribution (dashed curve), the sample ground truth particle sizes (yellow curve), and the empirical watershed measurement (blue curve) for a single synthetic micrograph drawn from distribution a . While the watershed method should be exact for images of isolated particles (i.e. low particle density), here the watershed PSD tends to shift toward larger particle sizes near the center of the distribution, with less deviation near the tails. This is most likely due to failure of the watershed segmentation to separate some of the overlapping and partially-occluded particles. Fig. 5b shows the distribution of particle circularities $C = 4\pi A/P^2$ for the same synthetic micrograph. Though the mode of this distribution is 1 as expected for circular regions, there is significant probability mass in the lower tail of

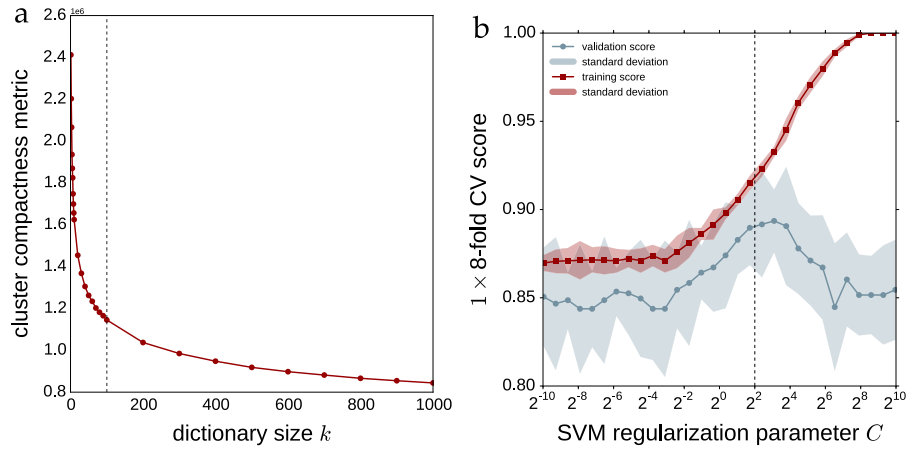


Fig. 4. a: Selecting the visual dictionary size $k = 100$ (indicated by the dashed line) via the graphical ‘elbow’ method. b: Selecting the SVM regularization parameter C (dashed line) based on the divergence of the SVM training (upper red curve) and validation (lower blue curve) accuracies. Confidence bands indicate sample standard deviations over cross-validation sets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

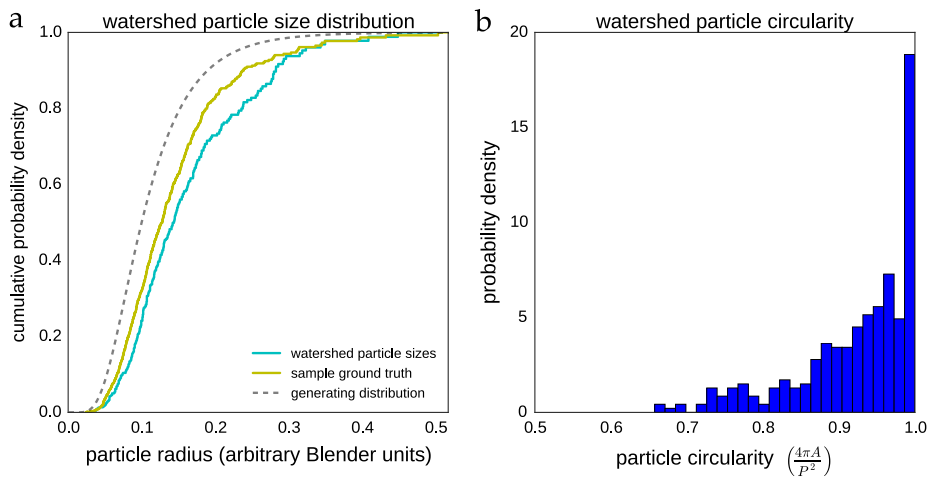


Fig. 5. a: Cumulative distributions for generating, sample ground truth, and empirical watershed PSD for a single synthetic powder micrograph drawn from distribution a. b: Histogram of the circularity for each watershed-segmented region in the image from a.

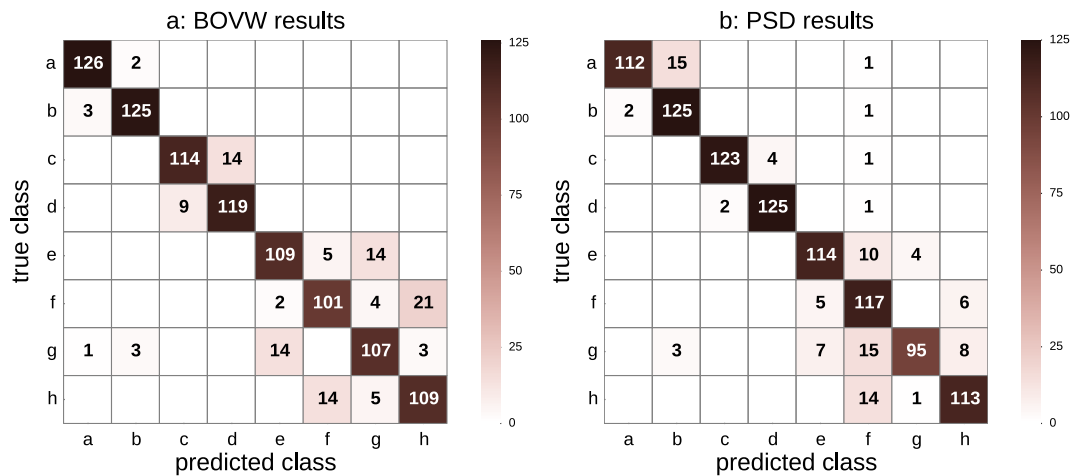


Fig. 6. a: Detailed classification results for bag of SIFT SVM classification (89.9% overall testing set accuracy). b: Detailed classification results for watershed particle size distributions (90.2% overall testing set accuracy). The vertical axis of each confusion matrix indicates the true class, and the horizontal axis indicates the class predicted by the SVM classifier.

the distribution, indicating that there are a substantial number of regions with more complex shapes, which can decrease the quality of the PSD estimate obtained via watershed segmentation.

We quantitatively compare the empirical particle size histograms obtained through watershed segmentation with the continuous ground truth PSDs using the Kolmogorov-Smirnov (KS) test [44]. We reject the null hypothesis (at the 5% significance level) that the empirical distribution and the ground truth distribution are equal for the $86 \pm 10\%$ of the 256 images from each generating distribution. In other words, the measured PSDs are statistically different from the ground truth PSDs. This result has important implications for computational modeling of AM powder feedstock. For example, model powders built from experimental powder PSDs determined via image analysis may not accurately represent the intended system.

In contrast, when comparing the ground truth particle size samples with their continuous generating distributions, the KS test (again at the 5% significance level) suggests we reject the null hypothesis for only $4 \pm 1\%$ of the 256 samples. Thus, the particle sizes used to create the synthetic powder images are an accurate representation of the particle size distributions; this establishes that sampling effects are not responsible for the poor recovery of the ground truth PSDs via watershed segmentation and region analysis.

3.2. Learning curve

Machine learning models may require different amounts of data to achieve good performance, depending on the flexibility of the model and the difficulty of the task [37]. Models that are simple or are based on strong (relevant) prior knowledge about the data typically need fewer training examples than richer, more flexible models [37]. To understand the data requirements of the BOVW model, we measured the change in training and testing accuracy as the size of the training set is increased over the range of $N = \{1, 2, \dots, 128\}$ examples per class. For each training set size that was evaluated, 25 random subsets of the original training set were used for training. In all cases, the full independent testing set (128 examples per class) was used for performance evaluation. For convenience, we reused the visual dictionary obtained from the original training set instead of extracting a new visual dictionary for each cross-validation iteration. We also use the SVM regularization parameter $C = 4$ that we selected via cross-validation over the training set.

The learning curve in Fig. 7 shows the improvement in classification accuracy as more images are included in the training set. The red (upper) curve shows the training set accuracy and the blue (lower) curve shows the testing set accuracy; the confidence bands indicate sample standard deviations. With small training sets, the accuracy on the training set is much higher than accuracy on the validation set. For example, training sets with only one or two instances per class yield a model that perfectly fits the training set, but generalizes poorly, indicating significant overfitting to the small training set. As the training set size increases, the average training set accuracy decreases and converges to about 92%. In contrast, the testing set accuracy begins low and increases with more training examples, converging to about 89% accuracy (dashed line) with around 80 training images per class (640 training images total). The classification accuracies converge to values that are consistent with the cross-validation accuracies obtained from the training set during the model parameter optimization procedure, as described above in Section 3.1.1.

4. Discussion

For this synthetic powder micrograph dataset, the BOVW method yields comparable classification results to segmenting

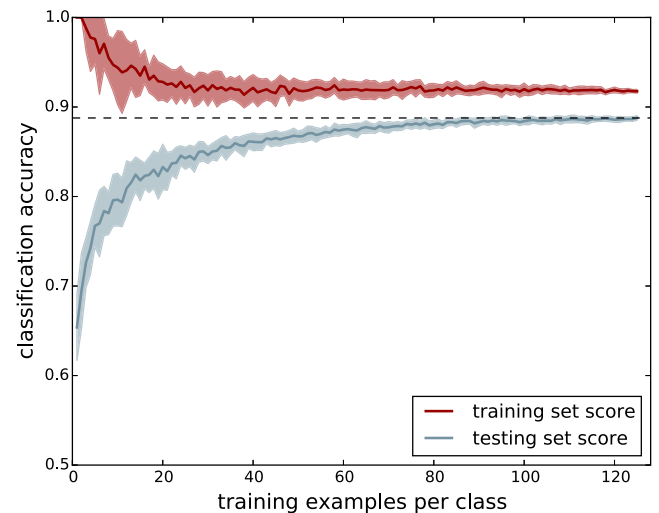


Fig. 7. Classification accuracy vs. number of training examples per class. The red (upper) curve shows the training accuracy and the blue (lower) curve shows the testing set accuracy; the confidence bands indicate sample standard deviations. For each training set size that was evaluated, 25 random subsets of the original training set were used for training, using the independent testing set (128 examples per class) only for performance evaluation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and measuring sphere-equivalent particle size distributions. This demonstrates that the BOVW model is not necessarily limited to applications involving qualitatively different microstructures, but has the potential to provide quantitative microstructural insight. We believe that the BOVW features for this particular synthetic powder dataset effectively capture the shape of the PSD by capturing statistical differences in the tendency for particles of different size to intersect and overlap in the image. Here, this is primarily a function of the shape of the PSD, because the particles in the synthetic powder images are spatially randomly distributed. We note that the BOVW model might not be discriminative for powder micrographs with sparse powder coverage, if the majority of key-point features were individual spherical particles with statistically equivalent surface texture.

Performance on our 8-way powder classification task converged to about $89 \pm 3\%$ with around 80 images per powder class. Generalizing this data requirement to a real-world powder classification task is confounded by several factors, but collecting a comparably sized initial training set of SEM powder micrographs should be well within the capabilities of an industrial or academic group attempting to develop a powder characterization system. Ideally, the collection of a training set of powder micrographs (a minimum of 80 images per class) is an up-front cost; potentially powder and/or machine manufacturers could provide a set of images to support material qualification. Once a training set is established, the data acquisition cost for this system is no greater than that of an image-based PSD characterization approach.

Care must be taken to ensure that sample preparation yields representative and repeatable images. Maintaining consistent SEM imaging conditions will pose a significant challenge, as changes in image contrast can affect resolution of topographical features on the surface of the powder particles. In practice, it is difficult to prepare sparsely dispersed powder samples with consistent spatial distribution of particles, even when using the same material source; the spatial distribution of particles can vary drastically across a single physical sample. For optimal image classification, it may be necessary and/or desirable to select regions of the sample on the basis of particle number density as is done presently. Alternatively, it may be more effective to characterize den-

sely dispersed powder samples using image texture methods, as this may eliminate some of the sample preparation effects on the spatial distribution of particles. One drawback of characterizing dense powder samples is that this could potentially make powder defects such as satellite particles difficult to resolve.

When comparing powder from different sources, it may be necessary to select the magnification to keep the average-sized particle consistent. This approach may help to capture the relative (but not absolute) shape of the PSD, morphological features, and to some extent the local spatial distribution of powder particles. It may not be possible to use a scale-invariant image texture representation to fully support inference about powder properties that depend heavily on the absolute physical scale of the particles. One potential approach to address this weakness is to include information about the physical scale of individual keypoint features. For example, a more robust microstructure descriptor might be obtained by appending the size (in microns) of each keypoint descriptor to the corresponding SIFT descriptor before performing the visual dictionary extraction and BOVW construction steps, similar to the location-augmented SIFT descriptors used by [47,48].

In real powder feedstock, common ‘defects’, including necked particles, broken or fragmented particles, surface roughness, and agglomerated clusters of fine particles can potentially influence the rheological properties of the powder, and thus overall build quality [14,18,15]. A realistic powder system encompasses particle size, shape, and spatial distributions; particle surface texture; and geometric defects such as partially fused, necked, and satellite particles. Rendering powder micrographs with these features may be challenging, so synthetic datasets will not likely serve as a useful training set for application of image texture methods to real powder materials. It is worth noting that this same set of features can make other PSD measurement approaches (including high-throughput methods such as laser particle size analysis) unreliable, because they violate the assumption of spherical particles implicit in many of these techniques. By associating these defects with visual words, the BOVW approach (and image texture methods in general) can be used to identify and flag them, and potentially link them to property metadata such as flowability.

In the present work, we use the same surface texture source image for each synthetic powder class. One objective for future research in this area is to consider synthetic powder classes with identical particle size distributions and different surface texture source images in order to decouple the contribution of powder material surface texture and particle size distribution.

Finally, we emphasize that the PSDs obtained through watershed segmentation of our synthetic powder micrographs are not statistically consistent with the ground truth PSDs. This indicates that a PSD obtained via image analysis is not necessarily a more objective representation than an implicit image texture representation. This is especially true for powders with a high degree of particle agglomeration and morphological heterogeneity, for which the implicit assumption of spherical particles does not apply. For powders displaying these features, it may be necessary to develop a computational model of the image formation process in order to infer a precise PSD via image-based methods.

5. Conclusions

We apply the bag of visual words (BOVW) computer vision technique to obtain a rich microstructure representation that can be used to explore process–structure–properties mappings. We find that BOVW classification performance saturates at $89 \pm 3\%$ with 80 synthetic powder micrograph instances for each of eight generating particle size distributions. The BOVW classification accuracy is comparable to conventional segmentation-based parti-

cle size analysis. Though watershed segmentation yields discriminative particle size distributions (PSD) for this synthetic dataset, we demonstrate that they are not statistically equivalent to the ground truth PSDs. The BOVW and other image texture approaches have the additional capability of characterizing important morphological and surface features of the powder particles. Image texture methods offer a promising image-based means of evaluating and qualifying powder feedstock for additive manufacturing (AM) processes, supplementing contemporary powder material characterization techniques. While we focus on powder feedstock for AM, image texture characterization methods are also applicable to other complex microstructure systems, such as those found in AM builds.

Acknowledgements

We gratefully acknowledge funding for this work through the National Science Foundation Grant Nos. DMR-1307138 and DMR-1507830, and through the John and Claire Bertucci Foundation. A.D. Rollett, R. Cunningham, and H. Jain graciously supplied the experimental powder micrographs in Fig. 1. Thanks to VLFeat [36], Scikit-Learn [41], Blender [20], and ImageJ [42] for the awesome open source code. We are also grateful for helpful and inspirational discussions with Prof. Abhinav Gupta and Xinlei Chen of the Carnegie Mellon Institute School of Computer Science.

References

- [1] Surya R. Kalidindi, Stephen R. Niezgod, Ayman A. Salem, Microstructure informatics using higher-order statistics and efficient data-mining protocols, *JOM* 63 (4) (2011) 34–41.
- [2] Stephen R. Niezgod, Anand K. Kanjarla, Surya R. Kalidindi, Novel microstructure quantification framework for databasing, visualization, and analysis of microstructure data, *Integrat. Mater. Manuf. Innovat.* 2 (1) (2013) 1–27.
- [3] Hongyi Xu, Ruoqian Liu, Alok Choudhary, Wei Chen, A machine learning-based design representation method for designing heterogeneous microstructures, *J. Mech. Des.* 137 (5) (2015) 051403-1–051403-10.
- [4] Lily Nguyen, Dong Wang, Yunzhi Wang, Marc De Graef, Quantifying the abnormal strain state in ferroelastic materials: a moment invariant approach, *Acta Mater.* 94 (2015) 172–180.
- [5] Brian L. DeCost, Elizabeth A. Holm, A computer vision approach for automated analysis and classification of microstructural image data, *Comput. Mater. Sci.* 110 (2015) 126–133.
- [6] Bart Van der Schueren, Jean-Pierre Kruth, Powder deposition in selective metal powder sintering, *Rapid Prototyp. J.* 1 (3) (1995) 23–31.
- [7] A. Simchi, The role of particle size on the laser sintering of iron powder, *Metall. Mater. Trans. B* 35 (5) (2004) 937–948.
- [8] John A. Slotwinski, Edward J. Garboczi, Metrology needs for metal additive manufacturing powders, *JOM* 67 (3) (2015) 538–543.
- [9] Randall M. German, *Powder Metallurgy and Particulate Materials Processing: The Processes, Materials, Products, Properties, and Applications*, Metal Powder Industries Federation Princeton, NJ, 2005.
- [10] D.D. Gu, W. Meiners, K. Wissenbach, R. Poprawe, Laser additive manufacturing of metallic components: materials, processes and mechanisms, *Int. Mater. Rev.* 57 (3) (2012) 133–164.
- [11] William E. Frazier, Metal additive manufacturing: a review, *J. Mater. Eng. Perform.* 23 (6) (2014) 1917–1928.
- [12] H.H. Zhu, J.Y.H. Fuh, L. Lu, The influence of powder apparent density on the density in direct laser-sintered metallic parts, *Int. J. Machine Tools Manuf.* 47 (2) (2007) 294–298.
- [13] Mahesh Mani, Brandon Lane, Alkan Donmez, Shaw Feng, Shawn Moylan, Ronnie. Fesperman, Measurement Science Needs for Real-time Control of Additive Manufacturing Powder Bed Fusion Processes, National Institute of Standards and Technology, Gaithersburg, MD, 2015, NIST Interagency/Internal Report (NISTIR), 8036.
- [14] V. Seyda, N. Kaufmann, C. Emmelmann, Investigation of aging processes of ti-6al-4 v powder material in laser melting, *Phys. Proc.* 39 (2012) 425–431.
- [15] Jamie Clayton, Doug Millington-Smith, Brian Armstrong, The application of powder rheology in additive manufacturing, *JOM* 67 (3) (2015) 544–548.
- [16] H.P. Tang, M. Qian, N. Liu, X.Z. Zhang, G.Y. Yang, J. Wang, Effect of powder reuse times on additive manufacturing of ti-6al-4v by selective electron beam melting, *JOM* 67 (3) (2015) 555–563.
- [17] A. Strondl, O. Lyckfeldt, H. Brodin, U. Ackelid, Characterization and control of powder properties for additive manufacturing, *JOM* 67 (3) (2015) 549–554.

- [18] J.A. Slotwinski, E.J. Garboczi, P.E. Stutzman, C.F. Ferraris, S.S. Watson, M.A. Peltz, Characterization of metal powders used for additive manufacturing, *J. Res. Natl. Inst. Stand. Technol.* 119 (2014) 460.
- [19] Vincent A. Hackley, Lin-Sien Lum, Vadas Gintautas, C.F. Ferraris, Particle Size Analysis by Laser Diffraction Spectrometry: Application to Cementitious Powders, US Department of Commerce, Technology Administration, National Institute of Standards and Technology, 2004.
- [20] Roland Hess, *The Essential Blender: Guide to 3D Creation with the Open Source Suite Blender*, No Starch Press, 2007.
- [21] B.L. DeCost, E.A. Holm, A large dataset of synthetic SEM images of powder materials and their ground truth 3D structures, *Data in Brief*, (2016) in press.
- [22] Svetlana Lazebnik, Cordelia Schmid, Jean Ponce, A sparse texture representation using local affine regions, *IEEE Trans. Pattern Anal. Machine Intell.* 27 (8) (2005) 1265–1278.
- [23] Gabriella Csurka, Christopher Dance, Lixin Fan, Jutta Willamowski, Cédric Bray, Visual categorization with bags of keypoints, *Workshop on Statistical Learning in Computer Vision*, vol. 1, ECCV, 2004, pp. 1–2.
- [24] Eric Nowak, Frédéric Jurie, Bill Triggs, Sampling strategies for bag-of-features image classification, in: *Computer Vision–ECCV 2006*, Springer, 2006, pp. 490–503.
- [25] Manik Varma, Andrew Zisserman, A statistical approach to material classification using image patch exemplars, *IEEE Trans. Pattern Anal. Machine Intell.* 31 (11) (2009) 2032–2047.
- [26] Navneet Dalal, Bill Triggs, Histograms of oriented gradients for human detection, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005 (CVPR 2005), vol. 1, IEEE, 2005, pp. 886–893.
- [27] David G. Lowe, Object recognition from local scale-invariant features, *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, 1999, vol. 2, IEEE, 1999, pp. 1150–1157.
- [28] David G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vision* 60 (2) (2004) 91–110.
- [29] Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool, Speeded-up robust features (surf), *Comput. Vision Image Understand.* 110 (3) (2008) 346–359.
- [30] Michael Calonder, Vincent Lepetit, Christoph Strecha, Pascal Fua, Brief: binary robust independent elementary features, in: *Computer Vision–ECCV 2010*, 2010, pp. 778–792.
- [31] Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski, Orb: an efficient alternative to sift or surf, in: *2011 IEEE International Conference on Computer Vision (ICCV)*, IEEE, 2011, pp. 2564–2571.
- [32] Stefan Leutenegger, Margarita Chli, Roland Y. Siegwart, Brisk: binary robust invariant scalable keypoints, in: *2011 IEEE International Conference on Computer Vision (ICCV)*, IEEE, 2011, pp. 2548–2555.
- [33] Stuart Lloyd, Least squares quantization in pcm, *IEEE Trans. Inform. Theory* 28 (2) (1982) 129–137.
- [34] David Arthur, Sergei Vassilvitskii, k-means++: the advantages of careful seeding, in: *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035.
- [35] Krystian Mikolajczyk, Cordelia Schmid, Indexing based on scale invariant interest points, *Proceedings of the Eighth IEEE International Conference on Computer Vision*, 2001 (ICCV 2001), vol. 1, IEEE, 2001, pp. 525–531.
- [36] Andrea Vedaldi, Brian Fulkerson, Vlfeat: an open and portable library of computer vision algorithms, in: *Proceedings of the International Conference on Multimedia*, ACM, 2010, pp. 1469–1472.
- [37] Trevor Hastie, Robert Tibshirani, Jerome Friedman, T. Hastie, J. Friedman, R. Tibshirani, *The Elements of Statistical Learning*, vol. 2, Springer, 2009.
- [38] Jianguo Zhang, Marcin Marszałek, Svetlana Lazebnik, Cordelia Schmid, Local features and kernels for classification of texture and object categories: a comprehensive study, *Int. J. Comput. Vision* 73 (2) (2007) 213–238.
- [39] Corinna Cortes, Vladimir Vapnik, Support-vector networks, *Machine Learn.* 20 (3) (1995) 273–297.
- [40] Bernhard Schölkopf, Alexander J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, 2002.
- [41] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Scikit-learn: machine learning in python, *J. Machine Learn. Res.* 12 (2011) 2825–2830.
- [42] Caroline A. Schneider, Wayne S. Rasband, Kevin W. Eliceiri, et al., Nih image to imagej: 25 years of image analysis, *Nat. Methods* 9 (7) (2012) 671–675.
- [43] Pierre Soille, Luc M. Vincent, Determining watersheds in digital pictures via flooding simulations, in: *Visual Communications and Image Processing*, International Society for Optics and Photonics, 1990, pp. 240–250.
- [44] Frank J. Massey Jr., The kolmogorov-smirnov test for goodness of fit, *J. Am. Stat. Assoc.* 46 (253) (1951) 68–78.
- [45] Robert L. Thorndike, Who belongs in the family?, *Psychometrika* 18 (4) (1953) 267–276.
- [46] Douglas Steinley, K-means clustering: a half-century synthesis, *British J. Math. Stat. Psychol.* 59 (1) (2006) 1–34.
- [47] Jorge Sánchez, Florent Perronnin, Teófilo De Campos, Modeling the spatial layout of images beyond spatial pyramids, *Pattern Recogn. Lett.* 33 (16) (2012) 2216–2223.
- [48] Ken Chatfield, Karen Simonyan, Andrea Vedaldi, Andrew Zisserman, Return of the devil in the details: delving deep into convolutional nets, *arXiv preprint arXiv:1405.3531*, 2014.