Mathematical Geology, Vol. 35, No. 7, October 2003 (© 2003)

Multiscale Roughness Analysis of Particles: Application to the Classification of Detrital Sediments[1]

Hervé Drolon,[2] Bernard Hoyez,[3] Fabrice Druaux,[4] and Alain Faure[4]

The shape of detrital quartz grains, mainly acquired under the effect of mechanical transport agents, has been observed and described for a long time by sedimentologists. The roughness of these grains is recognized as an important parameter for analyzing sediments, but to quantify this parameter with a discriminant measure can be difficult. By using the concept of multiresolution analysis and the wavelet formalism, we develop a roughness descriptor for particles, possessing multiscale analysis capabilities. It allows analyses of wear and erosion phenomena acting on particles by decomposing the information on the shape into different resolution levels, each level corresponding to a different scale. Thus it becomes possible to differentiate coarse roughness from finer roughness of a grain's outline. This descriptor is invariant under affine geometric transformation and can therefore be used to compare sands stemming from a wide range of natural environments. An application to the analysis and the classification of detrital sediments, along with a comparison with Fourier and fractal grain shape analysis, illustrates the performance of the method.

KEY WORDS: grain shape analysis, roughness analysis, wavelet transform, classification, neural networks.

INTRODUCTION

The shape of sedimentary particles has for a long time been recognized as being an important parameter in helping to improve the understanding of the sedimentological processes. The shape of quartz grains, for example, reflects the genesis of a given source and are distinct from the shape of quartz grains from other sources (Mazzullo and Magenheimer, 1987). Thus, it gives sedimentologists information about the physical agent which fashioned the grain and about its transport and deposit conditions.

---

[1]Received 4 October 2002; accepted 7 July 2003.

[2]AQL Groupe Silicomp, rue de la Châtaigneraie, BP 127, 35513 Cesson Sévigné Cedex, France; e-mail: herve.drolon@aql.fr

[3]Laboratoire de Mécanique, Université du Havre, Quai Frissard, 76600 Le Havre, France.

[4]GREAH, Université du Havre, IUT, Place R. Schuman, 76610 Le Havre, France.

0882-8121/03/1000-0805/1

The shape of a particle is a difficult characteristic to quantify, because shapes that one meets are often complex, and very different from mathematical objects such as spheres, ellipsoids, and cubes. Nevertheless, numerous methods have been proposed to describe particles outlines (see, e.g., Clark, 1981; Diepenbroek, Bartholomä, and Ibbeken, 1992, for a review). The basic idea consists of characterizing the contour by one or several parameters and to link these parameters to a studied physical property: elongation, angularity, roughness, or roundness, and more generally the degree of wear of the particle. Among existing techniques, the most used are Fourier analysis of the contour (Boon, Evans, and Hennigar, 1982; Diepenbroek, Bartholomä, and Ibbeken, 1992; Ehrlich and Weinberg, 1970; Mazzullo and others, 1992), fractal analysis (Hoyez, 1994; Kaye, 1978; Kennedy and Wei-Hsiung, 1992; Vallejo, 1995), or mathematical morphology (Pirard, 1992, 1994).

The notion of shape depends partly on the scale of observation. For this reason, most authors have attempted to isolate its different constituents which are the global aspect (sphericity, elongation) and details (angularity, roughness, or roundness) by characterizing them by independent coefficients. This decomposition in different levels of the information on the shape suggests that the notion of scale of observation is essential to characterize a particle. This intuitive concept of scale of observation is an important one, which appears frequently in pattern recognition (Cesar and Costa, 1996; Mokhtarian and Mackworth, 1992). For shapes being composed of structures occurring on different scales, it is often important to identify the right scale for the analysis of a physical process. Since the required scales are not known in advance, one solution to this representation problem is to reproduce the contour at multiple scales, so that each structure can be seen at its appropriate scale. This technique, which is related to the field of *multiscale shape analysis*, has been addressed recently by using the wavelet transform (Chuang, Kuo, and Jay, 1996) and has allowed new shape descriptors to be derived, called wavelet descriptors, that provide an efficient method for contour analysis and description.

By using the harmonic wavelet transform (Newland, 1993), we developed a new shape descriptor, *the multiscale roughness descriptor*, that proves to be perfectly adapted to the description of particles (Drolon and others, 1999). This descriptor acts as a mathematical microscope and allows the analyses of the contour of a grain and its roughness at different scales of resolution. Coefficients provided by this descriptor are invariant under translation, rotation, and change of scale of the contour, and can therefore be used to compare sands stemming from a wide range of sedimentary environments.

MULTISCALE ROUGHNESS DESCRIPTOR

Principle

Harmonic wavelet descriptors (HWD) are contour descriptors based on the harmonic wavelet transform (Newland, 1993). They are computed from Fourier descriptors (FD) of the contour (Granlund, 1972), with a fast algorithm which uses only FFT operations. They are invariant under translation, rotation, scaling, and independent of the starting point chosen to trace the contour. Although they are close to FD, HWD possess both multiresolution analysis capacities and spatial localization properties. For example, they can be used to decompose the shape of a contour in different resolution levels, thus allowing the analysis of structures appearing at different scales. Low resolution levels then represent the coarse aspects of the contour, while the higher resolution levels represent its increasingly fine details. This multiresolution decomposition allows the analysis of contour irregularities corresponding to a given scale. Moreover, a local modification of the contour results in a local modification of the wavelet coefficients. This spatial localization property can be used, for example, to localize a discontinuity on the contour. This is the main difference with Fourier analysis where a small contour irregularity would modify the whole spectrum of the shape. To illustrate these properties, Figure 1(A) shows an artificial contour composed of three kinds of roughness. Its wavelet representation is shown in Figure 1(B), for resolution levels 2, 3, and 4. Each resolution level allows the extraction of roughness corresponding to a given scale, and also provides the position of the roughness on the contour.

Calculation of the Descriptor

Let $z(n) = x(n) + iy(n)$ be the coordinates of a contour represented in the complex plane and composed of N points, where $N = 2^J$. Let $(a_{2^j+k}, \tilde{a}_{2^j+k})$ be the HWD corresponding to the resolution level j and to the position k on the contour, with $j = 0 \ldots J-2$ and $k = 0 \ldots 2^j - 1$. According to Parseval's theorem, the total energy of the contour is equal to the sum of energies in each resolution level j, according to formula (Newland, 1994):

$$\frac{1}{N}\sum_{n=0}^{N-1}\left|z(n)\right|^2 = \left|a_0\right|^2 + \sum_{j=0}^{J-2}\frac{1}{2^j}\sum_{k=0}^{2^j-1}\left(\left|a_{2^j+k}\right|^2 + \left|\tilde{a}_{2^j+k}\right|^2\right) + \left|a_{N/2}\right|^2 \qquad (1)$$

The coefficient a_0 is null, by translation normalization of the contour. The coefficient $a_{N/2}$ represents an additional wavelet at the Nyquist frequency and can be set to zero without affecting the reconstruction of the contour. The global roughness corresponding to a resolution level j is then evaluated by calculating the energy of this level, according to formula:

$$R_j = \frac{1}{2^j}\sum_{k=0}^{2^j-1}\left(\left|a_{2^j+k}\right|^2 + \left|\tilde{a}_{2^j+k}\right|^2\right), \quad \text{for} \quad j = 0 \ldots J-2 \qquad (2)$$

The multiscale roughness descriptor (MRD) is then defined by the (J-1) roughness coefficients R_j. It measures the distribution of the contour's energy between the

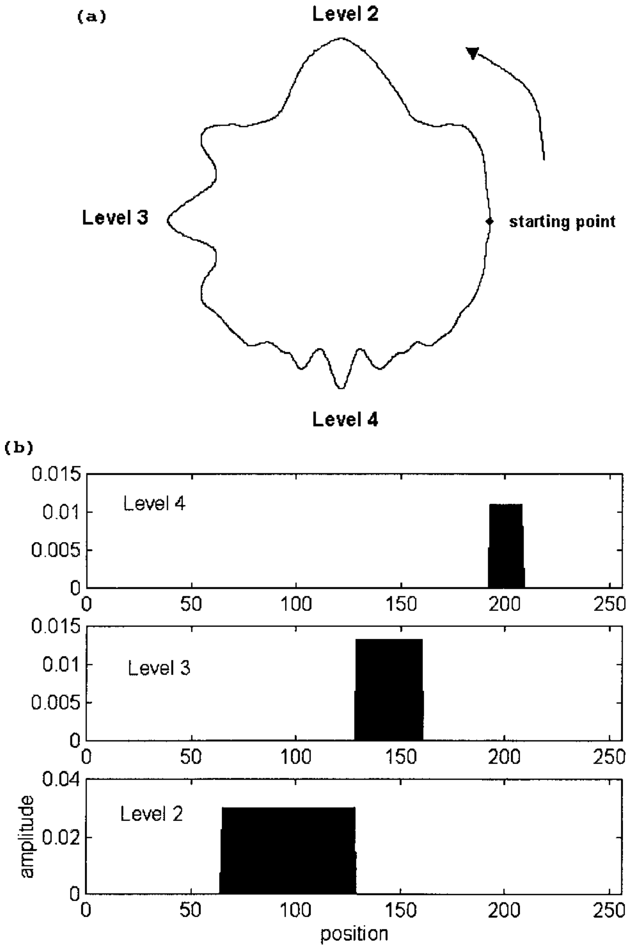

Figure 1. (A) Artificial contour. (B) Representation of the modulus of wavelet coefficients ($|a_{j,k}|^2 + |\tilde{a}_{j,k}|^2$) according to the position on the contour, for resolution levels 2, 3, and 4.

different resolution levels and thus characterizes the importance of each type of roughness, from the coarsest to the finest. Notice that coefficient R_j does not take into account the position of roughness on the contour. Indeed, the shape of particles results from a chaotic process, and roughness is randomly distributed along their boundary. It is therefore preferable to characterize these particles by the importance of roughness at a given scale, rather than by their importance at a given position.

Table 1. Roughness Coefficients of Grains Shown in Figure 2, for the 5 First Levels of Resolution

Level	0	1	2	3	4
Contour 0	1.00	0×10^{-3}	0×10^{-3}	0×10^{-4}	0×10^{-4}
Contour 1	1.07	41×10^{-3}	3×10^{-3}	5×10^{-4}	1×10^{-4}
Contour 2	1.00	9×10^{-3}	7×10^{-3}	3×10^{-4}	0×10^{-4}
Contour 3	1.00	3×10^{-3}	4×10^{-3}	15×10^{-4}	1×10^{-4}
Contour 4	1.02	22×10^{-3}	7×10^{-3}	5×10^{-4}	4×10^{-4}

Interpretation

As an example, Table 1 shows roughness coefficients obtained for five contours represented in Figure 2. Each grain is represented by five coefficients corresponding to the five first resolution levels. Levels superior to 5 are not used because they correspond to noise linked to the low magnification ($\times 100$) used to observe the grains. A strongly worn grain will have the majority of its energy localized in the first level (level 0). The value of level 0 will then characterize its elongation (close to 1 for a round grain, and 1.2 for a very elongated grain). On the other hand, a rough contour will have its energy distributed between the different resolution levels, according to the relative importance of each type of roughness. For example, angular contours will have high coefficients in level 1 or 2, while jagged contours will have high coefficients in level 3 or 4. Thus, MRD allows the characterization of the roughness of a contour at different scales. It measures contributions made by increasingly finer structures and can differentiate small contour roughness from coarser ones.

APPLICATION TO THE STUDY OF SANDY SAMPLES

Samples

One hundred eighty sandy samples from various sources have served as support to this study of form. These samples were collected so as to represent the three

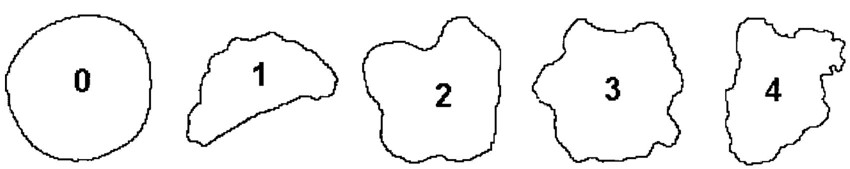

Figure 2. Examples of grain contours, from the most worn (0) to the most rugged (4).

main classes: nonworn sands, marine sands, and eolian sands.

- Nonworn sands (63 samples) were collected in disaggregation arenites, glacial drifts deposits, river lag deposits, alluvial terraces, torrents, lakes, and small beaches. The contour of these grains is often angular, and their faces are plane, smooth, or slightly streaked. One observes breakages, or irregular faces forming between them sharp, obtuse, or right angles.
- Marine sands (63 samples) were collected in recent marine beds or in ancient marine beds. Most of these grains have contours with blunt angles, or even rounded, and their faces are generally shiny.
- Eolian sands (54 samples) were collected on ergs, Saharan dunes, terrestrial, and coastal dunes. The form of these grains is blunt or round, more rounded on average than in the case of marine sands. Moreover, their surfaces are often frosted.

Each sample is treated in warm HCl and the 315- to 400-μm fraction is retained by sieving. One hundred quartz grains are then examined under the microscope with a magnification of $\times 100$ and their contour is digitalized at a resolution of 768×576 pixels. Contours are resampled on 1024 points and the multiscale roughness descriptor is then computed for each grain.

Statistical Analysis

Roughness analysis does not offer a real sedimentological interest except in the case of a population of grains that represents a sample of sand. Actually, a sandy sample is composed of a multitude of grains, each of them having its particular form. The characterization of a sample is then carried out by calculating a frequency histogram, for each resolution level. Each sample is then represented by six histograms, corresponding to resolution levels 0–5. Scales of resolution being independent of each other, each histogram yields different information on the sample (degree of sphericity, angularity, roundness. . .) and quantifies increasingly finer aspects of the contour of grains.

Figure 3 shows some examples of histograms obtained for the level of resolution 4. Roughness values are decomposed in 18 classes. Roughness classes are plotted on abscissa and the percentage of grains in each class is plotted on ordinate. After analysis, level 4, which characterizes fine grain irregularities, appeared as the most discriminant level. Eolian sands, strongly worn, have a roughness coefficient R_4 very low and often less than 0.5×10^{-4}. They present a marked mode with a percentage of grains generally exceeding 60% in the first class. Marine sands are a bit more heterogeneous, with a lower mode (40%–50%), and their roughness coefficient can reach 1×10^{-4}. On the other hand, nonworn sands (river or lake deposits) have very extensive statistical distributions, with roughness coefficients

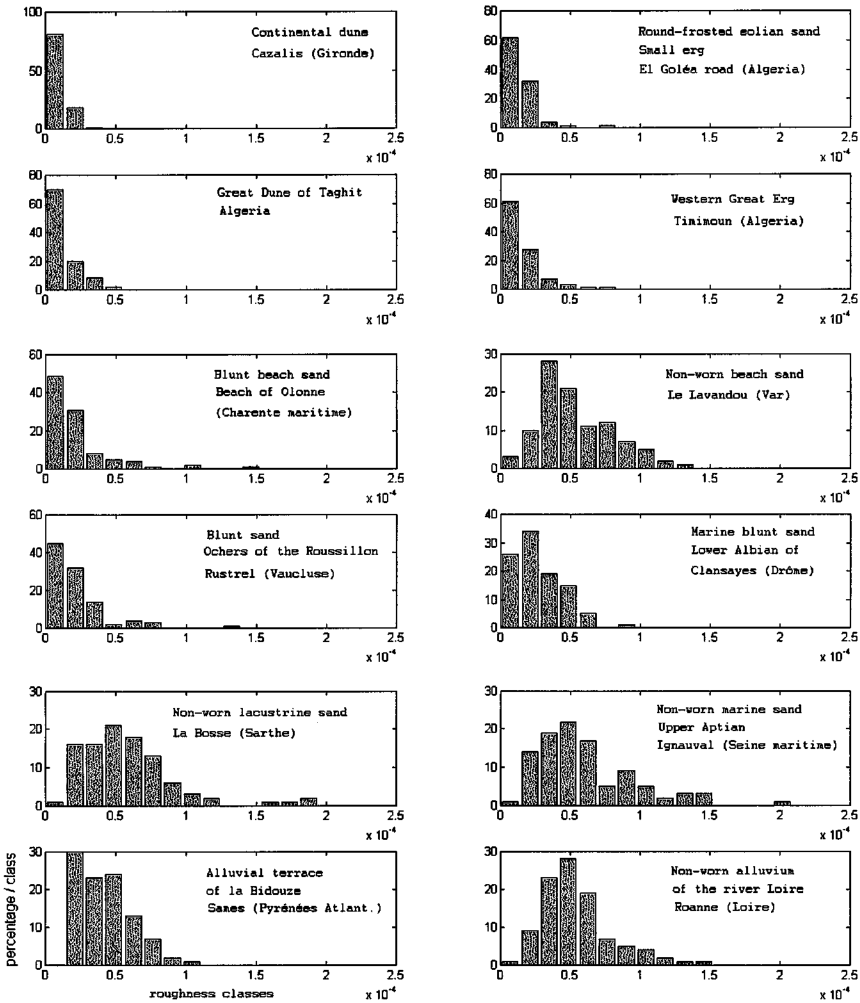

Figure 3. Examples of histograms of MRD (level 4) for sands of different milieus (X-axis: roughness classes; Y-axis: percentage of grains in each class).

often greater than 1×10^{-4}. Furthermore, the percentage of grains in the first class is always inferior to 5%.

Data Visualization

The form of frequency histograms is an indicator of the degree of sand evolution and of their heterogeneity. Nevertheless, it can be difficult to judge the likeness between the different histograms and therefore between the different statistical distributions. To measure this resemblance, the use of the Euclidean distance between histograms seems to be a natural estimator. Each histogram being made up of 18 classes, the distance between two histograms is therefore given by the distance between two vectors with 18 components. To study proximity relationships between samples, we have therefore used a Sammon mapping. The Sammon mapping (Sammon, 1969) is a topographic nonlinear mapping that seeks to maintain the structure of the data when it is projected from a high-dimensional space onto a lower dimensional space (generally 2D or 3D). The preservation of this structure is achieved by maintaining the distance between patterns under projection. The distance measure between two patterns is commonly the Euclidean metric.

Let d^*_{ij} be the distance between points x_i and x_j in the data space and d_{ij} be the distance between their corresponding images y_i and y_j in the projected space. The projection of N points is generated by minimization of the *Sammon stress* error, defined as

$$E = \frac{1}{\sum_{i<j} d^*_{ij}} \sum_{i<j}^{N} \frac{\left(d^*_{ij} - d_{ij}\right)^2}{d^*_{ij}} \tag{3}$$

This dimensionless error is a measure of how well the interpattern distances are preserved during the projection. Since $d_{ij} = \|y_i - y_j\|$, the mapping can be determined by adjusting the points y iteratively by a gradient descent method, so as to minimize the error.

The Sammon mapping thus attempts to keep points that are close together in the input space close together in the projected space, and similarly for distant points. In this manner, any clustering is approximately preserved and it becomes possible to visualize them. Figures 4 and 5 show mappings obtained for resolution levels 0 and 4 respectively.

Level 0, which characterizes elongation, allows the identification of strongly rounded and of Eolian origin sands. These are generally sands that have undergone a joint evolution, marine, and eolian (Pyla, Arcachon, Oléron), or sands having undergone a strong ancient water wearing and a recent wind wearing (Taghit, Cazalis). It is difficult to distinguish the other sands, whether they are eolian, marine, or nonworn, because degrees of sphericity are generally highly variable and this diversity is also found within the same sample. This phenomenon is explained by the fact that transport agents do not modify the general form of grains, except in the case of an intense eolian abrasion.

Level 4 on the other hand (Fig. 5) allows to translate with fidelity the degree of sand evolution from most rounded and polished grains (joint evolution, marine and eolian) to most irregular and most rugged grains (torrents, disaggregation

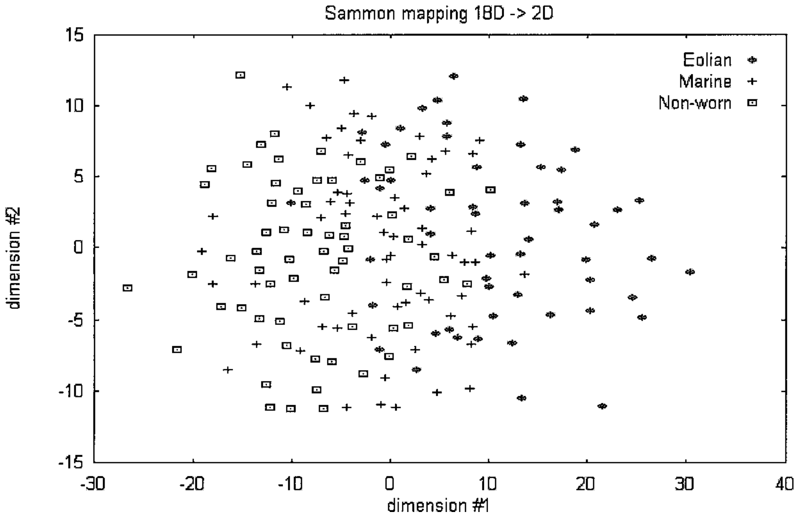

Figure 4. Sammon mapping obtained from histograms of level 0.

arenites). Between sands from wind evolution and nonevoluted sands is situated an area representing sands with a more complex or less known evolution, generally marine. When several samples stemming from the same bed are analyzed, their representative points are clustered in a satisfactory manner. The method

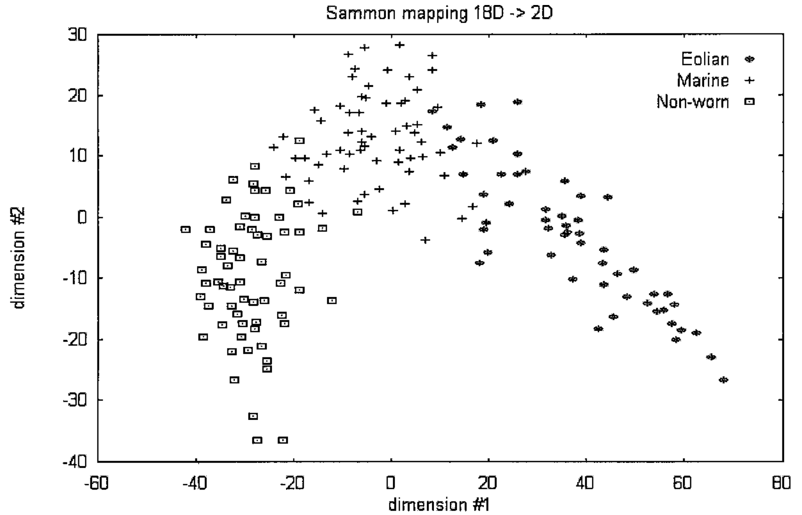

Figure 5. Sammon mapping obtained from histograms of level 4.

therefore seems reproducible and able to characterize the degree of evolution of a bed.

CLASSIFICATION

Principle

Histograms obtained from level 4 have been used to conceive a classifier implemented by a multilayer perceptron (MLP) (see, e.g., Bishop 1995; Hertz, Krogh, and Palmer, 1991, for an introduction to neural networks). MLPs are well-known global function approximators and their performance is based on a well-developed mathematical background. The basic idea is to use the MLP to model wear and erosion phenomena acting on quartz grains. The network, which uses sigmoid functions, consists of 18 input units, 6 hidden units, and 3 output units, corresponding to the three classes (eolian, marine, and nonworn). The database has been randomly divided into two sets: one for training the network (89 samples) and the other for testing the classifier (91 samples). The network is trained with a conjugate gradient algorithm (Møller, 1993).

Results

After learning, the network classifies correctly 100% of samples. Nevertheless, the goal of network training is not to learn an exact representation of the training set itself, but rather to build a statistical model of the process which generates the data. In this manner, the network is usually capable to *generalize*, i.e., to make good predictions for new inputs. The validity of the classifier has therefore been tested on the test set. The classification results are given as a confusion matrix in Table 2. The network recognizes relatively well eolian and nonworn grains, with a classification rate superior to 93%. On the other hand, it has more difficulty to recognize grains of marine type (average rate of 87.5% of classification), and has a tendency to classify them as eolian (9.38% of grains). Nevertheless, the global correct classification rate (92.22%) indicates that the statistical model built by the

Table 2. Confusion Matrix for the Classification Results

	Eolian	Marine	Nonworn
Eolian	96.15%	3.85%	0%
Marine	9.38%	87.5%	3.12%
Nonworn	0%	6.25%	93.75%

Note. Rows indicate the true classes, whereas columns give the classification performed by the network.

network is perfectly valid. By using MLP, it seems therefore possible to model wear phenomena acting on particles and thus to reconstitute the history of sands in sedimentary basins.

Comparison With Other Mathematical Methods

As a comparison, we carried out the same study by using the Fourier descriptors and the fractal dimension.

Fourier Descriptors

The descriptors used are these described in Thomas, Wiltshire, and Williams (1995), where they are used to classify quartz grains. Each contour is resampled on 1024 points and Fourier coefficients are calculated for each grain. Harmonics -127 to $+128$ are then retained for the analysis. The most discriminant harmonic is then identified by a relative entropy analysis of the amplitude frequency distributions of each harmonic (Full, Ehrlich, and Kennedy, 1984; Mazzullo, Sims, and Cunningham, 1986). Each sample is then represented by a frequency histogram, made up of 18 classes, corresponding to the selected harmonic. This analysis led to keep harmonic -42 to represent samples.

Fractal Dimension

The fractal dimension is calculated for each grain by using the method of Richardson (Hoyez, 1994). Each sample is then represented by a frequency histogram of fractal dimensions, and made up of 18 classes. Dimensions vary from 1.00 to 1.04 following the roughness of grains.

Histograms obtained by these two methods are presented at the input of two neural networks possessing the same architecture as for the MRD. Learning and test procedures are the same as those described previously. Table 3 indicates results obtained by the three methods and gives the global correct classification rate obtained for the learning phase and the test phase. Fourier and fractal analysis are

Table 3. Comparison of Performances of the Different Description Methods

Method	Learning	Test
Fourier descriptors	100%	81.11%
Fractal dimension	97.7%	71.11%
MRD level 4	100%	92.22%

efficient methods when the problem is to separate sands having undergone very different degrees of wear. Statistical analysis shows that eolian and nonworn samples are clearly separated. On the other hand, marine sands overlap both eolian and nonworn sands and are consequently more difficult to separate. This overlapping is more pronounced when the fractal dimension is used, which explains the low classification result obtained during the test.

CONCLUSION

By using the harmonic wavelet transform, we have proposed a new method for particle characterization, allowing the quantification of the roughness of grains, that is their irregularities at different scales. The MRD decomposes the contour into independent components, each component corresponding to a different resolution level. Low resolution levels bring information on the global form (elongation, angularity) while higher levels of resolution quantify finer roughness of the contour. This descriptor possesses good invariance properties, a low sensitivity to noise and the interpretation of its coefficients is relatively simple.

Its application to sediments analysis shows that level 4 is the most discriminant resolution level. It allows the identification of the three main sand types, namely eolian sands, marine sands, and nonworn sands. When used for classification, the descriptor proves to be far more effective than methods usually used for particle description. This method of description therefore seems to allow for the characterization of the degree and the type of wear of sands, and should be able to be applied to other types of particles, from the simplest to the most complex.

REFERENCES

Bishop, C. M., 1995, Neural Networks for Pattern Recognition: Oxford University Press, Oxford, 482 p.

Boon, J. D., Evans, D. A., and Hennigar, H. F., 1982, Spectral information from Fourier analysis of digitized quartz grain profiles: Math. Geol. v. 14, no. 6, p. 589–605.

Cesar, R. M., Jr., and Costa, L. da F., 1996, Towards effective planar shape representation with multiscale digital curvature analysis based on signal processing techniques: Pattern Recognit., v. 29, no. 9, p. 1559–1569.

Chuang, G. C.-H., and Kuo, C.-C. J., 1996, Wavelet descriptor of planar curves: Theory and applications: IEEE Trans. Image Process., v. 5, no. 1, p. 56–70.

Clark, M. W., 1981, Quantitative shape analysis: A review: Math. Geol., v. 13, no. 4, p. 303–320.

Diepenbroek, M., Bartholomä, A., and Ibbeken, H., 1992, How round is round? A new approach to the topic "roundness" by Fourier grain shape analysis: Sedimentology, v. 39, p. 411–422.

Drolon, H., Hoyez, B., Druaux, F., and Faure, A., 1999, Analyse par ondelettes de la rugosité des grains de sable : C.R. Acad. Sci. Paris, Sciences de la terre et des planètes, v. 328, p. 457–461.

Ehrlich, R., and Weinberg, B., 1970, An exact method for characterization of grain shape: J. Sedimentary Petrol., v. 40, p. 205–212.
Full, W., Ehrlich, R., and Kennedy, S., 1984, Optimal configuration and information content of sample data generally displayed as histograms or frequency plots: J. Sedimentary Petrol., v. 54, p. 117–126.
Granlund, G. H., 1972, Fourier preprocessing for hand print character recognition: IEEE Trans. Comput., v. 21, p. 195–201.
Hertz, J., Krogh, A., and Palmer, R. G., 1991, Introduction to the theory of neural computation: Addison-Wesley, Reading, MA, 327 p.
Hoyez, B., 1994, The roughness of sand grains: An application of Fourier analysis and of fractal dimension: Ann. Soc. Géol. du Nord, v. 3, 2ème série, p. 73–83. (In French)
Kaye, B. H., 1978, Specifications of the roughness and/or texture of a fine particle by its fractal dimension: Powder Technol., v. 21, p. 1–16.
Kennedy, S. K., and Wei-Hsiung, L., 1992, A comparison of Fourier and fractal techniques in the analysis of closed forms: J. Sedimentary Petrol., v. 62, no. 5, p. 842–848.
Mazzullo, J., Alexander, A., Tieh, and T., and Menglin, D., 1992, The effects of wind transport on the shapes of quartz silt grains: J. Sedimentary Petrol., v. 62, no. 6, p. 961–971.
Mazzullo, J., and Magenheimer, S., 1987, The original shape of quartz sand grains: J. Sedimentary Petrol., v. 57, no. 3, p. 479–487.
Mazzullo, J., Sims, D., and Cunningham, D., 1986, The effects of eolian sorting and abrasion upon the shape of quartz sand grains: J. Sedimentary Petrol., v. 56, no. 1, p. 45–56.
Mokhtarian, F., and Mackworth, A. K., 1992, A theory of multiscale, curvature-based shape representation for planar curves: IEEE Trans. Pattern Anal. Mach. Intellig., v. 14, no. 8, p. 789–805.
Møller, M. F., 1993, A scaled conjugate gradient algorithm for fast supervised learning: Neural Networks, v. 6, no. 4, p. 525–533.
Newland, D. E., 1993, Harmonic wavelet analysis: Proc. R. Soc. Lond. Ser., A, v. 443, no. 1917, p. 203–225.
Newland, D. E., 1994, An introduction to random vibrations, spectral and wavelet analysis, 3rd edn.: Longman Harlow, England, 480 p.
Pirard, E., 1992, Roughness analysis on powders using mathematical morphology: Acta Stereol., v. 11, Suppl. 1, p. 533–538.
Pirard, E., 1994, Shape processing and analysis using the calypter: J. Microsc. v. 175, no. 3, p. 214–221.
Sammon, J. W., Jr., 1969, A non linear mapping for data structure analysis: IEEE Trans. Comput. v. C-18, no. 5, p. 401–409.
Thomas, M. C., Wiltshire, R. J., and Williams, A. T., 1995, The use of Fourier descriptors in the classification of particle shape: Sedimentology, v. 42, p. 635–645.
Vallejo, L. E., 1995, Fractal analysis of granular materials: Geotechnique, v. 45, p. 159–163.