Determining wood chip size: image analysis and clustering methods

Paolo Febbi, Corrado Costa, Paolo Menesatti, Luigi Pari

Consiglio per la Ricerca e la sperimentazione in Agricoltura, Unità di ricerca per l'ingegneria agraria, Monterotondo (Roma), Italy

Abstract

One of the standard methods for the determination of the size distribution of wood chips is the oscillating screen method (EN 15149-1:2010). Recent literature demonstrated how image analysis could return highly accurate measure of the dimensions defined for each individual particle, and could promote a new method depending on the geometrical shape to determine the chip size in a more accurate way. A sample of wood chips (8 litres) was sieved through horizontally oscillating sieves, using five different screen hole diameters (3.15, 8, 16, 45, 63 mm); the wood chips were sorted in decreasing size classes and the mass of all fractions was used to determine the size distribution of the particles. Since the chip shape and size influence the sieving results, Wang's theory, which concerns the geometric forms, was considered. A cluster analysis on the shape descriptors (Fourier descriptors) and size descriptors (area, perimeter, Feret diameters, eccentricity) was applied to observe the chips distribution. The UPGMA algorithm was applied on Euclidean distance. The obtained dendrogram shows a group separation according with the original three sieving fractions. A comparison has been made between the traditional sieve and clustering results. This preliminary result shows how the image analysis-based method has a high potential for the characterization of wood chip size distribution and could be further investigated. Moreover, this method could be implemented in an online detection machine for chips size characterization. An improvement of the results is expected by using supervised multivariate methods that utilize known class memberships. The main objective of the future activities will be to shift the analysis from a 2-dimensional method to a 3-dimensional acquisition process.

Introduction

The size distribution of wood chips is recognized as one of the most important parameters for efficient combustion since it affects the storage, the efficiency of energy conversion and environmental emissions (Nati *et al.*, 2010). The dimensions of wood chips are specified by the international standard EN 14961-1:2010 (CSN, 2010) (Table 1).

For fuel chips, mechanically or manually operated screening devices are commonly applied and there is a large variety of applicable systems. The reference methods for size classification of samples are EN 15149-1:2010 (CSN, 2010; oscillating screen method) and CEN/TS 15149-3:2006 (CNS, 2006; rotary screen method). The average relative repeatability limits for horizontal (< 2 w-%) and rotary screenings are exceedingly low. Generally, the relative reproducibility limits based on the median values for horizontal (< 10 w-%) and rotary screening results seem acceptable. On average, reproducibility is better for horizontal screenings than for rotary screening. However, if different measuring principles are applied, particle size analysis of biofuels is associated with high measuring uncertainties (Hartmann *et al.*, 2006).

Since the chip form (shape and size) influences the sieve results, the image analysis could provide a method which is sensitive to the geometrical shape for determining a more accurate measure of size integrated with shape (Fernlund, 1998).

Consiglio per la Ricerca e la sperimentazione in Agricoltura, Unità di ricerca per l'ingegneria agraria (CRA-ING) is conducting a research activity for developing a new method, based on the image analysis, for the classification of wood particles, depending on both size and shape.

Materials and Methods

A sample of wood chips (8 liters) was sieved through horizontally oscillating sieves, using five different screen hole diameters (3.15, 8, 16, 45, 63 mm). The wood chips were sorted in decreasing size classes and the size distribution of wood chips was determined. The result is expressed as a percentage of the total mass of all fractions.

The test sample was divided and a wood chip subsample of 1 litre (706 chips) was processed in sequential sieving operations for the determination of its size distribution. Three fractions have been considered: 8÷16 mm, 16÷45 mm, 45÷63 mm. Digital images of all chips in the subsample have been acquired using a high resolution digital scanner A3 (600 dpi; Figure 1). 2-D numerical data have been processed in Matlab (rel. 7.1) environment, specifically developed by CRA-ING. After image segmentation, the following parameters from each object have been extracted: size descriptors (area, perimeter, Feret diameters, eccentricity, etc.) and shape descriptors (99 Fourier descriptors). The Feret diameter is defined as the distance between two parallel tangential lines restricting the object perpendicular to that direction; it measures a particle size along a specified direction. To determine the length and width of a particle area, the algorithm measures Feret diameters in 32 azimuth directions and takes the longest distance as D_{Max} (*length*) and the shortest distance as D_{min} (*width*). The Fourier coefficients can summarize the shape of an object in the frequency domain. Complex shapes can be represented with a small number of invariant coefficients, which can be

viewed as features extracted from the original shapes. Generally, a subset of the components is often enough to capture the overall features of the shape and discriminate different shapes (Zhang and Lu, 2002).

Wang (Wang, 1994) presented a theory suggesting that geometric forms, rectangles, triangles and diamonds, should fall into different areas of a plot of the shape factor, Sh (the ratio between the particles perimeter and area, P/A), against the ratio between D_{Max}/D_{min}.

In the present work, some ideal geometric shapes have been considered: rectangles, diamonds, squares and triangles. In order to obtain scale invariant contours from forms, each of them has been normalized by the length of $D_{min}=1$; so, the shape classification is size independent. This is represented in Figure 2, where diamonds, rectangles and two different kinds of triangles are reported at ratio D_{Max}/D_{min} varying in the range $\sqrt{2} \div 10$.

Fernlund (Fernlund, 1998) reported a plot of the shape factor (Sh) against the ratio between length and width made for rectangular, triangular and diamond shaped particles. Mathematically the different shapes fall into different areas in this diagram. However, the geometric forms are not easily differentiated using the 32 azimuth directions to compute the Feret diameters. The computed values do not agree with the theoretical ones and so there is a great deal of overlap. The reason for this is assumed to be that the particles are more irregular than they are similar to the typical geometric forms and their perimeters and areas do not agree with the theoretical perimeters and areas (Jansson and Muhr, 1995). It has been demonstrated that in a two dimensional plot it's quite difficult to distinguish enough the different shapes. Wang (Wang, 1994) considered four descriptors (perimeter, area, D_{Max}, D_{min}) to represent forms in a 2-dimensional space; the relative positions of objects in that space cannot be always distinguished. This method reduces the reference space, consequently reducing the distances between objects. Distances between shapes in a 2-dimensional space are smaller than in a n-dimensional space.

To discriminate the real forms of chips, it could be helpful to represent the objects in a multidimensional way, with many axes. This is equivalent to dilating the reference space, but augmenting the complexity of the analyses. The shape of chips can be transformed in similarity coefficients in order to quantify the resemblance between objects. Similarities are higher when the two objects are identical and lower when the two objects are completely different. A similarity index (S) can be transformed into a distance (D) by computing its one-complement. For a similarity measure varying between 0 and 1, the corresponding distance may be computed as (Legendre, 1998).

In an Euclidean space, the objects (chips) may be represented along axes using shape and size descriptors (Costa et al., 2011). The relative position of an object with respect to another one may be used to obtain clusters. Clustering objects is a multivariate operation which partitions the objects into two or more subsets (clusters), using pre-established rules of agglomeration or division, such that each object belongs to one and only one subset of that partition. The subsets form a series of mutually exclusive classes, among which the objects are included.

The adopted clustering method is based on the Euclidean distances among objects, computed using shape and size descriptors (109 variables). The hierarchical clustering routine produces a dendrogram showing how objects can be clustered. A dendrogram is made of branches that meet at nodes which are drawn at the similarity value where fusion of branches takes place.

The dendrogram based on Euclidean distance was performed with the Unweighted Pair-Group Average (UPGMA) algorithm. The highest similarity, or smallest distance, identifies the next cluster to be formed. Following this event, the method computes the arithmetic average of the similarities or distances between a candidate object and each of the cluster members or, in the case of previously formed cluster, between all members of the two clusters. All objects receive equal weights in the computation. Clusters are joined based on the average distance between all members in the two groups. Clustering proceeds by agglomeration as the similarity criterion is relaxed. Each cluster collects objects (chips) that are sufficiently similar, given the variables considered.

Table 1. Specification of the dimensions for wood chips.

Dimensions (mm) for wood chips (UNI EN 14961-1:2010)			
Particle size distribution	Main fraction (min. 75 w-%)	Cross sectional area	Coarse fraction, max length of
P16	3,15÷16 mm	< 1 cm^2	< 31,5/120 mm
P45	8÷45 mm	< 5 cm^2	< 120/350 mm
P63	8÷63 mm	< 10 cm^2	< 350 mm
P100	16÷100 mm	< 18 cm^2	< 350 mm

Results

The result of cluster analysis (UPGMA based on Euclidean distance) on both shape descriptors (Fourier descriptors) and size descriptors (area, perimeter, Feret diameters, eccentricity, etc.) has been reported in Figure 3 to observe the chips distribution according to their descriptors derived from image analysis. For graphical convenience, vertical lines are used to connect branches at the similarity levels of the nodes.

Figure 1. Scanned image of chips (left) and segmented image (right).

Figure 3 represents a summarized version of the whole dendrogram, being it composed by 706 observations. The separation of the main groups is in accordance with the original three sieving fractions. The percentage of chips classified according to the sieving procedure grouping is reported.

Comparing the original assignment of chips made by the sieve method to the assignment made by the clustering method (Table 2), it is possible to determine the number of chips classified in accordance. The method quality is expressed through the percentage of correctly classified units.

Conclusions

The examined data sets consist of chips that belong to three different fractions or groups. A comparison has been made between the tra-

Table 2. Chips counts in the three fractions using sieving and clustering methods.

Sieving	Fraction 8÷16	Fraction 16÷45	Fraction 45÷63	Total
Fraction 8÷16	483	-	-	483
Fraction 16÷45	-	220	-	220
Fraction 45÷63	-	-	3	3

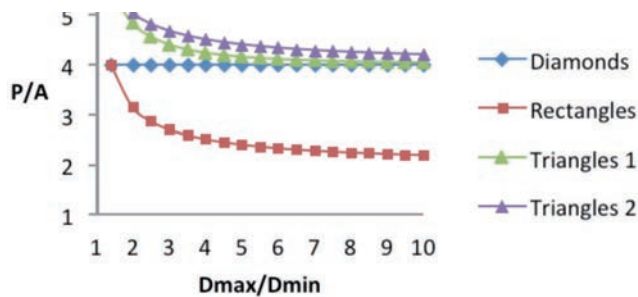| Clustering	Fraction 8÷16	Fraction 16÷45	Fraction 45÷63	Total
Fraction 8÷16	478	36	-	514
Fraction 16÷45	5	184	-	189
Fraction 45÷63	-	-	3	3

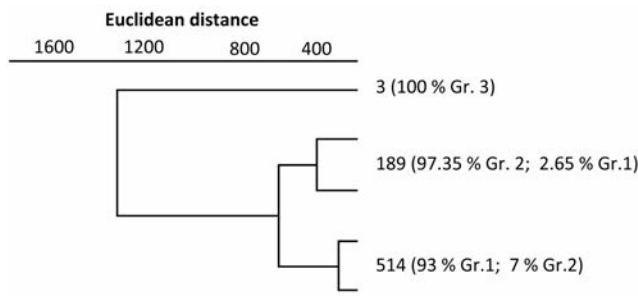**Figure 2. Different areas occupied by the geometric shapes.**

Figure 3: Summarized dendrogram for the cluster analysis on the chips descriptors (size and shape). The groups numbers refers to the sieving procedure grouping

ditional sieving and clustering results. The preliminary result shows that the image analysis-based method has a high potentiality for the characterization of wood chip size distribution and could be further investigated. The ability to observe particle groups in accordance with the fractions determined by the screen methods allows the development of a new method for the determination of chip size distribution compatible with the standards EN 15149-1/3. Moreover, this new method could be implemented in an online detection machine for chips size characterization.

Cluster analysis is an unsupervised method that attempts to identify classes (or groups) without using pre-established class memberships. Other works within different scientific fields demonstrated how the integration of complex shape descriptors (such as Fourier descriptors; Menesatti *et al*., 2009) with basic morphometric descriptors (Feret diameters, area, perimeter, etc.) could return better and more accurate descriptions of a phenomena.

In light of this, an improvement of the results is expected by using supervised multivariate methods that utilize known class memberships. Cumulative curves of the size distribution based on image analysis could be constructed. The main objective of the future activities will be to shift the analysis from a 2-dimensional method to a 3-dimensional acquisition process.

References

Costa, C., Antonucci, F., Pallottino, F., Aguzzi, J., Sun, D.-W., Menesatti, P., 2011. Shape analysis of agricultural products: A review of recent research advances and potential application to computer vision. Food and Bioprocess Technology, 4, 673-692.

CSN P CEN/TS 15149-3:2006. Solid biofuels - Methods for the determination of particle size distribution - Part 3: Rotary screen method.

CSN EN 14961-1:2010. Solid biofuels - Fuel specifications and classes - Part 1: General requirements.

CSN EN 15149-1:2010. Solid biofuels - Determination of particle size distribution - Part 1: oscillating screen method using sieve apertures of 1 mm and above.

Fernlund, J.M.R., 1998. The effect of particle form on sieve analysis: a test by image analysis, Engineering Geology, 50: 111-124.

Hartmann H., Thorsten B., Jensen P.D., Temmerman M., Rabier F., Golser M., 2006. Methods for size classification of wood chips. Biomass and Bioenergy 30: 944-953.

Jansson, M., Muhr, T., 1995. Geometrisk form hos ballast, inverkan på siktresultat och kvalitet. Examensarbete. Avdelningen för Teknisk geologi, Institutionen för Anläggning och miljö, Kungl Tekniska Högskolan.

Legendre P., Legendre L., 1998. Numerical Ecology. 2nd English, ed. Elsevier.

Menesatti P., Aguzzi J., Costa C., García J.A., Sardà F., 2009. A new morphometric implemented video-image analysis protocol for the study of social modulation in activity rhythms of marine organisms. Journal of Neuroscience Methods, ELSEVIER SCIENCE, 184(1): 161-168, ISSN 1872-678X.

Nati C., Spinelli R., Fabbri P., 2010. Wood chips size distribution in relation to blade wear and screen use, Biomass Bioenerg, 34(5): 583-87.

Wang, W.X., 1994. Form analysis of aggregate, for KTH-BASLLAST-standards. June, 1994.

Zhang D. and Lu G., 2002. "A Comparative Study of Fourier Descriptors for Shape Representation and Retrieval," Proc. Fifth Asian Conf. Computer Vision (ACCV '02), pp. 646-651, Jan. 2002.