Math Geosci (2016) 48:835–870
DOI 10.1007/s11004-016-9648-8

Analysis and Classification of Natural Rock Textures based on New Transform-based Features

Rodrigo Lobos[1,2] · Jorge F. Silva[2] · Julián M. Ortiz[3] · Gonzalo Díaz[4] · Alvaro Egaña[4]

Received: 25 June 2015 / Accepted: 6 July 2016 / Published online: 26 July 2016

Abstract This work develops a mathematical method to extract relevant information about natural rock textures to address the problem of automatic classification. Classical methods of texture analysis cannot be directly applied in this context, since rock textures are typically characterized by both stationary patterns (a classic kind of texture) and geometric forms, which are not properly captured with conventional methods. Due to the presence of these two phenomena, a new classification approach is proposed in which each rock texture class is individually analyzed developing a specific low-dimensional discriminative feature. For this task, multi-scale transform domain representations are adopted, allowing the analysis of the images at several levels of scale and orientation. The proposed method is applied to a database of digital pho-

✉ Rodrigo Lobos
rlobos@usc.edu; rlobos@ing.uchile.cl

Jorge F. Silva
josilva@ing.uchile.cl

Julián M. Ortiz
jortiz@ing.uchile.cl

Gonzalo Díaz
gonzdiaz@ing.uchile.cl

Alvaro Egaña
aegana@alges.cl

1 Department of Electrical Engineering, University of Southern California, Los Angeles, USA

2 Department of Electrical Engineering, Information and Decision Systems Group, Universidad de Chile, Av. Tupper 2007, 412-3, Room 508, Santiago, Chile

3 Department of Mining Engineering, Universidad de Chile, Av. Tupper 2069, Santiago, Chile

4 ALGES Lab, Advanced Mining Technology Center (AMTC), Universidad de Chile, Santiago, Chile

tographs acquired in a porphyry copper mining project, showing better performance than state-of-the-art techniques, and additionally presenting a low computational cost.

Keywords Rock textures · Texture classification · Texture and object source separation · Feature design · Wavelets and multi-resolution analysis · Pattern recognition

1 Introduction

Resource evaluation is of great interest in mining because it is the first stage in the evaluation of the project, where the quantity and quality of the economic ore must be determined. Subsequent stages such as mine design, planning and process optimization are performed based on the block model defined during this stage. Currently, models go beyond the quantification of concentrations of the elements of interest. A full material characterization is required to assess the performance that the rock will have in a given process during mining or at the metallurgical plant; therefore, in addition to a proper estimate of the tonnage and grades of interest, other characteristics of the rock must also be determined. The block model is built considering different kinds of information, such as geophysical data, geological mapping and logging, results from analytical and metallurgical tests over the rock, and geotechnical surveys. Visual information takes a leading role in this characterization process. Many of the characteristics of the rock are judged from visual inspection: lithology types require a judgement about the texture of the rock; alteration types are typically defined based on the presence of alteration minerals, which may appear on the rock with diverse levels of intensity; clast size and shape are relevant to define the fabric of the rock, etc. Most of this information will finally be stored as classes in a database, and carried downstream through the mining process. A proper characterization of some of these features has been found to be extremely relevant but quite difficult to perform, especially when considering that a consistent characterization requires a trained geologist. This motivates the study of rock textures as a key visual feature of the rock (Mock and Jerram 2005; Higgins 2006; Higgins and Chandrasekharam 2007), and also because the relationship with metallurgical performance has been previously established in many studies (Vink 1997; Bojcevski 2004; Croop and Goodall 2013). A good rock characterization allows us to have a better prediction of mining and metallurgical process responses (Dunham and Vann 2007), which has a direct impact on the financial performance of a project (Dunham et al. 2011).

Despite its great potential, rock texture information is not typically used in mining. This may be explained due to difficulty in defining, characterizing and quantifying textures (Bonnici 2012). However, textures have been widely investigated in areas such as image processing and computer vision (Do and Vetterli 2002; Ojala et al. 2002; Zujovic et al. 2013; Petrou and Sevilla 2006), which offers the possibility of bringing many new algorithms and methods to the geosciences. Nevertheless, it should be noted that the definition of textures in geology, mining engineering and image processing can be different, which must be taken into account before applying any method.

Several other problems in the mining context have been addressed using computer-aided techniques and digital image processing, such as ore sorting (Tessier 2007;

Perez et al. 2011), lithology classification (Chica-Olmo and Abarca-Hernandez 2000; Thomas et al. 2011), bubble sizing estimation on froth flotation processes (Bonifazi et al. 2001; Kracht et al. 2013), and detection of big rocks during operation (Cabello et al. 2002). Furthermore, digital image processing of rocks has been also used for rock classification (Lepisto et al. 2003; Singh et al. 2004). For rock texture classification, different approaches have been considered, including the use of gray-level co-occurrence matrices (Partio et al. 2002), the use of spectral features and natural network classifiers (Kachanubal and Udomhunsakul 2008), and a hierarchical neuro-fuzzy class method based on binary space partitioning (Gonçalves and Leta 2010). These techniques explore the use of textures features; however, their settings exclusively include classes with homogeneous (stationary) textures, and the context in which these techniques are applied does not correspond to mining.

This works develops new mathematical methods to extract low-dimensional discriminative features with the objective of performing automatic rock texture classification. A new approach is proposed, which rests on the extraction of low-dimensional discriminative features from the digital images based on transform domains and source separation algorithms. Six classes of rock textures frequently found in porphyry copper deposits are studied. These low discriminative features are designed on a case by case basis and, thus, the proposed method is customized to the specific deposit and texture classes. In this context, a classification scheme composed of a concatenation of five binary detector is built. The proposed approach allows one to deal with the different nature that a texture image could present, that is, as a stationary or as a structural pattern. For that, the classes are divided into two groups, which will be called: stationary textures and structural textures. The first group, which is a typical case studied in image processing (Petrou and Sevilla 2006), contains rock textures defined by a stationary pattern formed by crystals with particular spatial distribution (shape and size). The second group contains rock textures that, as a consequence of the behavior of the minerals, present structures of different shape, size and orientation, such as clasts (rock fragments) and veins (fractures in rocks filled with minerals) (Tarbuck et al. 2000). The existence of the aforementioned two groups does not allow the application of standard methods of texture analysis, since the structural information of the second group is not properly treated (Do and Vetterli 2002; Ojala et al. 2002; Zujovic et al. 2013; Petrou and Sevilla 2006). Furthermore, other methods, specially developed to process rock texture information, do not consider the presence of structures (Jungmann et al. 2011; Donskoi et al. 2007; Chatterjee 2013; Lepistö et al. 2005).

The focus of this work is on the feature extraction (FE), which is the development of low-dimensional discriminative features. Specifically, a set of features is developed for each binary classifier and is specially designed to extract characteristic information (signature) of a particular class. For this task, transform domains given by the wavelet transform (WT) (Mallat 2008) and the shearlet transform (ST) (Kutyniok et al. 2012) are used to properly analyze the stationary patterns and geometric forms of rock textures, while considering different levels of scale and orientation. As these two phenomena (stationary patterns and geometric forms) might coexist in an image, source separation algorithms (Szolgay and Szirányi 2012; Starck et al. 2005) are integrated in our feature design. This classification approach shows better performance than the state-of-the-art methods used for texture discrimination of natural images.

The superiority of the proposed features can be explained by the fact that we took into consideration stationary patterns and structural information at different levels of scale and orientation, depending on the nature (signature) of the considered rock classes.

The rest of the paper is organized as follows. Section 2 introduces the rock texture classification problem and gives specific details about the database. Section 3 gives a general description of the image classification problem and our classification approach. Section 4 presents the main contribution, where the development process for each low-dimensional discriminative feature is specified. Section 5 explains the mathematical expressions needed to assess the performance of the proposed classification scheme, and shows the experimental results and their analysis. Section 6 provides a summary of our results and some final remarks.

2 The Database of Natural Rock Textures

For our study, textures of rocks captured from drill hole cores digital photographs from a porphyry copper mining project are used to construct a database. Drill hole cores are stored in trays, which are photographed at high resolution. Although it is well known that textures are correlated to some of the process performance responses, given the difficulty in measuring and classifying them, they are seldom used as a rock feature in the construction of models of the ore deposits. Nonetheless, geologists have a description of the different types of textures that are of interest and, thus, a database was constructed, where a number of cases were labeled with one of six texture categories. Typically, the rock texture can be described in terms of the matrix texture which represents the background of the rock, and in terms of the objects present in the matrix, namely crystals, clasts, and structures. For our application, six categories are identified (Fig. 1), and described as: (1) aphanitic class, characterized by a matrix of the rock with the presence of small crystals, those of which only can be observed through a microscope, leads to a homogeneous surface (the matrix) with slight and regular local variations; (2) phaneritic class, characterized by larger crystals that constitute the matrix of the rock and which can be observed through the naked eye, leads to a surface that shows a homogeneous character with rougher variations (at larger scale) than the aphanitic class; (3) porphyry class, features a matrix of small sized crystals (microscopic scale), combined with larger crystals, which are visible by the naked eye scale, called porphyds, much larger than those found in the phaneritic class; (4) vein class, characterized by the presence of one or more of veins, usually depicting the same orientation, that is fractures on the rock with a sheetlike form that are filled with other minerals; (5) stockwork class, features a random array of veins, with varying orientations; and (6) breccia class, characterized by a homogeneous matrix containing large clasts that are objects with angular form.

The database has 16 gray-scale images of 128×128 pixels per class, in total 96 images. The images selected show pure classes. Nevertheless, we know, in practice, there could be an overlap of two or more classes in a single image.

For analysis purposes, it is relevant to identify two groups within the set of images. Firstly, a group called stationary textures is identified, composed of samples from the aphanitic, phaneritic and porphyry classes, where low-scale elements form a stationary

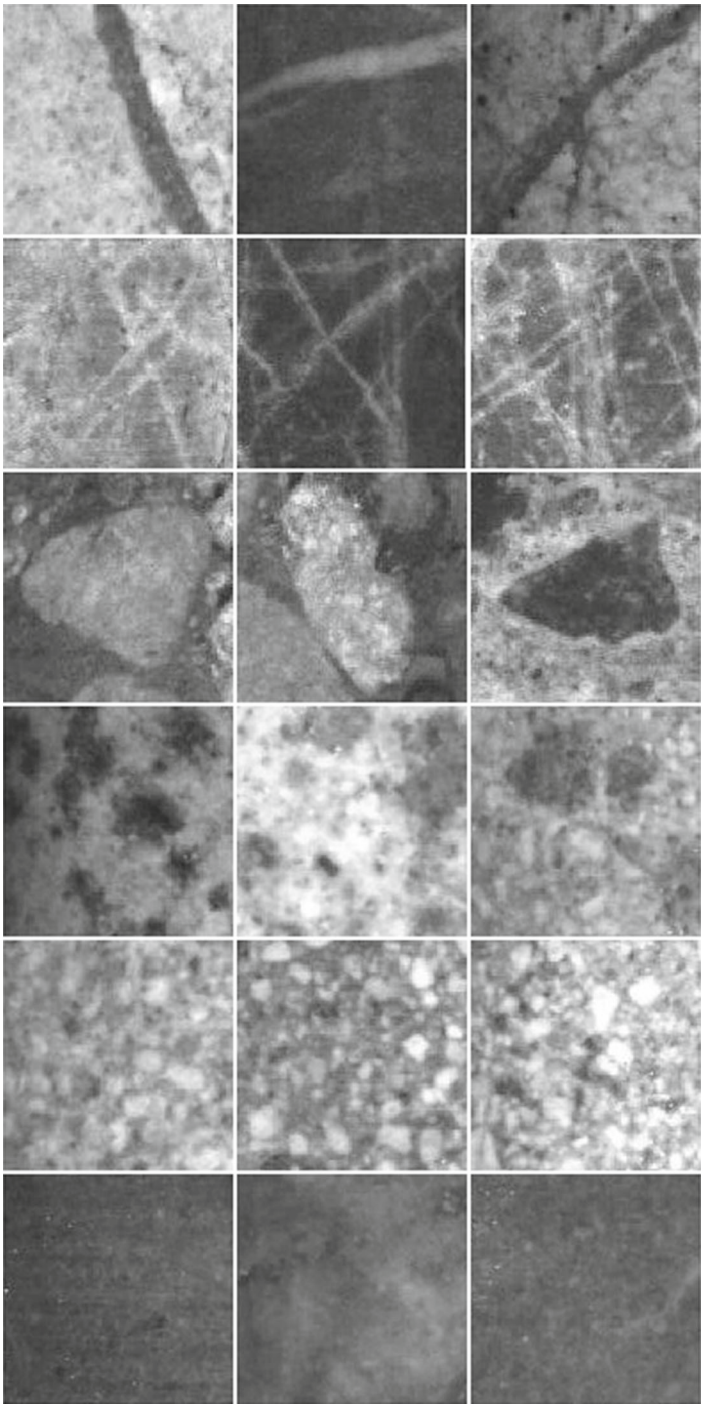


Fig. 1 Examples of each of the six classes analyzed in this work. (From *left*) *Each column* shows three images of aphanitic, phaneritic, porphyry, breccia, stockwork and vein, respectively

pattern formed by crystals with a particular distribution, shape and size. This group corresponds to textures associated with igneous rocks. It should be noted that the term "texture" in this case is applied to describe the shape, size and ordering of the crystals that form a stationary pattern. Therefore, the main difference between textures in this set is related to their crystals. In the aphanitic class, there are small crystals which only can be observed through a microscope; in the phaneritic class, there are larger crystals that can be observed by the naked eye; in the porphyry class, there are two types of crystals: small size crystals (microscopic scale) and larger crystals (naked eye scale) called porphyds. Furthermore, this last class has the biggest size crystals among the three classes of the group. Secondly, a group called structural textures is composed of samples from the vein, stockwork, and breccia classes, where structural features define the corresponding texture, based on the shape and relative orientation of these objects within the rock (Tarbuck et al. 2000). These textures are also associated with igneous processes, where intrusion of magmas within already formed rock generates the structural feature. In the vein class, the rock shows a fracture that is filled by a different material; therefore, the characteristic feature is the presence of one or few structures with similar orientation. When the rock has been broken by several veins in random orientations, then the class is labeled as stockwork. Finally, samples of the breccia class are characterized by the presence of large clasts (fragments of rocks with sharp edges), within a homogeneous matrix.

2.1 The Imaging Method

Images used for this study were obtained from the photographs that are routinely recorded from the drill hole cores obtained at a porphyry copper deposit in northern Chile, during the exploration campaign of mining projects. A database of photographs of full trays containing 4.5 m of core length was obtained with a regular digital camera, usually ranging from 4 to 8 megapixel of resolution under controlled illuminating conditions. These images are stored for future consultation in case a core sample needs to be identified and recovered from the drill hole core storage. For our study, sections of the recovered core with pure textures or with clear structures of the six types were selected and labeled by a experienced field geologist of the project. From these sections, a library of 16 images for each texture class considered in this study was built, with 128×128 pixels in resolution each. It should be emphasized that this control library was built in this manner to simplify the study and allow the design and development of specific classifiers considering the characteristic features for the six texture classes.

3 The Chain Sequential Classification Approach

In this section,we explain the classification approach adopted in this work. First, the image classification problem is presented in general to, then, show the adopted sequential classification scheme.

3.1 The Image Classification Problem

An image z can be represented as a two-dimensional vector $z = (z(i, j))_{\{(i,j)\}}$ where $(i, j) \in \{1, \ldots, n_1\} \times \{1, \ldots, n_2\}$. For simplicity, the entries of an image (or pixels) can be re-sorted in a one-dimensional vector of dimension $n = n_1 \times n_2$, where each image z belongs to a subspace $\mathscr{Z} \subset \mathbb{R}^n$. Associated with each image z, there is a hidden (non-observed) class label that is denoted by $y \in \mathscr{Y} = \{1, \ldots, M\}$, which represents the identity (or class) of the image. The classification task consists of finding a mapping from $\mathscr{Z}$ to $\mathscr{Y}$, which from the observation stipulates the identity or class of the image. In the Bayesian setting, the image is modeled as a realization of a random vector $Z(w)$ with values in $\mathscr{Z}$ and the class label is a random variable $Y(w)$ with values in $\mathscr{Y}$. Then, considering the distribution of the joint vector (Z, Y), the optimal rule (that minimizes the probability of error) is given by (Vasconcelos (2004) and Silva and Narayanan (2012))

$$g^*(z) = \underset{i \in \mathscr{Y}}{\operatorname{argmax}} P(Y = i | Z = z), \quad \forall z \in \mathscr{Z}, \tag{1}$$

where $P(Y = i|Z = z)$ denotes the a posteriori probability of class being i given the observation z. This is the well-known maximum a posteriori (MAP) decision rule (Vasconcelos 2004).

Unfortunately, to determine $g^*(\cdot)$, it is necessary to know the joint distribution $P(Z, Y)$, which is not available in practice. In contrast, we can assume that we have access to independent and identically distributed (i.i.d.) samples

$$\mathscr{D} = \{(z_1, y_1), (z_2, y_2), \ldots, (z_N, y_N)\},$$

which corresponds to the supervised (training) data. This raises a learning problem, which corresponds to approximate the optimal rule empirically from the data $\mathscr{D}$. For this, the problem of dimensionality reduction of the observation space $\mathscr{Z}$ needs to be addressed (van der Maaten et al. 2009). More precisely, starting from a big observation space of dimension n, a feature transformation (FT) $f : \mathscr{Z} \to \mathscr{X} \subset \mathbb{R}^m$ needs to be designed where $m < n$. In this context, $\mathscr{X}$ denotes the feature space, and $X = f(Z)$ is the new feature vector to do the inference about Y. The application of the FT is known as feature extraction (FE) process. FE is essential to reduce the learning problem to something that is tractable in terms of dimensionality, given the limited data $\mathscr{D} = \{(z_1, y_1), (z_2, y_2), \ldots, (z_N, y_N)\}$. The ideal objective is that $f(Z)$ keeps (on average) almost all the information of Z to discriminate the hidden variable Y (Vasconcelos 2004; Silva and Narayanan 2012).

3.2 The Chain Sequential Classifier

For our problem, a sequential approach conformed by a bank of binary classifiers (or detectors) is proposed. The basic idea was to simplify the decision scheme by a significant reduction in the dimensionality of the problem (in the FE phase), and the adoption of a sequence of simpler binary detectors. Therefore, the sequential

classification strategy decomposes the global decision in a set of binary detectors (partial decisions), where the objective in each step of the chain is to put the focus on one of the texture classes at a time, and design a description (feature extraction) that captures what makes this image class distinctive from the rest. The assumption used to adopt this divide to conquer approach is that every rock texture has a distinctive identity that can be represented with a parsimonious representation that discriminates this class from the rest with high accuracy. Under this assumption, a set of binary classifiers acting one after the other (on the designed features) offers an attractive solution for the global decision problem. On the other hand, the incremental nature of this approach provides the possibility to extend it to more texture classes, while keeping the complexity of the strategy constrained, as every binary detector is acting individually on a relatively small feature space by design.

Regarding the binary classifier, under the assumption that every texture class can have a distinctive small dimensional feature vector (that discriminates it from the rest), non-parametric discriminative classifiers can work perfectly well at every stage in the chain as the celebrated support vector machine (SVM). SVM classifiers benefit from the discrimination power of good features and, consequently, it is a perfect match to the feature design oriented strategy adopted in this work. In particular, SVM offers the capacity of defining a large margin decision region in the designed feature space with a simple decision rule based on the use of reproducible kernels Hilbert spaces and a linear decision approach that avoid overfitting to training conditions (Burges 1998; Cherkassky and Mulier 2007; Duda et al. 2012). This learning issue is important in our problem, as limited supervised training data are available. Then, the proposed discriminative approach is the best choice to avoid major estimation errors by reducing the learning to the definition of a large margin decision region (in the feature space) instead of more classical non-discriminative approaches that requires the estimation of the joint distribution (Burges 1998).

The proposed chain sequential classifier (CSC), denoted by $g_{\text{chain}}(\cdot)$, works as follows: for every class index $i \in \{1, 2, 3, 4, 5\}$, we design a FT

$$f_i : \mathscr{Z} \to \mathscr{X}_i, \tag{2}$$

where $X_i = f_i(Z)$ is the resulting feature vector, and its associated binary classifier is given by

$$g_i : \mathscr{X}_i \to \{0, 1\}, \quad i \in \{1, 2, 3, 4, 5\}. \tag{3}$$

The general idea of the decision scheme is the following: the first classifier $g_1(\cdot)$ decides or "detects" the presence of the class 1, the second $g_2(\cdot)$, based on the decision of the first stage, decides or detects the presence of the class 2, until the last $g_5(\cdot)$ that decides if the observation belongs to the class 5 or 6. In other words, given the raw observation Z our scheme receives as an entry a set of 5 partial binary decisions

$$(g_1(f_1(Z)), g_2(f_2(Z)), \ldots, g_5(f_5(Z))) \in \{0, 1\}^5, \tag{4}$$

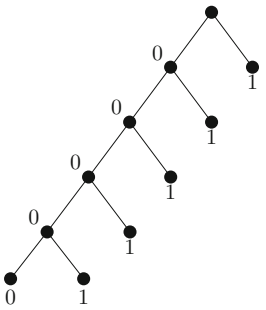


Fig. 2 Decision-tree-decoding structure of the sequential classification approach

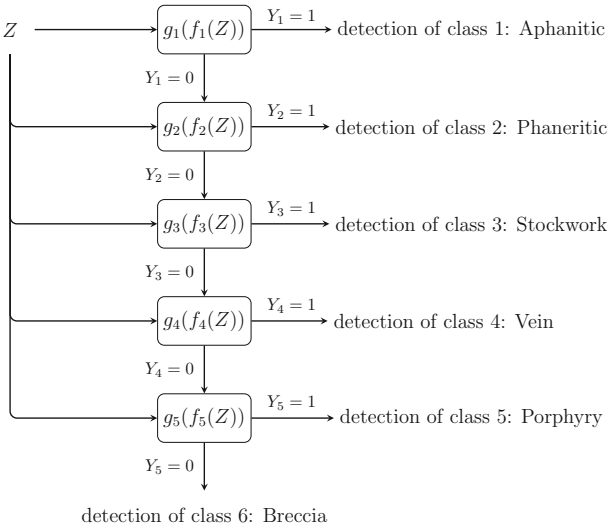


Fig. 3 Block diagram of the proposed chain sequential classifier

where the final decision decodes that path by the tree structure illustrated in Fig. 2. More formally, if $\hat{Y}_i = g_i(f_i(Z)) \in \{0, 1\}$ denotes the output of the i-classifier (with $i \in \{1, 2, 3, 4, 5\}$), and $\hat{Y}_6 = 1$ is fixed, then the final decision is given by

$$\hat{Y}(Z) = g_{\text{chain}}(Z) = \min \left\{ i \in \{1, \ldots, 6\} : \tilde{Y}_i = 1 \right\}. \tag{5}$$

4 Transform-based Feature Design for Binary Detection

This work puts the focus on the FE, which is the problem of designing $\{f_i\}_{i=1}^5$ in Eq. (2) for each binary classifier in the sequential strategy illustrated in Fig. 3. The objective for each stage is to find a parsimonious representation with a good discriminative power to discriminate the rock class of interest. In this section, we develop the feature transformation (FT) in Eq. (2) for each geological texture class in Eq. (3), and we evaluate its detection rate adopting SVM classifiers.

4.1 Aphanitic Class Detection: First Binary Classifier

Stationary textures present stationary patterns at different scales (Tarbuck et al. 2000). To detect the presence of these stationary patterns, it is necessary to have a technique that analyzes them at different scales or resolutions. For that, the texture-cartoon separation technique proposed by Szolgay and Szirányi (2012) is adopted. This technique is capable of separating the stationary patterns (or texture) from the geometric forms (or cartoon) according to a pre-established filter of width $\sigma > 0$. The parameter σ is associated with the scale of the texture-cartoon separation (for more details see Appendix 3). Therefore, changing the value of the width parameter σ allows us to separate the stationary patterns at different scales.

In particular, the Aphanitic class can be seen as a stationary texture whose stationary pattern is present at a very small scale (microscope scale). It is found empirically that for a small σ (according to the size of the images a small filter width would correspond to σ in the range [0.8–1.2]) the texture-cartoon separation technique captures the whole image in the texture component, while leaving the cartoon component empty. On the contrary, the cartoon part of the other five remaining classes (phaneritic, porphyry, vein, stockwork and breccia) is not significantly altered. Therefore, focusing just on the cartoon part, the relative change of image borders before and after the separation can be determined as a key feature. It is expected that for the aphanitic images this relative change should be very high, while the other classes would be very small.

4.1.1 Designed Feature

Denoting an image by z and its texture and cartoon parts associated with a filter σ by z^t_σ and z^u_σ, respectively, the quantity of image borders can be quantified by the total variation (TV) operator (Rudin et al. 1992). Specifically, for an image its TV is given by $\| \nabla z \|_{\ell_1}$, where $\| \cdot \|_{\ell_1}$ denotes the ℓ_1-norm and ∇ is the gradient operator. Then, the relative change previously described can be quantified by

$$f_1(z) \equiv \frac{| \, \| \nabla z \|_{\ell_1} - \| \nabla z^u_\sigma \|_{\ell_1} \, |}{\| \nabla z \|_{\ell_1}} \in [0, 1], \quad \forall z \in \mathscr{Z}. \tag{6}$$

To illustrate the trend of this feature, Fig. 4 shows the texture and cartoon components for an aphanitic sample and a stockwork sample, as well as the value of f_1 in both cases.

4.1.2 Detection

Concerning the detection of aphanitic class, f_1 is adopted as the only feature vector to train a collection of SVM classifiers. In Fig. 5, the discriminative ability of f_1 is shown presenting its value for each image in the aphanitic class and for images in the remaining classes. It can be seen how members of the aphanitic class are located at high values close to one while the others have small values close to zero. The presence of these two separated clusters in the feature space gives evidence that decision regions can be established with good discrimination and, consequently, low probability of

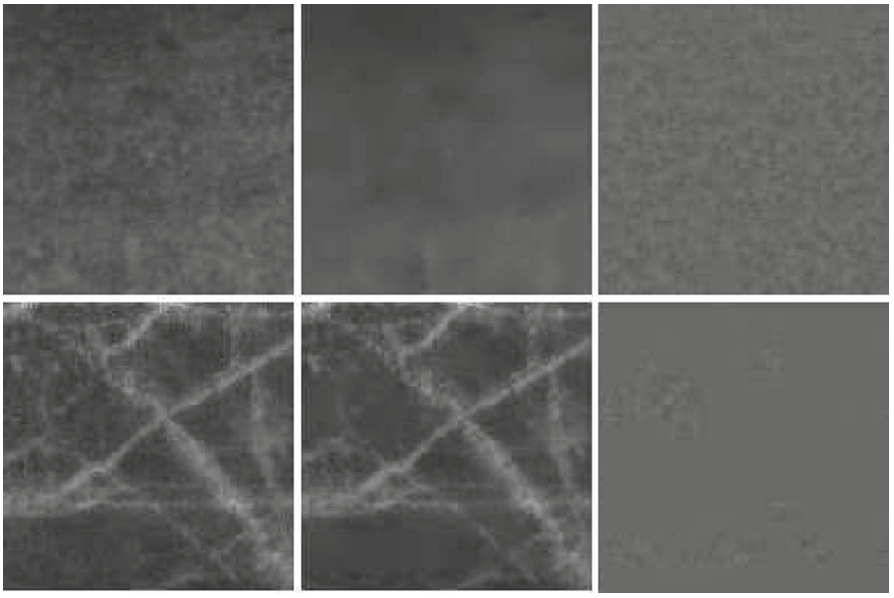


Fig. 4 *Each row* shows an original image and its cartoon and texture components for an aphanitic and a stockwork samples, respectively, setting $\sigma = 1$. In the first case $f_1 = 0.811$ and in the second one $f_1 = 0.134$

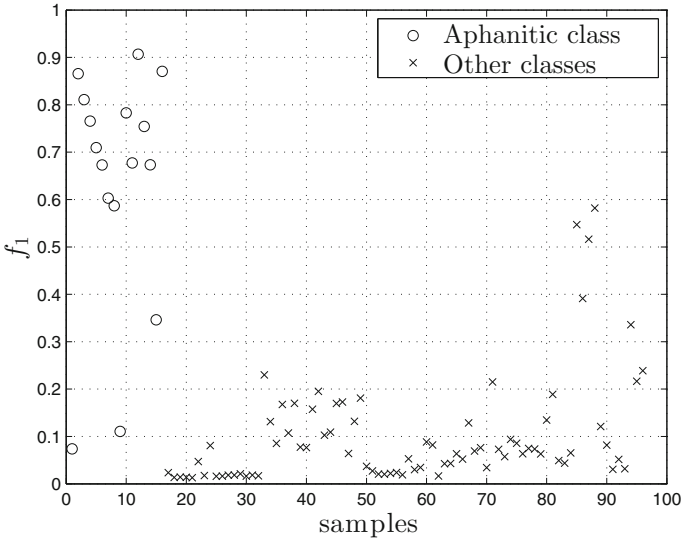


Fig. 5 Discriminative ability of f_1. The x-axis shows different samples. 1–16 correspond to aphanitic samples, 17–96 correspond to samples of the other classes. The y-axis shows the value of f_1 for each sample

error. The width of the filter chosen was $\sigma = 1$, which allows an adequate isolation of the small-scale stationary patterns of the aphanitic texture (Fig. 4). In general, the optimum value of σ should be a function of the scale at which the images of

the database were acquired. In our context using $\sigma = 1$, a considerable change in the border content of our target image class is observed between the original and its cartoon version (after separation).

Equipped with this one-dimensional feature, a family of SVM classifiers is built. As database 16 samples for our target aphanitic class and 80 samples for the remaining 5 classes (the alternative hypothesis) are considered. Three types of kernel types were studied obtaining the best performance with the linear kernel (over the polynomial or radial function basis (RBF) kernels). This fact is consistent with the almost linear separability of the classes in the f_1 feature domain (Fig. 5). For performance, a cross-validation is used to make an efficient use of the data. In particular, a leave-one-out approach is adopted, where each sample is separated and a model is trained with the remaining samples. Next, the trained model is tested over the previously separated sample. For this case, if a sample of the aphanitic class is separated the trained model is obtained using 15 samples of this class and 80 of the other. Conversely, if one sample of the remaining classes is separated, the trained model is obtained using 16 samples of the aphanitic class and 79 of the other. Finally, 93.8 % of accuracy is obtained that is consistent with the distribution illustrated in Fig. 5, and the ability of SVM to obtain a large margin decision boundary. Therefore, in this binary detection, the hypothesis that the aphanitic class has a low-dimensional signature (that can be represented in this case by a scale feature) is validated.

4.2 Phaneritic Class Detection: Second Binary Classifier

The phaneritic class can be considered as a classical texture class because it has stationary patterns. In contrast with the aphanitic class, these patterns occur at several scales (Fig. 1). Importantly for feature design, there exists a specific scale at which just the phaneritic class presents a stationary pattern and the others remaining classes do not. Then, the objective of the feature extraction part of this stage is to take advantage of this observation. For that, we first isolate the stationary pattern using the texture-cartoon separation technique indicated in Sect. 4.1 by fixing a particular σ in the range [5–6], according to the image size. More precisely, a filter size in the range [5–6] assigns a considerable proportion of stationary patterns of the phaneritic samples to the cartoon part. Meanwhile, the cartoon component of the other remaining four classes should not present significant stationary information. Consequently, in this range better discrimination of the phaneritic class is achieved that translates in better classification performance.

4.2.1 Designed Features

To quantify the presence of the particular stationary pattern of the phaneritic class, a transform domain representation is proposed, the wavelet transform (WT) (Mallat 2008) (more details in The Wavelet Transform in Appendix 1). For an image z in the pixel-based domain, the WT corresponds to a decomposition (or projection) of its information at several scales and locations. That decomposition is carried out by an orthonormal basis whose elements, called wavelets, are denoted by $\psi_{j,k}$, where j

and k correspond to a specific scale and location indexes, respectively. Then, after the projection of z onto the family $\{\psi_{j,k}\}$, the WT is the set of wavelet coefficients $\{d_{j,k} = \langle \psi_{j,k}, z \rangle\}$ given by the inner product between z and each wavelet $\psi_{j,k}$. Additionally, by fixing the scale index j, the j-wavelet sub-band corresponds to the wavelet coefficients at all possible locations at scale j.

It is known that the histogram of coefficients in a wavelet sub-band can be modeled by a generalized Gaussian density (GGD) (Do and Vetterli 2002), which is parametrized by a spread and a shape parameter, denoted by α and β, respectively. An interesting fact is that these parameters allow one to infer the form of the probability density function (PDF) associated with a sub-band. A high spread parameter means a wide PDF and a small shape parameter means a fast decay in the tail trend (more details are available in Appendix 2). With this statistical model, we have a probabilistic distribution for each wavelet sub-band of the image. As a result, it is possible to quantify the presence of the isolated stationary pattern for the phaneritic class because when a stationary pattern is present in an image, the PDF of the associated wavelet sub-bands tends to be wide. This means a high value for the spread and the shape parameters. On the other hand, in the rest of the remaining four classes, where there is not such a stationary pattern, the PDF shape is thin, having smaller spread and shape parameters. This distinctive feature is clearly illustrated in Figs. 6, 7 and 8.

Finally, for the feature extraction of the second binary classifier, the WT with three scales is computed, which gives a set of nine spread parameters and nine shape parameters. As a result, the feature vector of the second binary classifier is given by the set of statistical parameters of each of the nine sub-bands

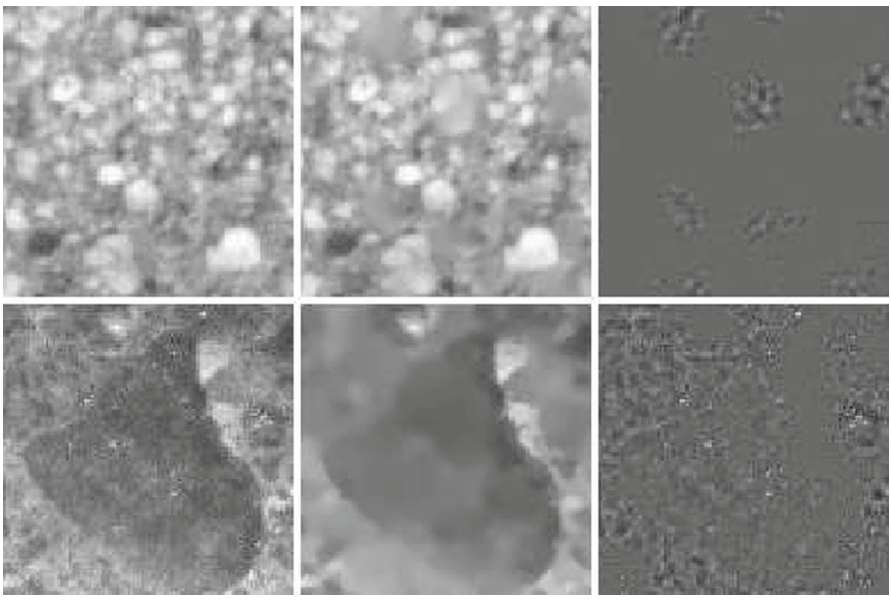


Fig. 6 *Each row* shows an original image and its cartoon and texture components for samples of the phaneritic and breccia classes, respectively, setting $\sigma = 5.5$

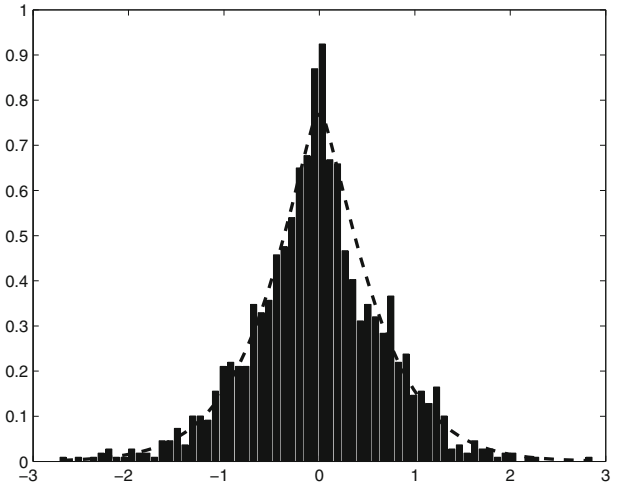


Fig. 7 Histogram and the fitted model ($\alpha = 0.683,\ \beta = 1.274$) of wavelet coefficients in a particular sub-band at the second scale for the cartoon image of a phaneritic sample

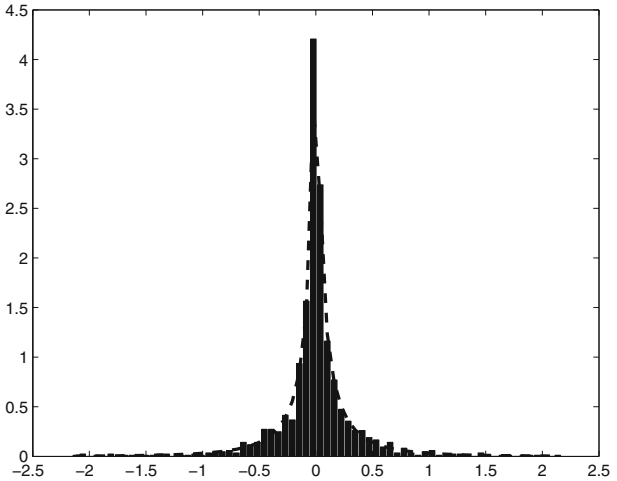


Fig. 8 Histogram and the fitted model ($\alpha = 0.014,\ \beta = 0.421$) of wavelet coefficients in a particular sub-band at the second scale for the cartoon image of a breccia sample

$$f_2 \equiv \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \\ \vdots & \vdots \\ \alpha_9 & \beta_9 \end{pmatrix}. \tag{7}$$

4.2.2 Detection

Concerning the problem of detecting the phaneritic class, Fig. 9 illustrates the discriminative ability of f_2 by showing the parameters of the fourth sub-band in the feature

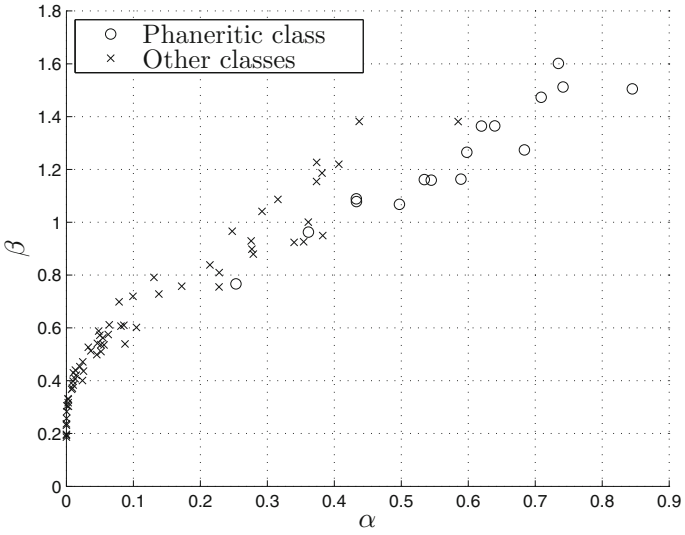


Fig. 9 Discriminative ability of f_2. Considering just one wavelet sub-band, the x-axis shows the value of the spread parameter for 16 samples of the phaneritic class and 64 samples of the other classes. Analogously, the y-axis shows the respective values of the shape parameter

vector, which is associated with the second level of resolution. These parameters are calculated using a wavelet decomposition of three scales after an isolation of the stationary pattern using a filter of width $\sigma = 5.5$. The chosen number of scales permits an appropriate resolution analysis of the images according to their size (128×128 pp.). Higher levels of decomposition of the wavelet were explored but they did not capture any stationary pattern and their contribution was negligible. As well as for the first binary detector, an adequate value for σ depends on the scale at which the images were acquired. Considering the scale used in the acquisition process (all photographs are acquired using the same resolution in the digital camera), it is found that $\sigma = 5.5$ allowed good isolation of the stationary patterns. In fact with this setting, two clusters can be obtained one for the phaneritic class and the other for the remaining four classes (porphyry, vein, stockwork and breccia). Remarkably, the other sub-bands present a similar discriminative trend, so clear decisions regions are expected when the parameters of all sub-bands are gathered in the designed nine-dimensional feature vector in Eq. (7).

Moving to the detection problem, the same family of SVM classifiers used for the first detector in Sect. 4.1 is trained. In this context, 16 samples for the phaneritic class and 64 samples for the alternative hypothesis are available. Again cross validation is adopted and the performance for all choices of kernels was obtained. The SVM with the polynomial kernel offered the best performances with a classification accuracy of 92.5 %. This high accuracy validates the feature design strategy developed in this section and the assumption of the existence of a discriminative signature for this particular class.

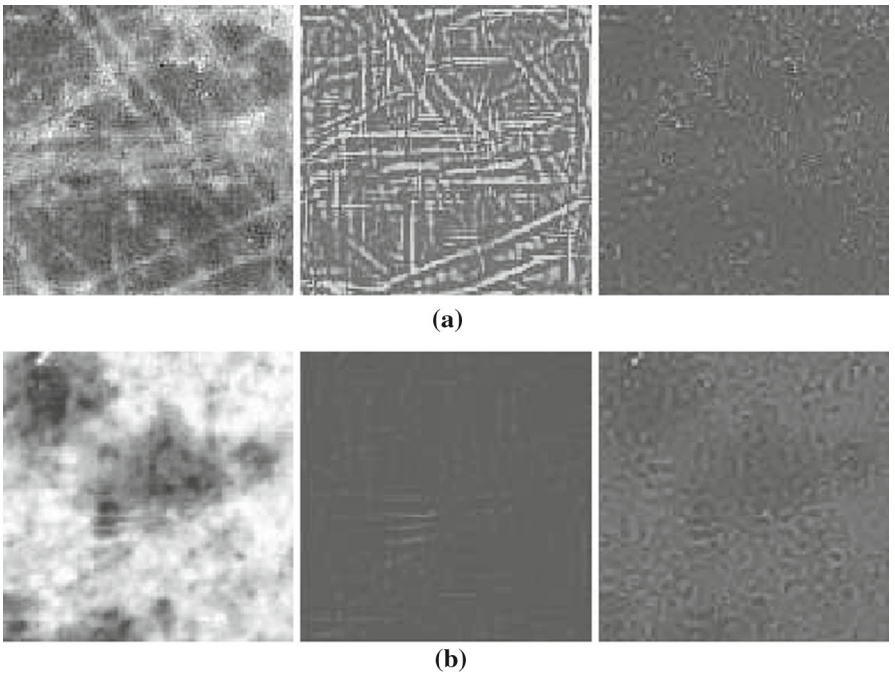


Fig. 10 **a** Sample of the stockwork class and its component of geometric forms and its component of stationary patterns, respectively ($f_3 = 599.53$). **b** Sample of the porphyry class and its component of geometric forms and its component of stationary patterns, respectively ($f_3 = 312.27$)

4.3 Stockwork Class Detection: Third Binary Classifier

The third binary classifier is intended to detect the stockwork class. This class can be considered as a structural texture. As a result, it is relevant to extract information associated with the presence of a geometric form, which is an array of veins of a sheetlike form, illustrated in Fig. 10a. In addition, the images of this class can present a stationary pattern in some parts, which hinders the analysis of the mentioned geometric forms. For that reason, the stationary patterns from the geometric forms need to be isolated. To do this, a source separation technique is used, which is based on representing the image in two alternative transform domains (Starck et al. 2005; Fadili et al. 2010; Kutyniok and Lim 2012).

4.3.1 Designed Feature

We model a stockwork image z by the sum of two independent sources

$$z = z_{\mathrm{s}} + z_{\mathrm{g}}, \tag{8}$$

where z_{s} corresponds to the stationary patterns (texture), and z_{g} corresponds to the geometric forms (veins). With this model, the source separation technique uses the fact

that these two independent sources are compressible in the wavelet and the shearlet domains, respectively (Appendix 4; Kutyniok and Lim 2012). Figure 10 shows the result of a separation using the adopted source separation technique, where it can be seen that z_g captures the geometric forms properly.

After the source separation is conducted, the ST is applied directly to z_g (see details on The Shearlet Transform in Appendix 1). If we fix a given scale j, and we compute the transform coefficient for each possible location, it is possible to calculate in which orientation the associated Shearlet coefficient achieves its maximum (Gibert et al. 2014). Next, all the maximum coefficients found can be added into one single value, which is called orientation energy (OE), and is denoted by ξ_j. More formally, considering the shearlet coefficients of z_g, denoted by $\{\mathcal{SH}_{j,l,k}(z_g)\}$, where j, l, and k are the scale, orientation and location parameters, respectively, for a pixel $k = (k_1, k_2)$ and a scale j, it is possible to define

$$\xi_{j,k}(l) \equiv |\mathcal{SH}_{j,l,k}(z_g)|, \quad l \in L_j, \tag{9}$$

where L_j corresponds to the orientations associated with the scale j (more details in The Shearlet Transform in Appendix 1). Then, the OE of z at scale j corresponds to

$$\xi_j(z) \equiv \sum_{k \in \mathbb{Z}^2} \max_{l \in L_j} \xi_{j,k}(l). \tag{10}$$

It happens that ξ_j tends to be high for the stockwork class, and small for the remaining three classes. Specifically, for our database just the OE of the second scale will be used, as it provides enough discrimination for the problem (Fig. 11). Thus, the feature used by the third binary classifier is given by

$$f_3 \equiv \xi_2. \tag{11}$$

4.3.2 Detection

In Fig. 11, the discriminative ability of f_3 is illustrated by showing the OE for each sample of the stockwork class and the remaining three classes (vein, porphyry and breccia). It should be mentioned that for the Shearlet transform only three levels of scale were considered. This is because higher decomposition levels do not show adequately the structures of the stockwork samples due to the size of the images (128×128). Furthermore, the use of the second scale to calculate the OE is due to the fact that the structures (sheetlike veins) of the stockwork class are well represented at this level of resolution, which translates in high value for the OE. This can be observed in Fig. 11, where samples of the stockwork class are located at high values and the other three classes show small values, providing a clear decision boundary in the feature space. To implement that decision region (or classifier), a family of SVM classifiers is trained with different kernel functions (linear, polynomial, and RBF). For training, 16 samples of the stockwork class and 48 samples of the remaining three classes are considered. For testing, cross-validation strategy was used as explained in Sect. 4.1.2.

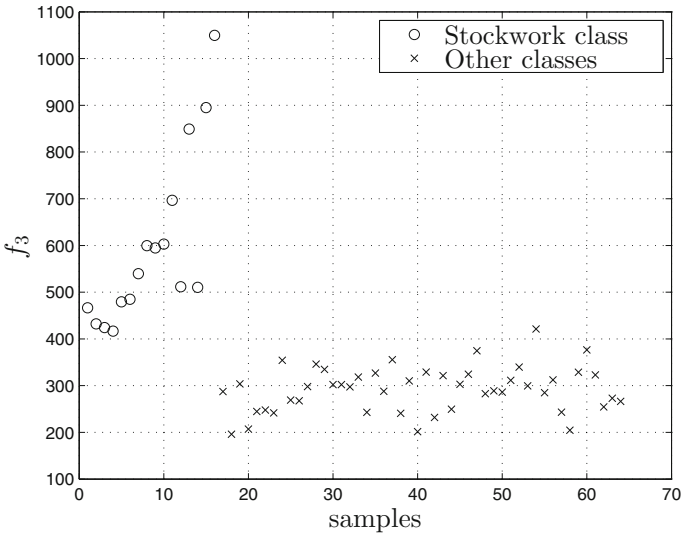


Fig. 11 Discriminative ability of f_3. The x-axis shows different samples. From 1–16 correspond to stockwork samples, and from 17–64 to samples of other classes. The y-axis shows the value of f_3 for each sample

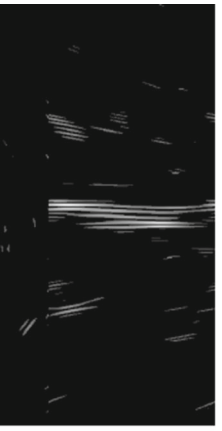
From the distribution of the feature presented in Fig. 11, a simple affine decision boundary is observed. Consistent with this observation, a linear SVM classifier shows the best performance for this task with 96.9 % of accuracy (correct classification rate). This performance supports the discrimination ability of f_3 for the accurate detection of the stockwork class.

4.4 Vein Class Detection: Fourth Binary Classifier

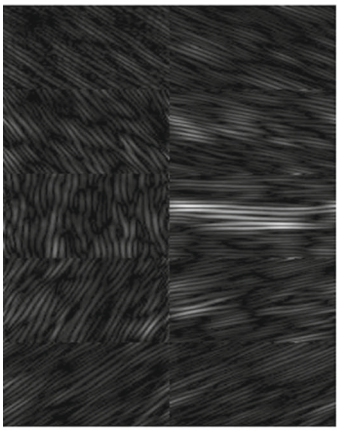
The fourth stage is designed to detect the vein class. Images of this class are characterized by the presence of a considerable thick vein, an object with a sheetlike form following an orientation. Figure 12a shows a sample of the class and its respective vein with a vertical orientation.

A feature extraction methodology is proposed with the purpose of quantifying the presence of this signature and differentiate it from the objects inside images of the other two remaining classes (breccia and porphyry). In fact, the other two classes are composed of clasts and porphyds (crystals), which have a more isotropic form than veins. Therefore, it is that isotropy difference the mean to discriminate the vein class from breccia and porphyry. For that purpose, the ST is adopted (The Shearlet Transform in Appendix 1).

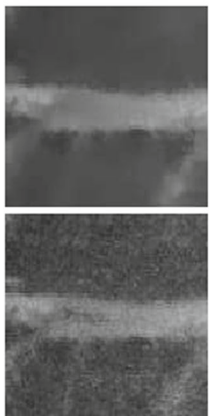
Yi et al. (2009) showed that a small proportion of the highest shearlet coefficients corresponds to the borders of the image. Therefore, if the image has an object with a particular orientation, the shearlet sub-band associated with that orientation should have large coefficients (located at the borders of the object). On the other hand, in the sub-band with orientation orthogonal to the object orientation, there should not



(c)



(b)



(a)

Fig. 12 **a–c** Sample of the vein class with its cartoon part ($\sigma = 8$), and its shearlet coefficients at scale $j = 2$ with ten orientations before and after considering 3 % of the most significant coefficients ($f_4 = 0.0011$)

be large coefficients. Figures 12, 13 and 14 illustrate this observation. Based on this, the sum of the absolute value of the coefficients of each sub-band is computed (or the ℓ_1-norm of the sub-band), which is known as the energy of the sub-band. The maximum of these computed energies should be associated with the sub-band with the same orientation of the object. Finally, the ratio between the energy of the sub-band orthogonal to the object orientation and the energy of the sub-band with the same orientation is computed. This ratio should be close to zero for our target class. Conversely, for objects with an isotropic form, the energy of each sub-band should be similar since the object has borders in each direction, and the ratio between any sub-band an its orthogonal counterpart should be close to one.

4.4.1 Designed Feature

More formally, given an image z, its Shearlet coefficients $\{\mathcal{SH}_{j,l,k}(z)\}$, and a percentage value $p \in (0, 1)$, we define

$$\mathcal{SH}^p(z) \equiv \{\mathbb{1}_{\mathscr{P}_p}(\mathcal{SH}_{j,l,k}(z)) \cdot \mathcal{SH}_{j,l,k}(z),\ j \leq 0,\ l \in L_j,\ k \in \mathbb{Z}^2\}, \quad (12)$$

where $\mathscr{P}_p$ is the set of the p-percentage most significant shearlet coefficients of z, and

$$\mathbb{1}_{\mathscr{P}_p}(\mathcal{SH}_{j,l,k}(z)) = \begin{cases} 1, & \mathcal{SH}_{j,l,k}(z) \in \mathscr{P}_p, \\ 0, & \mathcal{SH}_{j,l,k}(z) \notin \mathscr{P}_p. \end{cases} \quad (13)$$

Therefore, $\mathcal{SH}^p(z)$ corresponds to the shearlet coefficients where the $1 - p$ less significant fraction is set to zero. We denote the elements of this set by $\mathcal{SH}^p_{j,l,k}(z) \equiv \mathbb{1}_{\mathscr{P}_p}(\mathcal{SH}_{j,l,k}(z))\mathcal{SH}_{j,l,k}(z)$, and the sub-band corresponding to the scale parameter j and the orientation parameter l by $\mathcal{SH}^p_{j,l}(z)$. Next, for a scale j, considering its associated orientations $l \in L_j$, we define the set $\mathscr{K}^p_j \equiv \{\kappa^p_{j,l} \in [0, 1],\ l \in L_j\}$, where

$$\kappa^p_{j,l} \equiv \left(\frac{\min(\| \mathcal{SH}^p_{j,l}(z) \|_{\ell_1}, \| \mathcal{SH}^p_{j,\bar{l}}(z) \|_{\ell_1})}{\max(\| \mathcal{SH}^p_{j,l}(z) \|_{\ell_1}, \| \mathcal{SH}^p_{j,\bar{l}}(z) \|_{\ell_1})} \right)^2, \quad (14)$$

with $\bar{l} \in L_j$ being the closest orientation to the orthogonal orientation of l. With this definition, we define

$$\kappa^{p*}_j \equiv \kappa^p_{j,l^*}, \quad \text{where} \quad (15)$$

$$l^* \equiv \arg\max_{l \in L_j} \| \mathcal{SH}^p_{j,l}(z) \|_{\ell_1}. \quad (16)$$

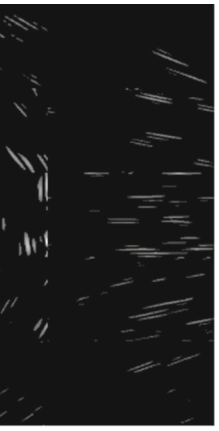
Finally, our designed feature is given by

$$f_4 \equiv \kappa^{p*}_j, \quad (17)$$

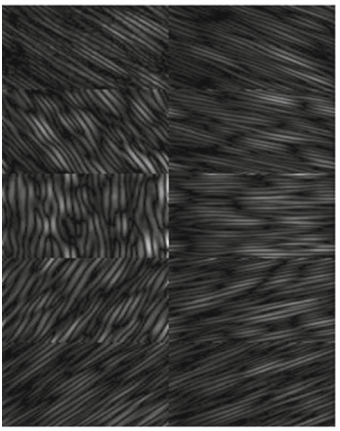
where $j = 2$ and $p = 0.03$.

Figures 12, 13 and 14 illustrate how high magnitude Shearlet coefficients are located at the borders of images, and how f_4 varies as the isotropy of objects increases. It should

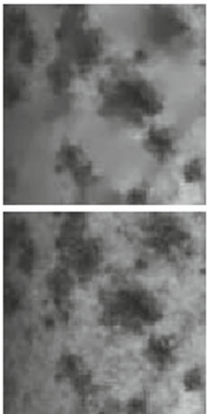
Fig. 13 **a–c** Sample of the breccia class with its cartoon part ($\sigma = 8$), and its shearlet coefficients at scale $j = 2$ with ten orientations before and after considering 3 % of the most significant coefficients ($f_4 = 0.7468$)



(c)



(b)



(a)

Fig. 14 **a–c** Sample of the porphyry class with its cartoon part ($\sigma = 8$), and its shearlet coefficients at scale $j = 2$ with ten orientations before and after considering 3 % of the most significant coefficients ($f_4 = 0.7024$)

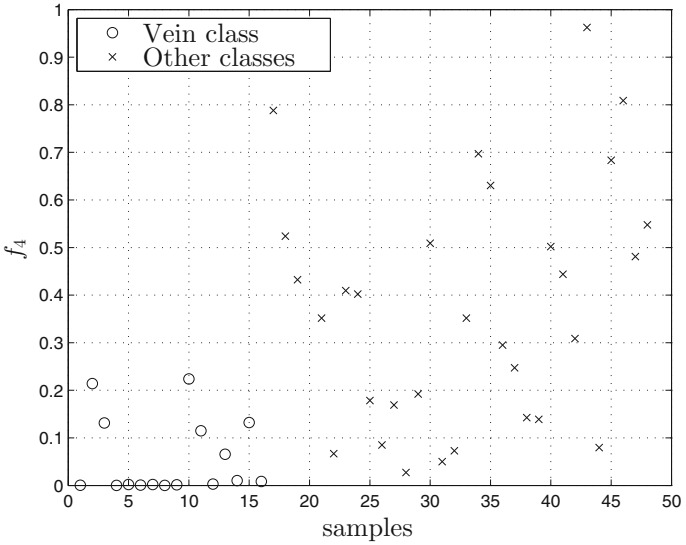


Fig. 15 Discriminative ability of f_4. The x-axis shows different samples. 1–16 correspond to vein samples, and 17–48 to samples of other classes. The y-axis shows the value of f_4 for each sample

be mentioned that to focus the analysis on the signature objects of the class, a texture-cartoon separation was done using $\sigma = 8$ before applying the ST (see Appendix 3). As well as for the first and second classifiers, the optimal value of σ depends on the scale at which the images were originally acquired. In our context after a systematic experimental analysis, we found that smaller values ($\sigma \ll 8$) do not isolate properly the objects, and higher values ($\sigma \gg 8$) blur the borders of the main vein in the vein class, which has a negative effect in terms of classification performance.

The decision of using the second scale of the shearlet decomposition (i.e., $j = 2$) relies on the fact the vein object is well represented at this level of resolution. Other levels of scale did not provide additional complementary information to discriminate our target signature (in the sense of quantifying the anisotropy of the target object). In fact, the effect of considering the isotropy coefficients of other levels in the feature vector did not translate in an improvement in the classification performance.

Finally, the percentage of coefficients p was chosen small enough to exclusively detect the borders of the image (Figs. 12, 13, 14).

4.4.2 Detection

The discriminative ability of f_4 is illustrated in Fig. 15. In general, the discrimination power of this feature is clear in the sense that vein samples are concentrated at low values (in the range [0, 0.2]) of the feature space while samples from breccia and porphyry are located in a wider range from 0.1 to 1. For a significant portion of the samples, no overlap is observed between the two classes. However, a zero error decision region is not possible to establish. The reason is that some images of breccia and porphyry present some anisotropic objects that translate in some overlap with vein

samples in the feature space. Concerning classification performances, the same family of SVM classifiers is trained with 16 samples of the vein class and 32 samples of the two alternative classes. For testing, cross validation was used as in Sect. 4.1.2, and again the minimum empirical risk principle was adopted for selecting the best SVM kernel. The polynomial kernel showed the best classification rate with a 79.2 % of accuracy. This smaller rate, compared with previous stages, is reflected in the distribution presented in Fig. 15.

4.5 Porphyry and breccia Classification: Fifth Binary Classifier

The fifth binary classifier is intended to discriminate the porphyry class from the breccia class. In the first case, the porphyry class corresponds to a stationary texture which presents stationary patterns of small and big scale. In the second case, the breccia class corresponds to a structural texture which is characterized by the presence of big objects called clasts. The objective here is to extract and quantify the information given by the small-scale stationary patterns present in the porphyry class, which are barely present in the breccia class. For that purpose, the local binary patterns (LBPs) are used (Ojala et al. 2002). LBP associates a binary pattern to each pixel of an image. Specifically, for each pixel its value is compared with the value of the other pixels in a neighborhood considering a certain profile. This process creates a binary vector. The entries in the vector depend on whether the compared pixel was higher or smaller than the neighbor pixel. Finally, a histogram is computed with the frequency of the binary patterns in the image.

The low scale stationary patterns present in the porphyry class create elevated zones in the LBP histogram (Fig. 16). On the other hand, the absence of stationary patterns in the breccia class results in a flat LBP histogram. The feature of the fifth binary classifier is designed to capture this difference. To do so, the kurtosis of the LBP histogram is proposed as feature, since this value is associated with the histogram flatness.

4.5.1 Designed Feature

More precisely, given an image and its LBP histogram denoted by h, f_5 is given by

$$f_5 = \frac{\frac{1}{n}\sum_{i=1}^{n}(h_i - \overline{h})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(h_i - \overline{h})^2\right)^2}, \tag{18}$$

where n is the number of bins in the histogram, h_i is the histogram value at i-bin, and $\overline{h}$ is the histogram mean. Note that there are no parameters to be chosen in f_5.

4.5.2 Detection

Concerning the classification of this final stage, Fig. 17 shows the values of f_5 for samples of the two classes porphyry and breccia. Remarkably, two clear cluster in this one-dimensional feature space can be observed, where the samples of the porphyry class are concentrated in the range [0.25, 12] of the feature space, while samples of

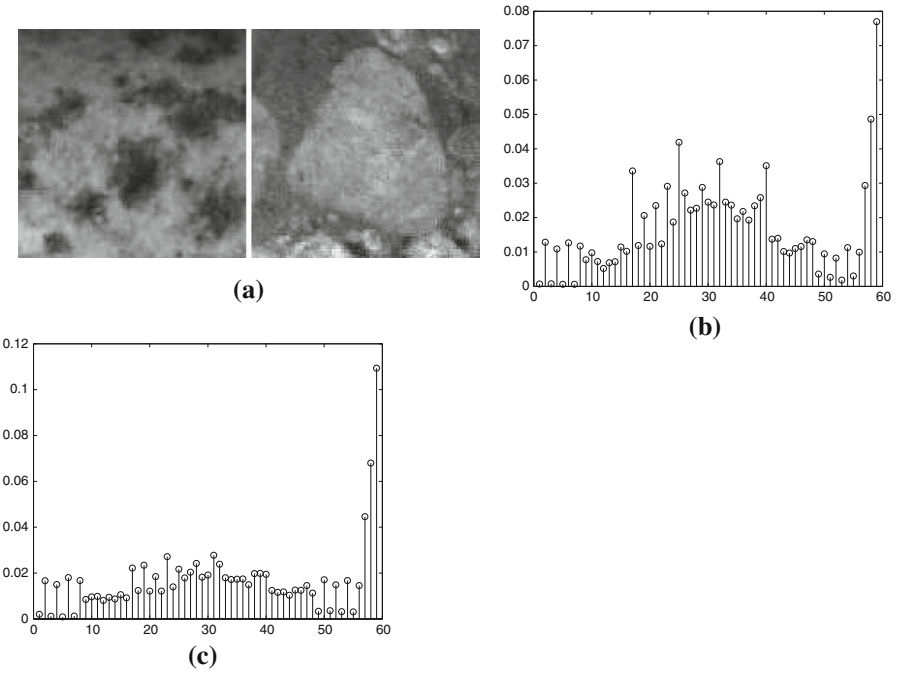


Fig. 16 **a** An example of porphyry (*left*) and breccia (*right*) classes, respectively. **b** The LBP histogram of the porphyry image ($f_5 = 8.465$). **c** The LBP histogram of breccia image ($f_5 = 21.572$)

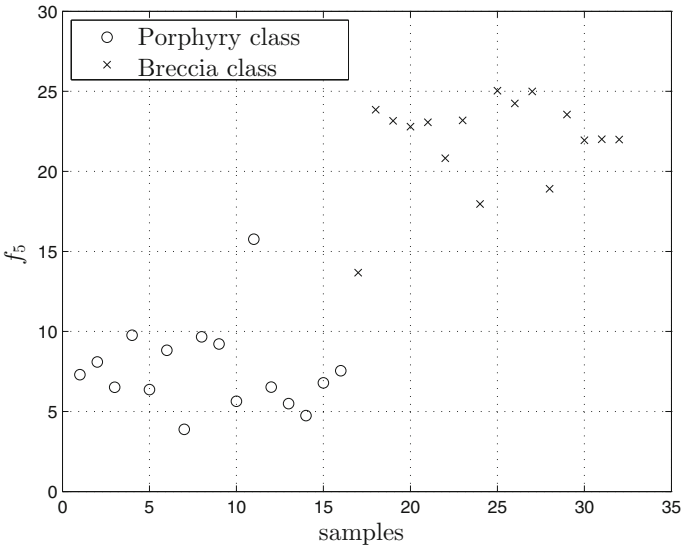


Fig. 17 Discriminative ability of f_5. The x-axis shows different samples. 1–16 correspond to porphyry samples, and 17–32 to breccia samples. The y-axis shows the value of f_5 for each sample

the breccia class are concentrated in a non-overlapping range [18, 26] of the feature space. The presence of these two well-established clusters allowed us to build an SVM classifier with good classification performance. For that, our family of SVM classifiers is trained in this case with 16 samples of the breccia class and 16 samples of the porphyry class. Cross validation is adopted as before, and the minimum empirical risk rule is used to select the best kernel. As expected from Fig. 17, where a linear boundary shows an almost perfect decision, the best performance was obtained by the linear kernel achieving 93.2 % of classification accuracy. This evidence supports the hypothesis of this work that a small dimensional feature is able to capture a signature that discriminates porphyry from breccia in this last decision stage.

5 Global Classification Performance

This section analyzes the classification performance of the chain sequential classifier (CSC) presented in Sect. 3. To compute the performances of this strategy, it is instrumental to know the performance of any of the partial decisions of the chain. In fact the performance of the CSC is a function of the partial performances reported and analyzed in Sect. 4.

From the binary sequential approach in Eq. (5), its probability of error is given by

$$\begin{aligned} P(g_{\text{chain}}(X) \neq Y) &= 1 - P(g_{\text{chain}}(X) = Y), \\ &= 1 - \sum_{l=1}^{6} P(g_{\text{chain}}(X) = Y | Y = l) P(Y = l). \end{aligned} \tag{19}$$

From Eq. (19), the probability of error reduces to calculating the conditional probabilities $\{P(g_{\text{chain}}(X) = Y|Y = l)\}_{l=1}^{6}$, which involve the knowledge of the success probability for each binary detector $\{P(g_i(X_i) = Y_i)\}_{i=1}^{5}$ in Eq. (3). More precisely, from the structure in Fig. 3, we have that

$$P(g_{\text{chain}}(X) = Y|Y = l) = \prod_{i=1}^{l} P(g_i(X_i) = Y_i), \tag{20}$$

where the random binary vector $(Y_1, ..Y_5) \in \{0, 1\}^5$ in Eq. (20) denotes the binary coding of the true class Y, using for that coding the path structure illustrated in Fig. 2. Finally, since the same number of images per class is available in our database, $P(Y = l) = 1/6, \ \forall l \in \{1, \ldots, 6\}$ is considered. Finally, replacing Eq. (20) in Eq. (19), it follows that

$$P(g_{\text{chain}}(X) \neq Y) = 1 - \frac{1}{6} \cdot \sum_{l=1}^{6} \left(\prod_{i=1}^{l} P(g_i(X_i) = Y_i) \right). \tag{21}$$

5.1 Analysis and Results

Equation (21) is used to determine the performance of our method globally. Therefore, the set of probabilities $\{P(g_i(X_i) = Y_i)\}_{i=1}^{5}$ of every binary detector, reported in Sect. 4, is used. As a summary of these results, Table 1 reports the performance of each binary classifier, where its success probability and the selected SVM kernel are listed. Here, it is possible to compare the quality of features designed for each of the stages in Sect. 4, where the performance for each binary detector is very good in general supporting the main assumption considered in this work that every class has a discriminative signature that is reflected in a small dimension feature vector that discriminates that particular class from the rest. The only exception of this rule is g_4, associated with the detection of the vein class where the designed feature shows lower accuracy. This can be attributed to the complexity of discriminating the vein class from the breccia and porphyry classes. In fact, from the way the feature extraction was designed for this detector (detecting if the objects inside the images have a particular anisotropic orientation), it happens that there are some images from the breccia class that contains structures with a certain degree of anisotropy, which causes confusions. This can be observed in Fig. 15.

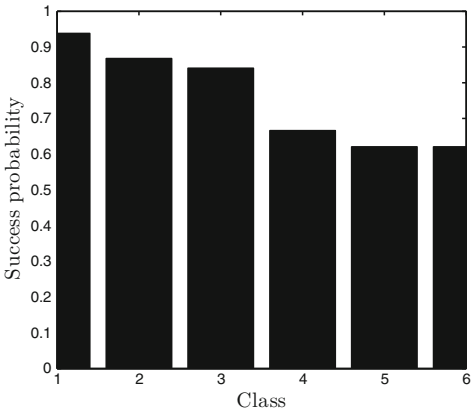
Finally, to evaluate the performances of our designed small dimensional discriminative features in the context of the proposed sequential strategy driven by SVM detectors, we consider two widely adopted and recognized state-of-the-art methods used to discriminate texture in natural images (Do and Vetterli 2002; Ojala et al. 2002), based on WT and LBP, respectively. Table 2 presents the probability of error of our scheme g_{chain} and the other two classifiers. Remarkably, the proposed method shows an important reduction in classification errors when compared with the methods based on WT or LBP. In particular, our method offers a reduction in error classification (a significant 32.1 % relative improvement) with respect to the wavelet-based method which is the most competitive for this task.

Table 1 Success probability and SVM kernel of each binary classifier

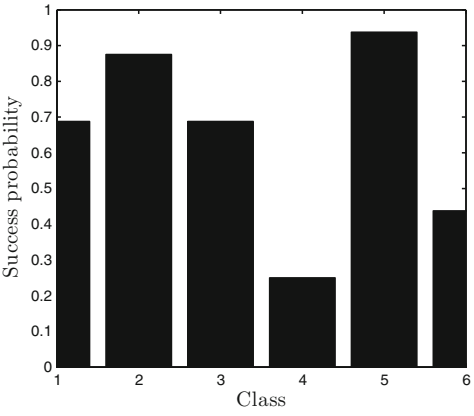
Binary classifier	Success probability ($P(g_i(X_i) = Y_i)$)	SVM configuration
g_1	0.938	Linear kernel
g_2	0.925	Polynomial kernel
g_3	0.969	Linear kernel
g_4	0.792	Polynomial kernel
g_5	0.932	Linear kernel

Table 2 Error probability of each method

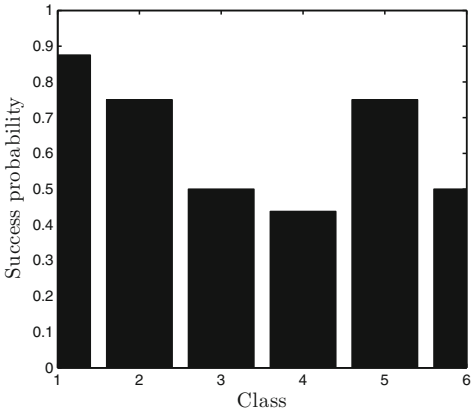
Method	Probability of error
CSC	0.241
LBP-based method	0.365
Wavelet-based method	0.355



(a)



(b)



(c)

Fig. 18 **a** Success probability of the CSC for each class. **b** Success probability of the wavelet-based method for each class. **c** Success probability of the LBP-based method for each class. The corresponding labels are: 1 aphanitic; 2 phaneritic; 3 stockwork; 4 vein; 5 porphyry; 6 breccia

To complement this analysis, Fig. 18 shows the success probability per class, where as expected the two natural-texture methods have good performances for stationary texture classes (aphanitic, phaneritic and porphyry). However for the structural textures (stockwork, vein and breccia), where there are a combination of stationary patterns (at a given scale) and objects, the performance is affected. This verifies the motivation of this work, in which conventional natural-texture discriminative methods have problems detecting objects with the presence of both texture patterns and cartoon components. In contrast, our approach based on the use of features that capture both texture and object as information for discrimination achieves a more uniform success probability among the classes, and, more importantly, better performance overall.

6 Conclusions

This work proposes a new set of discriminative features for detecting natural rock texture based on the extraction of texture information (related to the matrix of the rock) and object information (related to the presence of large objects within the matrix) in transform-based domains. In addition, a chain sequential classification approach is proposed, that, on the one hand, facilitates the design of isolated features, and on the other, provides a way to address the practical issue of limited data.

The proposed classification scheme has the ability to discriminate stationary textures from structural textures, by means of a separation into texture and cartoon components. Because of this ability, the proposed framework offers very good performance for the problem of automatic rock classification for the six classes addressed, showing significant improvements with respect to well-known state-of-the-art techniques used conventionally for texture classification (indexing) of natural images. These findings stress how relevant is for natural rock image classification to develop specific features for this task.

It is important to emphasize that the assumption that it is possible to design a set of small dimensional discriminative features for every rock class is validated in this work. A concrete methodology to extract that discriminative information from the image is proposed for every stage of the chain illustrated in Fig. 3. The existence of this set of small dimension and high discriminative features is the main justification that supports the fact that the sequential classification strategy (using SVM classifiers) is performing better than state-of-the-art methods. SVM has the ability to learn (in a non-parametric fashion) large margin decision regions with limited supervised samples from each class. Then, the features proposed in this work facilitate and make this simple classification strategy possible.

On the direction of adopting more complex classifiers, much work can be done to address the classification problem using a global strategy, instead of the sequential approach proposed here. In theory, a global strategy is optimal when the joint distribution is known between the features and the class label (Sect. 3.1). This was not feasible in our problem considering that very limited data were available. However, the sequential approach shows to be an excellent alternative under the condition that the classification problem can be addressed in a small dimensional feature space in each stage of the chain. To move from the sequential to the joint strategy, more data

are needed as the natural way to address this problem is to construct a large feature space (the union of the features designed for each detector can be an initial alternative) to address the decision problem from there. This work offers concrete directions to construct this large features space, and we believe that equipped with more data improvements and better performances can be expected.

Acknowledgments This material is based on the work supported by grants of Conicyt-Chile, Fondecyt 1140840, and by Advanced Mining Technology Center (AMTC). In addition, the work of J. F. Silva is supported by the Advanced Center for Electrical and Electronic Engineering (AC3E), Basal Project FB0008. The authors wish to express their deepest gratitude to the anonymous reviewers for their rigorous and constructive comments and to Andrea Vidal for proofreading part of this material.

Appendix 1: The Wavelet and Shearlet Transforms

In this section, two mathematical tools are presented: the wavelet transform (WT) and the shearlet transform (ST), which allow to analyze an image at different scales, locations and orientations. A brief mathematical background is introduced showing only the continuous case for the sake of brevity (Mallat 2008; Easley et al. 2008).

The Wavelet Transform

The WT corresponds to a decomposition of a one-dimensional function f by a family of functions $\{\psi_{j,k}\}_{(j,k)\in\mathbb{Z}^2}$, which form an orthonormal basis of $\mathbf{L}^2(\mathbb{R})$, and provide a partition of $\mathbf{L}^2(\mathbb{R})$ in several scales (Mallat 1989, 2008). These functions are obtained through the dilation and translation of a kernel function ψ, called mother wavelet, as follows:

$$\psi_{j,k}(x) = 2^{-j/2}\psi(2^{-j}x - k), \quad (j,k) \in \mathbb{Z}^2. \tag{22}$$

The mother wavelet is constructed from a scaling function ϕ which satisfies the two-scale difference equation (Strang 1989). The construction is given by

$$\psi(x) = \sqrt{2}\sum_{n\in\mathbb{Z}} g(n)\phi(2x - n), \tag{23}$$

where $g(n) = (-1)^n h(1-n)$ and h is a filter that must meet several conditions to obtain an adequate family $\{\psi_{j,k}\}_{(j,k)\in\mathbb{Z}^2}$ (Strang 1989). Thus, the decomposition of a function f is given by a set of wavelet coefficients $\{d_j[k] \equiv \langle f, \psi_{j,k}\rangle = \int_{-\infty}^{+\infty} f(x)\psi_{j,k}(x)\mathrm{d}x : (j,k) \in \mathbb{Z}^2\}$, where each coefficient captures the information that the function has between two subsequent resolution levels at a particular position. Fortunately, to calculate these coefficients it is not necessary to know $\psi_{j,k}$ explicitly, and it is enough to have access to the filter h. In fact, given a scale j it is possible to calculate the coefficients of the next scale $j+1$ by

$$d_{j+1}[k] = \sum_{n=-\infty}^{+\infty} g[n-2k]a_j[n], \tag{24}$$

where

$$a_{j+1}[k] = \sum_{n=-\infty}^{+\infty} h[n-2k]a_j[n]. \tag{25}$$

The coefficients $\{a_j[k] \equiv \langle f, \phi_{j,k}\rangle = \int_{-\infty}^{+\infty} f(x)2^{-j/2}\phi(2^{-j}x-k)\mathrm{d}x,\ (j,k) \in \mathbb{Z}^2\}$ are known as approximate coefficients, and capture the content that a function has at a particular scale level.

In the case of a two-dimensional function $f(x)$ (like an image), where $x = (x_1, x_2)$, to keep an orthonormal decomposition there are three kinds of wavelet coefficients given by

$$d^1_{j+1}[k] = \sum_{(n_1,n_2)\in\mathbb{Z}^2} h_{HH}(n_1-2k_1, n_2-2k_2)a_j[n_1,n_2], \tag{26}$$

$$d^2_{j+1}[k] = \sum_{(n_1,n_2)\in\mathbb{Z}^2} h_{HL}(n_1-2k_1, n_2-2k_2)a_j[n_1,n_2], \tag{27}$$

$$d^3_{j+1}[k] = \sum_{(n_1,n_2)\in\mathbb{Z}^2} h_{LH}(n_1-2k_1, n_2-2k_2)a_j[n_1,n_2], \tag{28}$$

where

$$a_{j+1}[k] = \sum_{(n_1,n_2)\in\mathbb{Z}^2} h_{LL}(n_1-2k_1, n_2-2k_2)a_j[n_1,n_2], \tag{29}$$

$k = (k_1, k_2) \in \mathbb{Z}^2$, $h_{HH}(k) = g(k_1)g(k_2)$, $h_{HL}(k) = g(k_1)h(k_2)$, $h_{LH}(k) = h(k_1)g(k_2)$, and $h_{LL}(k) = h(k_1)h(k_2)$. Furthermore, the approximate coefficients are given by $\{a_j[k] = \int_{-\infty}^{+\infty} f(x)2^{-j/2}\phi^2(2^{-j}x-k)\mathrm{d}x, (j,k) \in \mathbb{Z}^3\}$. It is important to mention that to use the algorithm given by Eqs. (26)–(29), the initial approximate coefficients $a_0[\cdot]$ are needed. In our problem, these coefficients are related to the resolution of the camera, and are given by a fraction of the digital image values (Mallat 2008). Finally, just considering the parameter $i \in \{1, 2, 3\}$ and the scale parameter j, the concept of wavelet sub-band is introduced corresponding to the set $W^{(i)}_j(f) = \{d^{(i)}_j[k],\ k \in \mathbb{Z}^2\}$.

The Shearlet Transform

The ST corresponds to a two-dimensional function decomposition by analyzing functions, the so-called shearlets, which are waveforms indexed by scales, orientations, and locations (Kutyniok et al. 2012; Easley et al. 2008).

The shearlet functions correspond to the set $\{\psi^{(d)}_{j,l,k} :\ d = 0, 1,\ j \geq 0,\ -2^j \leq l \leq 2^j - 1,\ k \in \mathbb{Z}^2\}$, where each element is obtained by the action of a family of operators on a single function ψ defined by the sum of two functions $\psi^{(1)}$ and $\psi^{(2)}$, whose frequency supports are restricted to $\mathscr{D}_0 = \{(w_1, w_2) : |w_1| \geq 1/8, |w_2/w_1| \leq 1\}$ and $\mathscr{D}_1 = \{(w_1, w_2) : |w_2| \geq 1/8, |w_1/w_2| \leq 1\}$, respectively (Easley et al. 2008).

The operations used on ψ correspond to: scaling operators, associated with matrices A_0 and A_1 called anisotropic dilation matrices; shearing operators, associated with

matrices B_0 and B_1 called shearing matrices; and the translation operator. Thus, a shearlet at scale j, orientation l, and location k is given by

$$\psi^{(d)}_{j,l,k}(x) = 2^{\frac{3j}{2}}\psi^{(d)}(B_d^l A_d^j x - k), \tag{30}$$

where $A_0 = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$, $A_1 = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$, $B_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, and $B_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. One should note the connection between shearlets and wavelets as both are derived from analogous processes, and both give a complete description of $\mathbf{L}^2(\mathbb{R}^2)$ in the sense that wavelets form an orthonormal basis and shearlets form a parseval frame (Easley et al. 2008). Nevertheless, shearlets allow a more refined analysis since they consider the orientation index (l).

The shearlet coefficients for a two-dimensional function f are given by

$$\{\mathcal{SH}^{(d)}_{j,l,k}(f) = \langle f, \psi^{(d)}_{j,l,k}\rangle : \ d = 0, 1, \ j \geq 0, \ -2^j \leq l \leq 2^j - 1, \ k \in \mathbb{Z}^2\}, \tag{31}$$

where $\langle\cdot\rangle$ denotes the inner product in $\mathbf{L}^2(\mathbb{R}^2)$. The different orientations are subject to which part of the support of ψ is being considered, which is associated with the parameter d. For this reason, it is convenient for the notation to eliminate this parameter by re-labeling the orientation parameter in the shearlet coefficient expression (Yi et al. 2009)

$$\mathcal{SH}_{j,l,k}(f) = \begin{cases} \mathcal{SH}^{(d)}_{j,(l-1-2^j),k}(f), & 1 \leq l \leq 2^{j+1}, \\ \mathcal{SH}^{(d)}_{j,(3\cdot 2^j - l),k}(f), & 2^{j-1} < l \leq 2^{j+2}, \end{cases}$$

then $l \in L_j = \{1, \ldots, 2^{j+2}\}$. Finally, considering just the scale and orientation parameters, the concept of shearlet sub-band is introduced corresponding to the set $\mathcal{SH}_{j,l}(f) = \{\mathcal{SH}_{j,l,k}(f), \ k \in \mathbb{Z}^2\}$ (Kutyniok et al. 2012; Easley et al. 2008).

Appendix 2: Wavelet Statistical Texture Modeling

Do and Vetterli (2002) showed that a good statistical model for a natural texture is that the wavelet coefficients for a given sub-band follow an i.i.d. distribution. More precisely, a suitable PDF for a specific wavelet sub-band $W_j^{(i)}(z)$ is given by the generalized Gaussian density (GGD)

$$P(W_j^{(i)}(z) = w) = p(w; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \mathrm{e}^{-(|w|/\alpha)^\beta}, \tag{32}$$

where $\Gamma(\cdot)$ is the Gamma function, and α as well as β are real parameters that describe the form of the PDF. α is called the spread parameter, related to the standard deviation, where a high value means a wide PDF. β is called the shape parameter and is related to the tail asymptotic behavior, where a small value implies a light tail. Both parameters can be estimated by a maximum-likelihood approach using the coefficients in $W_j^{(i)}(z)$.

For instance, if a wavelet decomposition of three scale levels is applied to a stationary texture image, it is possible to assign the parameters α and β to each wavelet sub-band. Thus, a model of 18 parameters is assigned to the image.

Appendix 3: Texture-Cartoon Separation

This method is composed of two stages (Szolgay and Szirányi 2012). In the first one, the image z is filtered by a σ-sized low-pass filter L_σ forming $z_\sigma = L_\sigma * z$, where $*$ denotes the convolution operator. For each pixel x, the value $\Delta LTV(x) = |\nabla z_\sigma(x) - \nabla z(x)|$ is calculated, which corresponds to the change of the local total variation (LTV) after the filtering process. If $\Delta LTV(x)$ has a high value, the cartoon image pixel $z_u(x)$ will be equal to the filtered image pixel $z_\sigma(x)$, if it does not, $z_u(x) = z(x)$. It should be noted that to have a good separation it is necessary to fix σ adequately. For this task, the following orthogonality measure between two images z_1 and z_2 is used

$$\mathrm{ADE}(z_1, z_2) = \left| \arcsin\left(\frac{\langle \bar{z_1}, \bar{z_2} \rangle}{\| \bar{z_1} \|_{\ell_2} \cdot \| \bar{z_2} \|_{\ell_2}} \right) \right|, \tag{33}$$

where $\bar{z_1}$ and $\| \cdot \|_{\ell_2}$ denote the vector formed by the concatenation of columns of z_1 and the ℓ_2-norm, respectively.

Next, an interval $[s_1, s_2]$ of possible σ-values is fixed, and the image z is divided in B blocks. Each block is separated using the previous process for each $\sigma_i \in [s_1, s_2]$ forming z^b_{u,σ_i} and z^b_{v,σ_i}. Next, the blocks z^b_u, $b \in \{1, \ldots, B\}$, that compose the cartoon image are given by

$$z^b_u = z^b_{u,\sigma_m}, \tag{34}$$

where

$$\sigma_m = \arg \min_{\sigma_i \in [s_1, s_2]} (\mathrm{ADE}(z^b_{u,\sigma_i}, z^b_{v,\sigma_i})). \tag{35}$$

Finally, the cartoon and texture images for this first stage are denoted by z^{I}_u and z^{I}_v, respectively, and it is possible to start the second stage. In this part, a new cartoon image denoted by z^{II}_u is created from z^{I}_u. The process begins with a partition of z^{II}_u in B blocks, and then for each $b \in \{1, \ldots, B\}$ the following diffusion equation is applied (Perona and Malik 1990)

$$z^{\mathrm{II},b}_u(x, t+1) = z^{\mathrm{II},b}_u(x, t) + \frac{\lambda}{|\eta(x)|} \sum_{(x') \in \eta(x')} \nabla^{(x')} \times (g(\nabla^{(x')} z^{\mathrm{I},b}_u(x))) \nabla^{(x')} z^{\mathrm{II}}_u(x, t), \tag{36}$$

where $\lambda \in \mathbb{R}$ is a scalar constant that determines the rate of diffusion, $|\eta(x)|$ is the number of elements of the spatial neighborhood of pixel x, g is a weigh function used to not blur cartoon edges, and $\nabla^{(x')} z^{\mathrm{II}}_u(x, t)$ is an approximation of the image gradient at a particular direction given by

$$\nabla^{(x')} z^{\mathrm{II}}_u(x, t) = z^{\mathrm{II}}_u(x', t) - z^{\mathrm{II}}_u(x, t). \tag{37}$$

The last step is to find the iteration t_{ADE} in which the separation carried out in (36) is optimum. For this process, the orthogonality measure given in (33) is used. Thus, the final blocks $u_{\mathrm{II}}^{b},\ b \in \{1,\ldots,B\}$, of the second stage are calculated as follows:

$$z_u^{\mathrm{II},b}(x) = z_u^{\mathrm{II},b}(x, t_{\mathrm{ADE}}), \tag{38}$$

where

$$t_{\mathrm{ADE}} = \arg \min_{i\in\{1,\ldots,I_{\max}\}} (\mathrm{ADE}(z_u^{\mathrm{II},b}(x,t_i), z_v^{\mathrm{II},b}(x,t_i))), \tag{39}$$

and $I_{\max}$ is a previously fixed number of diffusion iterations. Finally, denoting by z_u and z_v the final cartoon and texture images, respectively, we have that

$$z_u = z_u^{\mathrm{II}}, \tag{40}$$

$$z_v = z - z_u. \tag{41}$$

7 Appendix 4: Source Separation Technique

For this method, it is supposed that the image model has the form

$$z = z_{S_1} + z_{S_2}, \tag{42}$$

where z_{S_1} and z_{S_2} are two information sources which are sparse in different transform domains (Starck et al. 2005). An important point to highlight is that each of these sources should not have a good sparsity in the domain in which the other source is sparse. Denoting by Φ_1 and Φ_2 the synthesis matrices associated with the domains in which the sources are sparse, the separation process is carried out by solving the following optimization problem (Fadili et al. 2010)

$$(\hat{z}_{S_1}, \hat{z}_{S_2}) = \arg \min_{z_{S_1}, z_{S_2}} \| \Phi_1^{\mathrm{T}} z_{S_1} \|_{\ell_1} + \| \Phi_2^{\mathrm{T}} z_{S_2} \|_{\ell_1} + \lambda \| z - z_{S_1} - z_{S_2} \|_{\ell_2}^2, \tag{43}$$

where $\lambda \in \mathbb{R}$ and $\| z - z_{S_1} - z_{S_2} \|_{\ell_2}^2$ correspond to a regularization parameter and a term that penalizes a possible noise component in the image, respectively.

References

Bojcevski D (2004) Metallurgical characterisation of George Fisher meso-textures and micro-textures. Master's thesis, University of Queensland

Bonifazi G, Serranti S, Volpe F, Zuco R (2001) Characterisation of flotation froth colour and structure by machine vision. Comput Geosci 27(9):1111–1117

Bonnici NK (2012) The mineralogical and textural characteristics of copper-gold deposits linked to mineral processing attributes. Ph.D. thesis, University of Tasmania

Burges CJ (1998) A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 2(2):121–167

Cabello E, Snchez M, Delgado J (2002) A new approach to identify big rocks with applications to the mining industry. Real-Time Imaging 8(1):1–9

Chatterjee S (2013) Vision-based rock-type classification of limestone using multi-class support vector machine. Appl Intell 39(1):14–27
Cherkassky V, Mulier FM (2007) Learning from data: concepts, theory, and methods. Wiley, New York
Chica-Olmo M, Abarca-Hernandez F (2000) Computing geostatistical image texture for remotely sensed data classification. Comput Geosci 26(4):373–383
Croop A, Goodall W (2013) The influence of rock texture on mineral processing. MinAssist Australia. http://www.minassist.com.au/site/downloads/The-Influence-of-Rock-Texture-on-Processing.pdf. Accessed 15 Apr 2016
Do MN, Vetterli M (2002) Wavelet-based texture retrieval using generalized gaussian densities and Kullback–Leibler distance. IEEE Trans Image Process 11(2):146–158
Donskoi E, Suthers S, Fradd S, Young J, Campbell J, Raynlyn T, Clout J (2007) Utilization of optical image analysis and automatic texture classification for iron ore particle characterisation. Miner Eng 20(5):461–471
Duda RO, Hart PE, Stork DG (2012) Pattern classification. Wiley, New York
Dunham S, Vann J (2007) Geometallurgy, geostatistics and project value—does your block model tell you what you need to know? In: Proceedings of the project evaluation 2007, pp 19–20
Dunham S, Vann J, Coward S (2011) Beyond geometallurgy—gaining competitive advantage by exploiting the broad view of geometallurgy. In: 1st Ausimm international geometallurgy conference
Easley G, Labate D, Lim WQ (2008) Sparse directional image representations using the discrete shearlet transform. Appl Comput Harmon Anal 25(1):25–46
Fadili JM, Starck JL, Elad M, Donoho D et al (2010) Mcalab: reproducible research in signal and image decomposition and inpainting. IEEE Comput Sci Eng 12(1):44–63
Gibert X, Patel VM, Labate D, Chellappa R (2014) Discrete shearlet transform on GPU with applications in anomaly detection and denoising. EURASIP J Adv Signal Process 1:1–14
Gonçalves LB, Leta FR (2010), Macroscopic rock texture image classification using a hierarchical neuro-fuzzy class method. In: Mathematical problems in engineering 2010
Higgins MD (2006) Quantitative textural measurements in igneous and metamorphic petrology, vol 276. Cambridge University Press, Cambridge
Higgins MD, Chandrasekharam D (2007) Nature of sub-volcanic magma chambers, Deccan province, India: evidence from quantitative textural analysis of plagioclase megacrysts in the giant plagioclase basalts. J Petrol 48(5):885–900
Jungmann M, Kopal M, Clauser C, Berlage T (2011) Multi-class supervised classification of electrical borehole wall images using texture features. Comput Geosci 37(4):541–553
Kachanubal T, Udomhunsakul S (2008) Rock textures classification based on textural and spectral features. Int J Comput Intell 4:240–246
Kracht W, Emery X, Paredes C (2013) A stochastic approach for measuring bubble size distribution via image analysis. Int J Miner Process 121:6–11
Kutyniok G, Lim WQ (2012) Image separation using wavelets and shearlets. In: Curves and surfaces. Springer, pp 416–430
Kutyniok G et al (2012) Shearlets: multiscale analysis for multivariate data. Springer, New York
Lepistö L, Kunttu I, Autio J, Visa A (2003) Rock image classification using non-homogenous textures and spectral imaging In: WSCG short papers proceedings. UNION agency, Plzen, Czech Republic, pp 3–7
Lepistö L, Kunttu I, Visa A (2005) Color-based classification of natural rock images using classifier combinations. In: Image analysis. Springer, pp 901–909
Mallat S (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 11:674–693
Mallat S (2008) A wavelet tour of signal processing: the sparse way. Academic Press, Paris
Mock A, Jerram D (2005) Crystal size distributions (csd) in three dimensions: insights from the 3d reconstruction of a highly porphyritic rhyolite. J Petrol 46(8):1525–1541
Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24:971–987
Partio M, Cramariuc B, Gabbouj M, Visa A (2002) Rock texture retrieval using gray level co-occurrence matrix. In: Proceedings of 5th nordic signal processing symposium, vol 75
Perez CA, Estévez PA, Vera PA, Castillo LE, Aravena CM, Schulz DA, Medina LE (2011) Ore grade estimation by feature selection and voting using boundary detection in digital image analysis. Int J Miner Process 101(1):28–36

Perona P, Malik J (1990) Scale-space and edge detection using anisotropic diffusion. Pattern Anal Mach Intell IEEE Trans 12(7):629–639

Petrou M, Sevilla PG (2006) Image processing: dealing with texture, vol 10. Wiley, Chichester

Rudin LI, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Phys D: Nonlinear Phenom 60(1):259–268

Silva JF, Narayanan SS (2012) On signal representations within the bayes decision framework. Pattern Recognit 45(5):1853–1865

Singh M, Javadi A, Singh S (2004) A comparison of texture features for the classification of rock images. In: Intelligent data engineering and automated learning—IDEAL 2004. Springer, pp 179–184

Starck JL, Elad M, Donoho DL (2005) Image decomposition via the combination of sparse representations and a variational approach. Image Process IEEE Trans 14(10):1570–1582

Strang G (1989) Wavelets and dilation equations: a brief introduction. SIAM Rev 31(4):614–627

Szolgay D, Szirányi T (2012) Adaptive image decomposition into cartoon and texture parts optimized by the orthogonality criterion. IEEE Trans Image Process 21(8):3405–3415

Tarbuck EJ, Lutgens FK, Tasa D (2000) Earth science. Prentice Hall, Upper Saddle River

Tessier J, Duchesne C, Bartolacci G (2007) A machine vision approach to on-line estimation of run-of-mine ore composition on conveyor belts. Miner Eng 20(12):1129–1144

Thomas A, Rider M, Curtis A, MacArthur A (2011) Automated lithology extraction from core photographs. First Break 29(6):103–109

van der Maaten LJ, Postma EO, van den Herik HJ (2009) Dimensionality reduction: a comparative review. J Mach Learn Res 10(1–41):66–71

Vasconcelos N (2004) Minimum probability of error image retrieval. Signal Process IEEE Trans 52(8):2322–2336

Vink L (1997) Textures of the hilton north deposit, Queensland, Australia, and their relationship to liberation. Ph.D. thesis, University of Queensland

Yi S, Labate D, Easley GR, Krim H (2009) A shearlet approach to edge analysis and detection. Image Process IEEE Trans 18(5):929–941

Zujovic J, Pappas TN, Neuhoff DL (2013) Structural texture similarity metrics for image analysis and retrieval. Image Process IEEE Trans 22(7):2545–2558