CHAPTER 1

Fundamentals of Texture Processing for Biomedical Image Analysis

A General Definition and Problem Formulation

Adrien Depeursinge[a,*,**], Julien Fageot[*] and Omar Al-Kadi[†]
[*] École Polytechnique Fédérale de Lausanne, Biomedical Imaging Group, Lausanne 1015, Switzerland.
[**] University of Applied Sciences Western Switzerland (HES-SO), Institute of Information Systems, Sierre 3960, Switzerland.
[†] University of Jordan, King Abdullah II School for Information Technology, Amman, Jordan.
[a] Corresponding: `adrien.depeursinge@hevs.ch`

Contents

Abstract

This chapter aims to provide an overview of the foundations of texture processing for biomedical image analysis. Its purpose is to define precisely what biomedical texture is, how is it different from general texture information considered in computer vision, and what is the general problem formulation to translate 2D and 3D textured patterns from biomedical images to visually and biologically relevant measurements. First, a formal definition of biomedical texture information is proposed from both perceptual and mathematical point of views. Second, a general problem formulation for biomedical texture analysis is introduced, considering that any approach can be characterized as a set of local texture operators and regional aggregation functions. The operators allow locally isolating desired texture information in terms of spatial scales and directions of a texture image. The type of desirable operator invariances are discussed, and are found to be different from photographic image analysis. Scalar-valued texture measurements are obtained by aggregating operator’s response maps over regions of interest.

Keywords: Quantitative image analysis, spatial stochastic process, texton, heterogeneity, texture analysis.

1. Introduction

Everybody agrees that nobody agrees on the definition of *texture* information. According to the Oxford Dictionaries[1] texture is defined as “the feel, appearance, or consistency of a surface or a substance”. The context in which the word *texture* is used is fundamental to attach unambiguous semantics to its meaning. It has been widely used in extremely diverse domains to qualify the properties of images, food, materials and even music. In the context of food and sometimes material sciences, characterizing texture information often involves measuring the response of the matter subject to forces such as shearing, cutting, compressing and chewing [63, 6]. Starting from the early developmental months of newborn babies, tactile perception of textured surfaces is an important stage of the human brain development [59]. It is an essential step to be successfully acquainted with the physical properties physical properties of the surrounding environment [39]. To some extent, estimating the properties and diversity of the latter without having to touch every surface can be efficiently carried out through vision. Human vision learns to recognize texture patterns through extensive experimentation confronting visual and tactile perception of textured surfaces [71]. This provides hints on why human visual texture recognition performs much beyond the use of low level descriptive terms such as *coarse*, *edgy*, *directional*, *repetitive*, and *random*.

In the context of biomedical imaging, texture information relates to the micro- and macro- structural properties of biomedical tissue. Radiologists, pathologists and biologists are trained to establish links between visual image patterns and underlying cellular and molecular content of tissue samples [67]. Unfortunately, very large variations of this complex mapping occur, resulting in image interpretation errors with potentially undesirable consequences [56, 68, 3]. These variations are partly due to the diversity of human biology and anatomy as well as image acquisition protocols and reconstruction, compounded by observer training. Important efforts were initi-

---

[1] `https://en.oxforddictionaries.com/definition/texture`, as of 10 October 2016.

ated by medical imaging associations to construct unified terminologies and grading scores in the context of radiology and histopathology, aiming to limit variations in image interpretation and reporting [36, 37, 17, 24]. However, in the particular context of biomedical texture information, the terms used (*e.g.*, *heterogeneous enhancement*, *hypervascular* [17]) are often as inadequate as low level descriptive terms of general textures (*e.g.*, *coarse*, *edgy*) while the perception of human observers is much richer (see Sections 4.1 and 4.2 of Chapter 9). When considering three-dimensional architectures of biomedical tissue, human observers have limited intuition of these 3D solid textures, because they cannot be fully visualized [16]. Only virtual navigation in Multi-Planar Rendering (MPR) and semi-transparent visualizations are made available by computer graphics and allow observing 2D projections.

Computer-based quantitative image texture analysis has a tremendous potential to reduce image interpretation errors and can make better use of the image content by yielding exhaustive, comprehensive and reproducible analysis of imaging features in two and three-dimensions [2, 35, 62]. Nevertheless, besides the lack of a clear definition of biomedical texture information, several challenges remain, such as: the lack of an appropriate framework for multi-scale, multi-spectral analysis in 2D and 3D; validation; and, translation to routine clinical applications. The goal of this book is to illustrate the importance of these aspects and to propose concrete solutions for optimal biomedical texture analysis. This chapter will first propose a definition of texture in the particular context of biomedical imaging (Section 2). Second, a general theoretic framework for Biomedical Texture Analysis (BTA) will be proposed in Section 3. The latter is designed to best leverage the specific properties of biomedical textures. Differences with the classical texture analysis paradigm in computer vision will be highlighted. Important aspects of texture operator and aggregation function design will be further discussed and illustrated through several concrete examples in Chapter 2. It will also recapitulate key aspects of biomedical texture processes and analysis, with an aim of raising awareness of limitations of popular texture operators used in the biomedical literature while providing directions to design the next generation of BTA approaches. Chapter 3 will use the comparison axes established in Chapters 1 and 2 to compare most popular modern biomedical texture analysis approaches. With the purpose of guiding neophyte or experienced users, a simple checklist is proposed to assess the relevance of a the BTA approach in a particular medical or biological applicative context.

2. Biomedical texture processes

This section proposes an extensive definition of biomedical texture information under biological, medical, physical, statistical and mathematical viewpoints.

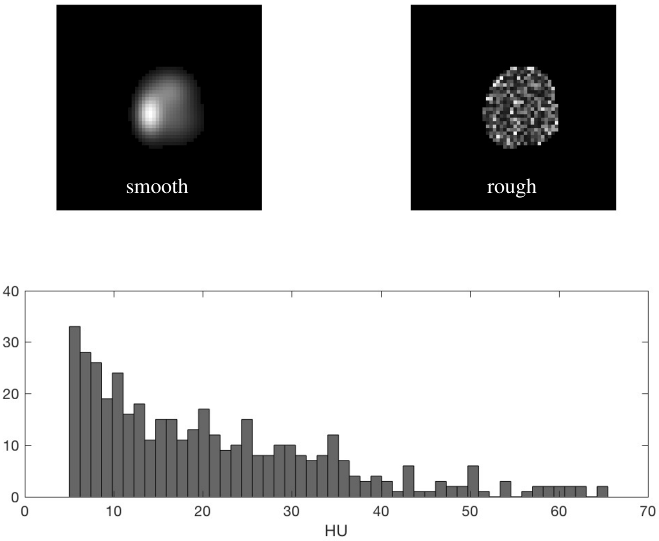


Figure 1.1: The two simulated tumors have identical distribution of the pixel’s values and cannot be differentiated using intensity image measures only. They differ in the spatial relationships between the pixels, which is specifically captured by image texture analysis.

2.1. Image intensity versus image texture

Low-level quantitative image analysis (*i.e.*, pixel-level[2]) can be separated into two main categories: intensity and texture. Image intensity relates to the statistical distribution of the pixel values inside a defined Region Of Interest (ROI). The pixel values can be either normalized across images (*e.g.*, Hounsfield Units (HU) in X-ray Computed Tomography (CT), Standardized Uptake Values (SUV) in Positron Emission Tomography (PET)), or unnormalized (*e.g.*, Hematoxylin and Eosin (H&E) stains in histopathology, Magnetic Resonance Imaging (MRI)). Classic quantitative measures of image intensity are the four statistical moments of the pixel values’ distribution (mean, variance, skewness and kurtosis). Other measures are specific to the considered imaging modality (*e.g.*, SUV max or Total Lesion Glycolysis (TLG) in PET) [49]. The latter are extremely useful to characterize the image content, but cannot measure the spatial relationships between pixel values (see Fig. 1.1). A qualitative keyword such as *tumor heterogeneity* is ambiguous because it is unclear if the heterogeneity

[2] The word *pixel* is used to design both 2D and 3D (*voxels*) image samples.

concerns pixel values (intensity) or their spatial organization (texture). It is though commonly used to describe the visual aspect of tumors in radiological images with ambivalent meaning [33, 12, 48].

The spatial relationships (*i.e.*, the transitions) between pixel values are precisely what texture information is encoding. Haidekker defined texture as “a systematic local variation of image values” [26]. Petrou stated that “the most important characteristic of texture is that it is scale dependent” and that “different types of texture are visible at different scales” [51]. This highlights the importance of the variation *speed* or *slope* or *oscillation* between pixel values, which will be different in *smooth* versus *rough* textures (see Fig. 1.1 left and right). This first notion of texture scale relates to the spatial frequencies in the image. The higher the spatial frequency, the finer the scale of the transition between proximate pixel values. A second important notion is the direction of the transition. These two notions of spatial scale and direction are fundamental for visual texture discrimination (see Fig. 1.2) [64].

Blakemore *et al.* provided initial evidence that the human visual system possesses neurons that are selectively sensitive to directional spatial frequencies [5], which has been widely confirmed later on [45]. Most approaches proposed for computerized texture analysis are leveraging these two properties either explicitely (*e.g.*, Gray-Level Co-occurrence Matrices (GLCM) [27], Gray-Level Run Length Matrices (GLRLM) [22], Gray-Level Size Zone Matrices (GLSZM) [66], directional filterbanks and wavelets [7], Histogram of Oriented Gradients (HOG) [40], Local Binary Patterns (LBP) [47], Scattering Transforms (ST) [8, 60]) or implicitely (*e.g.*, Convolutional Neural Networks (CNN) [34, 57], Dictionary Learning (DL) [23, 42, 28]).

A natural mathematical tool to study directional spatial frequency components in D-dimensional signals and images is the Fourier transform and is defined in Eq. (1.1). It is straightforward to see that the Fourier transform for $\boldsymbol{\omega} = \mathbf{0}$ computes the mean of the function, which is not considered as texture information since it relates the mean intensity of the pixels in the image. For $||\boldsymbol{\omega}|| > 0$, the modulus of the Fourier transform quantifies the magnitude of the transitions, where $||\boldsymbol{\omega}||$ is inversely proportional to the scale and the orientation of the vector $\boldsymbol{\omega}$ defines the direction of the spatial frequencies. The correspondence between texture images $f_i(\boldsymbol{x})$ and their Fourier transforms $\hat{f}_i(\boldsymbol{\omega})$ is illustrated in Fig. 1.2.

Another important notion of texture information relates to the order of the transition between the pixel values (*i.e.*, the order of directional image derivatives). For instance, first-order (gradient) transitions will characterize step-like transitions, whereas second-order transitions (Hessian) describe local image curvature. For instance, modeling these two types can identify various types of transitions either at the margin or inside a tumor region [17, 10].

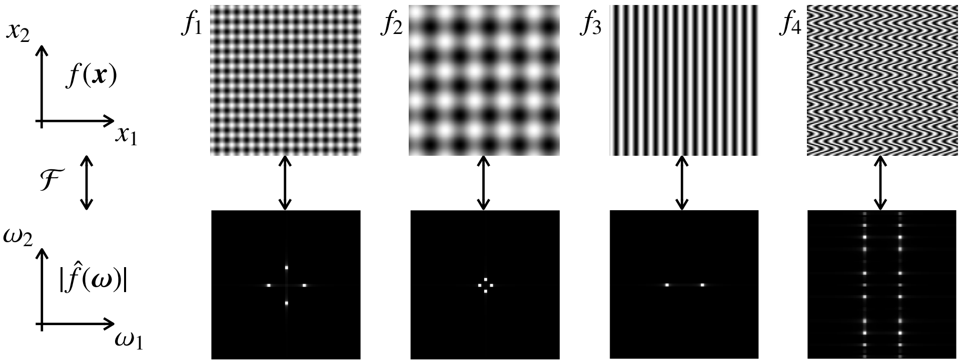


Figure 1.2: Importance of image scales and directions to unequivocally describe texture functions $f(\boldsymbol{x})$. The latter are efficiently represented by the modulus of their dual Fourier representation $\hat{f}(\boldsymbol{\omega})$ in terms of their directional frequency components. $\boldsymbol{\omega} = \mathbf{0}$ corresponds to the centers of the images of the bottom row. f_1 and f_2 have identical directional information but differ in image scales. f_1, f_3, and f_4 have identical spatial frequencies along x_1, but differ in their frequency components along x_2.

2.2. Notations and sampling

Throughout this section, images are defined on an unbounded continuum. This is motivated by the fact that many physical phenomena are intrinsically continuous and results in analog signals. Also, important mathematical operations arising in texture analysis, including geometric transformations (*e.g.*, shifts, rotations, scalings) and random processes, are better defined in the continuous domain without boundaries. For instance, a discrete rotation is defined as the approximation of its continuous domain counterpart. Moreover, stationary random processes (see Section 2.3.1) are implicitly defined over an unbounded domain, since any shifted process should be well-defined. The design of algorithms then calls for the discretization of continuous images, together with their restrictions on bounded domains.

An image is modeled as a D-dimensional function of the variable $\boldsymbol{x} = (x_1, \dots, x_D) \in \mathbb{R}^D$, taking values $f(\boldsymbol{x}) \in \mathbb{R}$. The Fourier transform of an integrable $f(\boldsymbol{x})$ is noted $\hat{f}(\boldsymbol{\omega}) \in \mathbb{C}$ and is defined as

$$f(\boldsymbol{x}) \overset{\mathcal{F}}{\longleftrightarrow} \hat{f}(\boldsymbol{\omega}) = \int_{\mathbb{R}^D} f(\boldsymbol{x})\; e^{-\mathrm{j}\langle \boldsymbol{\omega}, \boldsymbol{x}\rangle} \mathrm{d}\boldsymbol{x}, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product.

As we said, in practice, images and their Fourier transforms have to be sampled and restricted to bounded domains. We introduce now the notations that goes together with these operations. In what follows, however, we rely mostly on the continuous

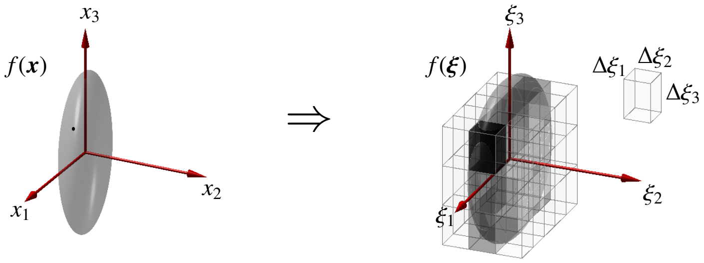


Figure 1.3: Image sampling of 3D texture functions (left: continuous $f(\boldsymbol{x})$, right: discretized $f(\boldsymbol{\xi})$). Most 3D bioimaging protocols have equal axial sampling steps, which differ from the depth sampling step ($\Delta\xi_1 = \Delta\xi_2 \neq \Delta\xi_3$). This results in rectangular voxels.

domain formulation. A D-dimensional discretized image function f of restricted domains is noted as $f(\boldsymbol{\xi}) \in \mathbb{R}$. Pixel positions $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_D)$ are indexed by the vector $\boldsymbol{k} = (k_1, \ldots, k_D)$, the sampling steps $\Delta\xi_1, \ldots, \Delta\xi_D$, and the dimension-wise offsets $c_1, \ldots, c_D$ as

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_D \end{pmatrix} = \begin{pmatrix} \Delta\xi_1 \cdot k_1 \\ \vdots \\ \Delta\xi_D \cdot k_D \end{pmatrix} - \begin{pmatrix} c_1 \\ \vdots \\ c_D \end{pmatrix}, \tag{1.2}$$

with $k_i \in \{1, \ldots, K_i\}$. The domain $\boldsymbol{F} \subset \mathbb{R}^D$ of f is $F_1 \times \cdots \times F_D$ with $F_i = \Delta\xi_i \cdot \{1, \ldots, K_i\} - c_i$. Image sampling of 3D texture functions is illustrated in Fig. 1.3. The sampling steps are most often varying between patients and dimensions. This needs to be taken into account when designing texture operators and is further discussed in Section 3.2.

The discrete Fourier transform of $f(\boldsymbol{\xi})$ is noted $\hat{f}(\boldsymbol{\nu}) \in \mathbb{C}$ and is defined as

$$f(\boldsymbol{\xi}) \overset{\mathcal{F}}{\longleftrightarrow} \hat{f}(\boldsymbol{\nu}) = \sum_{\boldsymbol{\xi} \in \boldsymbol{F}} f(\boldsymbol{\xi})\, e^{-\mathrm{j}\langle \boldsymbol{\nu}, \boldsymbol{\xi} \rangle}, \tag{1.3}$$

with $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_D) \in \boldsymbol{\Omega}$ in the discrete Fourier domain. In the context of discrete texture functions, the realization of the texture process is defined on the sampling grid determined by the sampling step Δx_i. Therefore, the sharpest possible transition can only be measured on a minimal physical distance corresponding to Δx_i. As a consequence, the maximum spatial frequency measurable along the image axes is $\Delta\nu_i = 1/\Delta x_i$. The domain $\boldsymbol{\nu} \subset \mathbb{R}^D$ of $\hat{f}$ is $\Omega_1 \times \cdots \times \Omega_D$ with $\Omega_i = \{1, \ldots, K_i\}/\Delta x_i$ (see Fig. 1.4). The number of elements in $\boldsymbol{\Omega}$ is the same as in $\boldsymbol{F}$: $|\boldsymbol{\Omega}| = |\boldsymbol{F}|$.

The discretization implies that sampled images are approximations of continuous images. Moreover, for the sake of simplicity, although the domains $\boldsymbol{F}$ of image functions are defined as bounded, the potential issues raising at the frontiers are not dis-

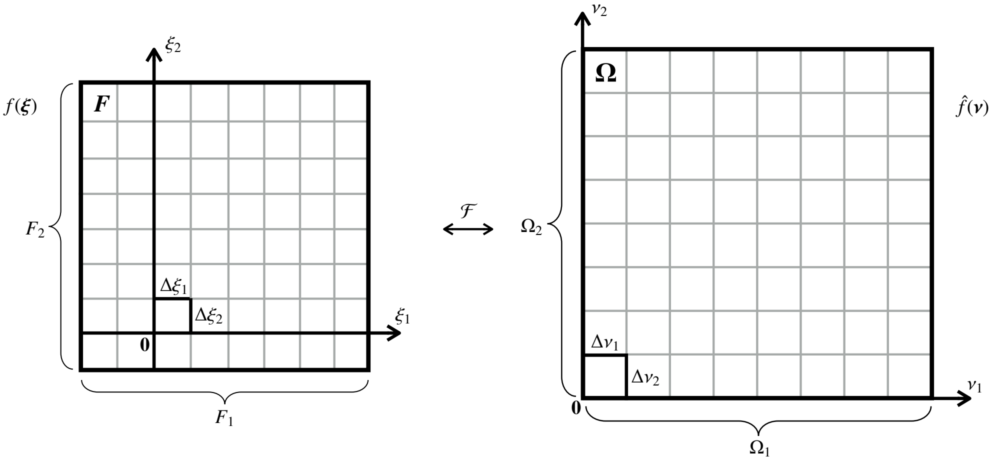


Figure 1.4: Sampled texture function $f(\boldsymbol{\xi})$, $\boldsymbol{\xi} \in \boldsymbol{F}$, and its corresponding Fourier plane $\hat{f}(\boldsymbol{\nu})$, $\boldsymbol{\nu} \in \boldsymbol{\Omega}$. The number of elements in $\boldsymbol{\Omega}$ is the same as in $\boldsymbol{F}$: $|\boldsymbol{\Omega}| = |\boldsymbol{F}|$. The sampling steps $\{\Delta\xi_1, \Delta\xi_2\}$ of the spatial domain are inversely proportional to the sampling steps $\{\Delta\nu_1, \Delta\nu_2\}$ of the Fourier domain: $\Delta\nu_i = 1/\Delta\xi_i$. Note that $\mathbf{0}$ need not lie on the grid for the spatial image domain $\boldsymbol{F}$.

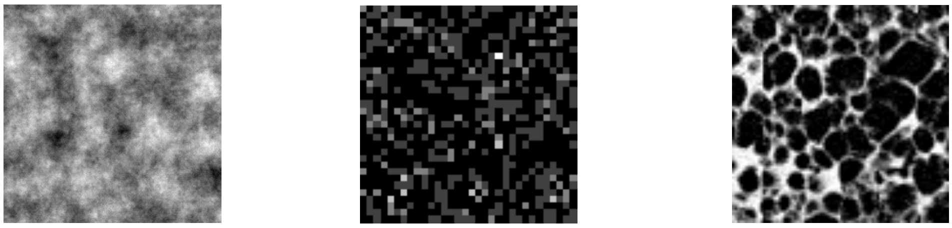
cussed here. An hypothesis that all processing operations are happening far away from the domain boundaries will be made. This will be particularly important when discussing the stationarity of texture processes in Section 2.3.1, as well as texture operators and their application to image functions in Section 3.1.

2.3. Texture functions as realizations of texture processes

Biomedical texture patterns are produced by stochastic biological processes, where randomness results from the diversity of human biology and tissue development. Therefore, it is important to clarify that observed texture functions $f(\boldsymbol{x})$ are realizations of spatial stochastic texture processes of various kind [58]. We define stochastic texture processes of $\mathbb{R}^D$ as

$$\{X(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^D\},$$

where $X(\boldsymbol{x})$ is the value at the position indexed by $\boldsymbol{x} \in \mathbb{R}^D$. The values of $X(\boldsymbol{x})$ follow a Probability Density Function (PDF) $p_{X(\boldsymbol{x})}(q)$, where q is the value of the pixel at the position $\boldsymbol{x}$. Classical examples of such processes are moving average Gaussian or pointwise Poisson (see Fig. 1.5), for which the probability laws are well controlled. However, biomedical texture processes (*e.g.*, lung fibrosis in CT, see Fig. 1.5 right) are much less studied and described in the literature, because we can only fit a PDF from sets of observation containing large variations caused by differences in anatomy



moving average Gaussian
$|\boldsymbol{F}| = 128 \times 128$

pointwise Poisson
$|\boldsymbol{F}| = 32 \times 32$

biomedical: lung fibrosis in CT
$|\boldsymbol{F}| = 84 \times 84$

Figure 1.5: Discretized texture processes $X(\boldsymbol{\xi})$ of $\boldsymbol{F} \subset \mathbb{R}^2$ (see Section 2.2 for a definition of $\boldsymbol{F}$). PDFs of moving average Gaussian and pointwise Poisson are known, which is never the case for biomedical texture processes (*e.g.*, lung fibrosis in CT).

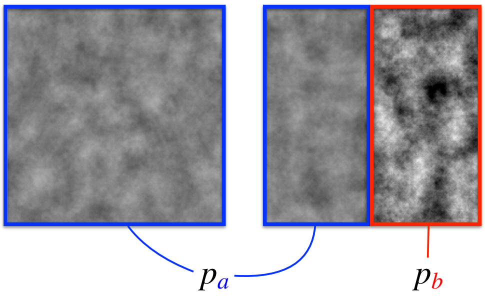
of individuals and imaging protocols.

2.3.1. Texture stationarity

A spatial stochastic process $X(\boldsymbol{x})$ is stationary in the strict sense if it has the same probability law than all its shifted versions $X(\boldsymbol{x} - \boldsymbol{x}_0)$ for any shift $\boldsymbol{x}_0 \in \mathbb{R}^D$. As a consequence, the PDF $p_{X(\boldsymbol{x})}(q) = p_X(q)$ does not depend on the position $\boldsymbol{x}$ anymore.

A discretized stochastic process inherits the stationary property by restricting to discretized shifts. The stationarity is not well-defined at the boundary of the finite domain $\boldsymbol{F}$ of discrete bounded image, but is still meaningful when we are far from the boundaries. For instance, let us consider a patch of dimension 3×3 sliding over the domain $\boldsymbol{F}$ while being fully included in $\boldsymbol{F}$ (no boundary conditions required). For any position of the patch and for each of its nine elements, the statistics of its pixels values will follow the same probability law. In images, texture processes may be piecewise-stationary in a well-defined ROI. Realizations of homo- and heteroscedastic (therefore non-stationary) moving average Gaussian processes are shown in Fig. 1.6.

Since the PDFs are not known in the context of biomedical texture processes, strict stationarity must be relaxed to rather assess the similarity between local PDFs. Even more generally, the notion of spatial stationarity of natural texture processes must be redefined in terms of visual perception and/or tissue biology. It is important to note that the notions of spatial stationarity and texture classes are not tightly bound, since natural texture classes can be constituted by distinct texture processes (see Fig. 1.7). Likewise, biomedical texture classes may be defined based on anatomical or disease levels and can include several cells and tissue types. The image analysis task when the textures processes of interest are stationary relate to texture classification. Texture segmentation of ROIs is considered when the processes are piecewise-stationary.



$$p_{a,\boldsymbol{X}(\boldsymbol{x})}(q) = \frac{1}{1\sqrt{2\pi}}e^{-\frac{(q-0)^2}{2\,1^2}}$$

$$p_{b,\boldsymbol{X}(\boldsymbol{x})}(q) = \frac{1}{3\sqrt{2\pi}}e^{-\frac{(q-0)^2}{2\,3^2}}$$

Figure 1.6: Texture stationarity illustrated with homo- (left) and heteroscedastic (center) moving average Gaussian processes. The central image contains the realizations of two texture processes that are piecewise-stationary over the blue or the red region.

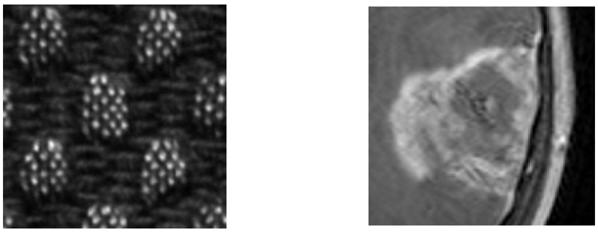


Figure 1.7: Relaxed spatial stationarity of natural texture processes defined in terms of visual perception and/or tissue biology. Left: The Outex texture class “canvas039” [46] consists of two distinct visual patterns. Right: Brain glioblastoma multiforme in an axial slice of gadolinium-injected T1-MRI includes both an enhancing region and a necrotic core defining distinct subregions. In these two examples, the textures are containing two clear distinct patterns that can be considered homogeneous although not stationary in the strict sense.

2.4. Primitives and textons

As opposed to stochastic approaches, more deterministic methods for texture recognition consider that homogeneous classes are constituted by sets of fundamental elementary units (*i.e.*, building blocks) called *texture primitives* [52]. Examples of such texture classes are shown in Fig. 1.8. The bottom row of Fig. 1.8 highlights the importance of the geometric relationships between primitives, where f_1 and f_2 differ only in terms of the density and geometric transformations (*e.g.*, local rotations) of the same primitive. Occlusions are observed in the upper left texture of Fig. 1.8.

Most biomedical texture patterns are of stochastic nature and are constituted by primitives that are neither well-defined, nor known in advance (see Fig. 1.10). A more general definition of fundamental texture micro-structures was proposed by Julesz in

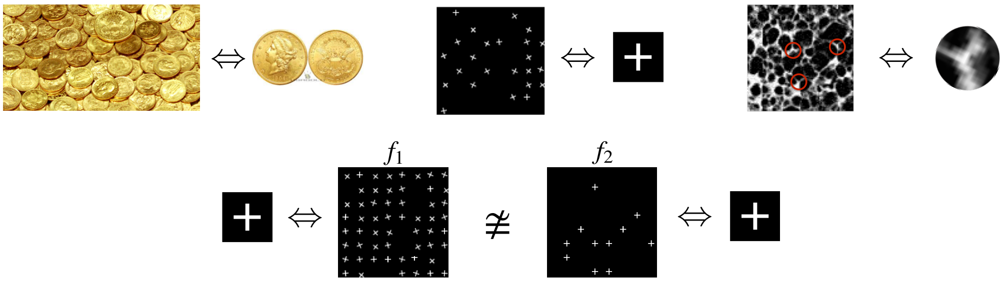
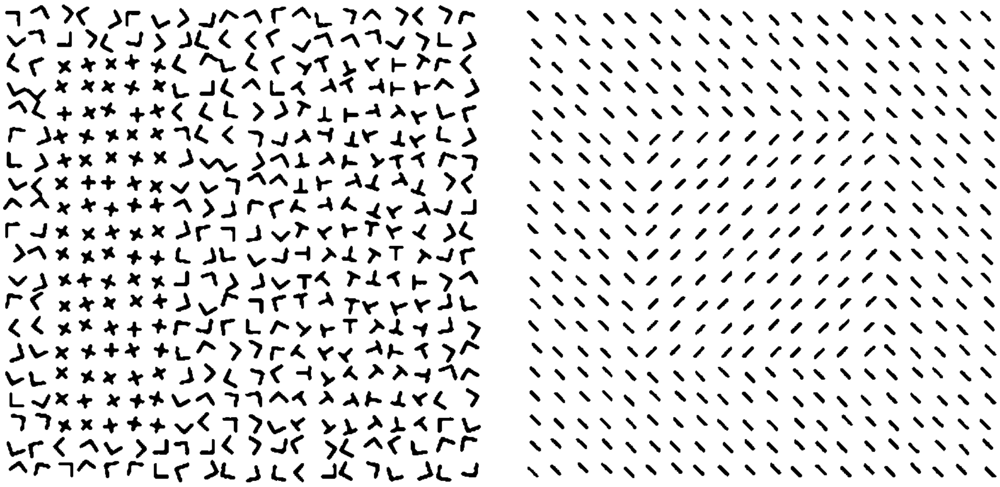


Figure 1.8: Top: various textures and their corresponding primitives or textons. Bottom: importance of the spatial organization (*e.g.*, density, local geometric transformations, occlusions) of the primitives.

1981, who coined the word *texton* [32]. This was one of the earliest texture perception studies of human vision, which laid down the fundamentals of preattentive vision [30, 32, 31, 29, 4]. The concept of texture being effortlessly or preattentively discriminable, *i.e.*, the spontaneous perception of texture variations without focused attention, is illustrated in Fig. 1.9 (a), where the left side area is considered preattentively discriminable while the right side is not and can only be distinguished after careful visual analysis. Similarly, for Fig. 1.9 (b), the inner square is easily discerned from the outer square. Julesz attributed this to the order of image statistics of texture patterns. Textures having equal second order statistics in the form of identical joint distribution of grey level pairs, tend to not be preattentively discriminable. An example of this focused attention is the case between the **L**-shaped background and the **T**-shaped right region of Fig 1.9 (a). This is known as Julesz conjecture, which he and his colleagues refuted afterwards using carefully constructed textures having equal second order and different third and higher order statistics [29]. Nevertheless, this conjecture gave a better understanding of texture perception which led to the proposal of the texton theory. The theory states that texture discrimination is achieved by its primitive or fundamental elements, named textons [32, 31] or sometimes called texels (stands for texture elements) [27]. Several approaches were proposed to derive collections of textons from texture images using texture operators [73, 69] (*e.g.*, Gabor, Laplacians of Gaussians, see Section 2.2 of Chapter 3), or directly from pixel values spatial organization in local neighborhoods [23, 61, 42] (*e.g.*, dictionary learning, see Section 2.3 of Chapter 3)).

Understanding the types of primitives and textons, as well as their (local) spatial organization in biomedical texture classes is of primary importance to efficiently design BTA approaches and in particular texture operators. This is further discussed in Section 3 as well as in Chapters 2 and 3.

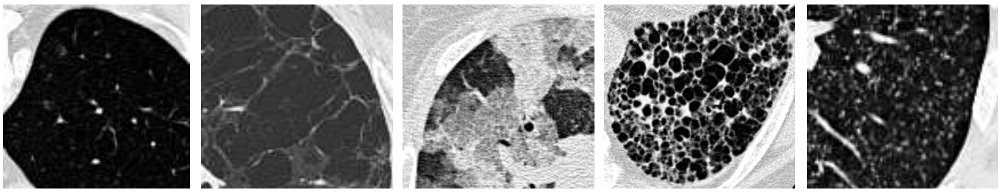


(a) (b)

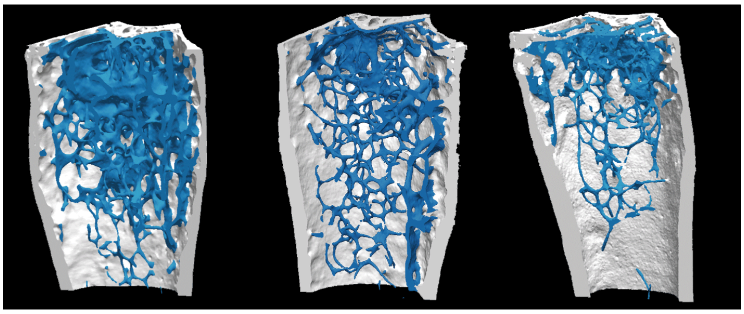
Figure 1.9: An example of preattentive discrimination [15]. (a) texture composed of two regions: left **+**-shaped textons are preattentively (*i.e.*, effortlessly) distinguishable against **L**-shaped textons, while in the right the **T**-shaped textons needs focused attention (*i.e.*, using long-term memory). (b) texture composed of line segments where the difference in orientations segregates preattentively the middle region from the outer region.

2.5. Biomedical image modalities

When compared to the broad research field of computer vision, analyzing texture in biomedical images can rely on highly specific properties of the latter. In particular, most biomedical imaging modalities rely on controlled acquisition protocols allowing the standardization of fundamental image properties such as pixel/voxel size and intensity, as well as image orientation and sample/patient position [20]. The setting is therefore fundamentally different from general computer vision based on photographic imagery resulting from scene captures obtained with varying viewpoints. The properties of most common biomedical imaging protocols are listed in Table 1.1. Imaging protocols with unknown or inaccurate physical size of pixels are rare but exist. Examples of the latter are scanned film-based X-ray radiographs (obsolete), digital photography for dermatology and ophthalmology, and endoscopic videos (see Section 3 of Chapter 11).



healthy emphysema ground glass honeycombing micronodules



Isodensity visualization of 3D trabeculae bone microarchitecture.

Figure 1.10: 2D and 3D biomedical textures are mainly stochastic and do neither have known nor well-defined primitives. Top: normal and altered 2D lung parenchyma from interstitial lung diseases in axial high-resolution CT [14]. Bottom: 3D bone microarchitecture of normal and osteoporotic trabeculae (blue) in micro CT [21].

3. Biomedical Texture Analysis (BTA)

In this section, a general theoretic framework for D-dimensional texture analysis is introduced. The latter leverages the specific properties of biomedical texture processes and functions defined in Section 2. A general formulation of texture operators and aggregation functions is introduced in Section 3.1. Popular texture analysis approaches are exemplified as particular cases of operators and aggregation functions. The importance of inter-patient and inter-dimension scale normalization, as well as operator robustness to rigid transformations is highlighted in Sections 3.2 and 3.3. Fundamental limitations, current approaches and future directions of multi-scale and multi-directional image analysis are further discussed in Sections 3 and 4 of Chapter 2, respectively. The operators and aggregation functions of most popular BTA approaches are described and qualitatively compared in Chapter 3.

Table 1.1: Typical specifications of common biomedical imaging modalities.

modality	signal measured	dimensionality	image dimensions	physical dimensions of pixels/voxels	image type
40× digital histopathology	light absorption	2D color	$\approx 50,000^2$	$\Delta x_1, \Delta x_2$: $\approx 0.275\mu m$.	structural
digital radiography	X-ray absorption	2D grayscale	$\approx 2048^2$	$\Delta x_1, \Delta x_2$: $\approx 0.1mm$.	structural
MRI	nuclear magnetic resonance	3D grayscale	$\approx 512^3$	$\Delta x_1, \Delta x_2$: 0.5-1mm, Δx_3: 0.5-5mm.	structural
CT	X-ray absorption	3D grayscale	$\approx 512^3$	$\Delta x_1, \Delta x_2$: 0.5-1mm, Δx_3: 0.5-5mm.	structural
3–D US	acoustic reflection	3D grayscale	$\approx 256^3$	$\Delta x_1, \Delta x_2, \Delta x_3$: 0.5-1$mm$.	structural
μCT	X-ray absorption	3D grayscale	$\approx 512^3$	$\Delta x_1, \Delta x_2, \Delta x_3$: $\approx 1\mu m$.	structural
PET	γ-rays emitted indirectly by a radiotracer	3D grayscale	$\approx 256^3$	$\Delta x_1, \Delta x_2, \Delta x_3$: 2-5$mm$.	functional
OCT	optical scattering	3D grayscale	$\approx 1024^2 \times 256$	$\Delta x_1, \Delta x_2, \Delta x_3$: $< 10\mu m$.	structural

3.1. Texture operators and aggregation functions

Without any loss of generality, we propose to consider that every D-dimensional texture analysis approach can be characterized by a set of N local operators $\mathcal{G}_n$ and their corresponding spatial supports $\boldsymbol{G}_n \subset \mathbb{R}^D$, assumed to be bounded. The value $\mathcal{G}_n\{f\}(\boldsymbol{x}_0) \in \mathbb{R}$ corresponds to the application of the operator $\mathcal{G}_n$ to the image f at location $\boldsymbol{x}_0$. $\mathcal{G}_n$ is local in the sense that $\mathcal{G}_n\{f\}(\boldsymbol{x}_0)$ only depends on the values of f on the shifted spatial support $\boldsymbol{G}_n + \boldsymbol{x}_0$ (see Fig. 1.11 for an illustration in 2D).

The operator $\mathcal{G}_n$ is typically applied to the input texture function $f(\boldsymbol{x})$ by *sliding* the spatial support of its function over all positions $\boldsymbol{x}_0$ in $\mathbb{R}^D$. This yields response maps[3] $h_n(\boldsymbol{x}_0)$ as

$$h_n(\boldsymbol{x}_0) = \mathcal{G}_n\{f\}(\boldsymbol{x}_0). \tag{1.4}$$

The structure of the output space of the operator will depend on the desired properties and invariances of operators [53]. The operator $\mathcal{G}_n$ can be linear (*e.g.*, linear filter) or non-linear (*e.g.*, max, linear filter combined with a rectifier in CNNs, GLCMs, LBPs). Accordingly, the response maps h_n can highlight desired properties of the input texture function (*e.g.*, spatial scales corresponding to a well defined frequency band, im-

---

[3] They are commonly called *feature maps* in CNNs.

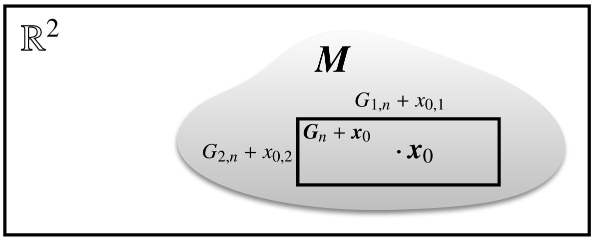


Figure 1.11: At a fixed position $\boldsymbol{x}_0$, texture functions $f(\boldsymbol{x})$ are analyzed by local texture operators $\mathcal{G}_n$ with spatial supports $\boldsymbol{G}_n = G_{1,n} \times \cdots \times G_{D,n}$, where $\boldsymbol{G}_n \subset \mathbb{R}^D$. When applied to all positions $\boldsymbol{x}_0 \in \mathbb{R}^D$, $\mathcal{G}_n$ yields response maps $h_n(\boldsymbol{x}_0) = \mathcal{G}_n\{f\}(\boldsymbol{x}_0)$, which can be aggregated over a ROI $\boldsymbol{M} \subset \mathbb{R}^D$ to obtain a scalar-valued texture feature η_n.

age gradients along x_1, co-occurrences, local binary patterns or circular frequencies). Examples of response maps are shown in Figures 1.12 and 1.13. The properties of popular texture operators are discussed and compared in Chapter 3. In the particular case of linear operators, the application of the operator $\mathcal{G}_n$ to the image function $f(\boldsymbol{x})$ at a given position is a *scalar product* between $f(\boldsymbol{x})$ and a function $g_n(\boldsymbol{x})$ with support $\boldsymbol{G}_n$ (see Eq. 3.2 of Chapter 3). Applying a linear operator $\mathcal{G}_n$ to the input texture function $f(\boldsymbol{x})$ by *sliding* the spatial support of its function over all positions $\boldsymbol{x}_0$ is called *convolution* (see Eq. 3.1 of Chapter 3). Convolutional texture operators are discussed in Section 2 of Chapter 3.

In order to extract collections of scalar measurements $\boldsymbol{\eta} = (\eta_1, \dots, \eta_N)$ from N response maps $h_n(\boldsymbol{x})$, an aggregation function is required to gather and summarize the operators’ responses over a defined ROI domain $\boldsymbol{M} \subset \mathbb{R}^D$ (see Fig. 1.11). The values of the vector $\boldsymbol{\eta}$ define coordinates of a texture instance in the feature space $\mathbb{R}^N$. Integrative aggregation functions are commonly used to extract estimations of features statistics (*e.g.*, counts, means, covariances). For instance, the mean can estimate the average responses of a given operator over $\boldsymbol{M}$ as

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_N \end{pmatrix} = \frac{1}{|\boldsymbol{M}|} \int_{\boldsymbol{M}} \Big(h_n(\boldsymbol{x})\Big)_{n=1,\dots,N} \mathrm{d}\boldsymbol{x}, \tag{1.5}$$

where $|\boldsymbol{M}| = \int_{\boldsymbol{M}} \mathrm{d}\boldsymbol{x}$ is the area[4] covered by $\boldsymbol{M}$. Aggregation functions are not limited to integral operations. For example, $\max\limits_{\boldsymbol{x} \in \boldsymbol{M}} \Big(h_n(\boldsymbol{x})\Big)$ is an aggregation function used in deep CNNs for the max-pooling of feature maps (see Section 2.4.1 of Chapter 4). It is

---

[4] or volume, hypervolume when $D > 2$.

worth noting that the aggregation operation can itself be seen as an operator applied to the feature maps $h_n(\boldsymbol{x})$. However, it differs from the definition introduced in Eq. (1.4) as it relies on irregular spatial supports defined by $\boldsymbol{M}$, which are not slid over the entire image domain.

Examples of simple convolutional texture operators, response maps and aggregation functions are shown in Fig. 1.12 (multi-scale circularly symmetric operators) and Fig. 1.13 (directional operators). Both figures show that specific operator design allow highlighting desirable or discriminatory texture properties (*e.g.*, scale, directionality). However, it appears clearly in Fig. 1.10 that the properties of biomedical textures are very subtle and have big intra-class variations. To some extent, the latter type of variation is due to local rigid transformations of the structures (*e.g.*, a local rotation of a vessel or bronchus following the anatomy of lung lobes). In order to limit the impact of these variations on the performance of BTA algorithms, we will define more precisely the kind of texture transformations that can be expected, as well as strategies to make BTA approaches more robust to them. In particular, the characteristics of texture normalization and operator invariances is introduced in Sections 3.2 and 3.3, respectively. Operator invariances are further discussed in Chapter 7. The management of intra- and inter-class complexity observed in the hyperspaces spanned by aggregated responses of the operators (texture features $\boldsymbol{\eta}$) is developed in Chapter 6.

Designing sets of texture operators that are able to accurately identify tissue type, function, genomics or response to treatment while accounting for intra-class variation is a very challenging task. How to design such operators? In the rare case where class-specific texture primitives are known, the ideal set of operators would be detectors of those primitives (*e.g.*, detectors of crosses, coins, or collagen junctions for the textures shown in Fig. 1.8), where aggregation functions would count the occurrences of the primitives to quantify their density. Knowing texture primitives is rare in practice and more general approaches for operator design are required. The most general strategy for operator design relies on understanding and/or characterizing the whole span of possible texture functions and, within this span, identifying most important combinations of image scales and directions with discriminative capabilities for each considered tissue class in the specific BTA task in hand. This can be done by exhaustively parceling the Fourier domain by groups of spatial frequencies and their directions (see Section 2.1 and Section 2.2 of Chapter 3). However, the exhaustive analysis of spatial scales and directions is computationally expensive for discrete texture functions $f(\boldsymbol{\xi})$ with densely sampled domains $F_1 \times \cdots \times F_D$ and choices are required (see Chapter 2). Making such choices has been studied in the literature by following essentially two opposing strategies: feature *handcrafting* versus feature *learning*. Feature handcrafting requires defining strong assumptions on expected discriminative types of image scales, directions and transitions. Classical examples of handcrafted texture operators are circularly/spherically symmetric or directional Gabor wavelets [1],

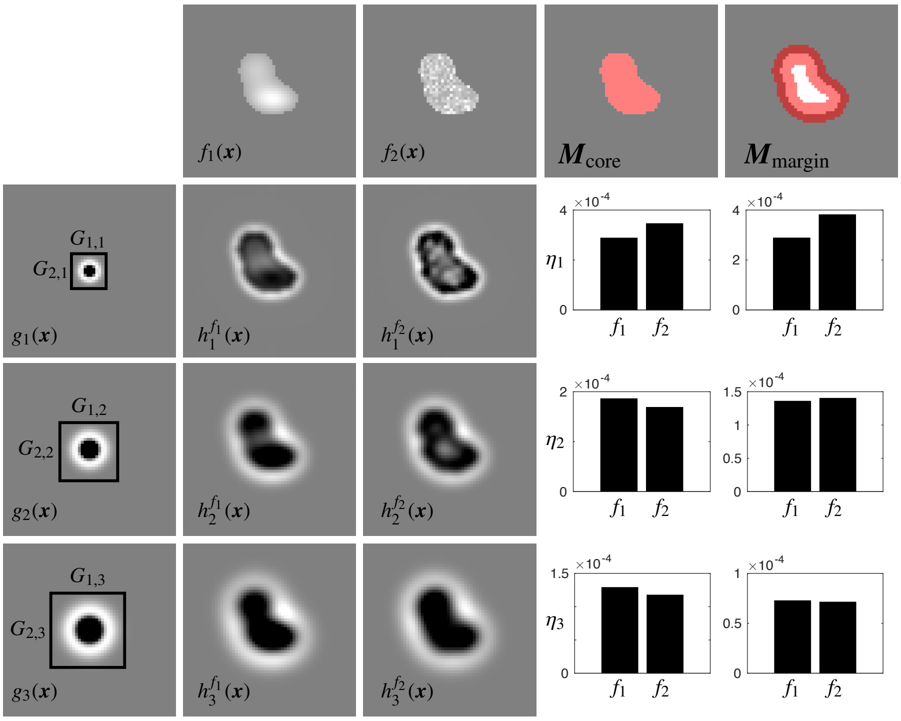


Figure 1.12: 2D texture functions f_1 and f_2 (synthetic tumors) are analyzed with a collection of convolutional circularly symmetric texture operators $\{\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3\}$ with functions $\{g_1, g_2, g_3\}$. The latter are 2D Laplacians of Gaussians (LoG, see Section 2.1 of Chapter 3) at different scales and have different spatial supports $G_{1,n} \times G_{2,n}$. The resulting response maps $h_n^{f_i}$ are revealing distinct properties of both the core and the margin of the tumors. In particular, $h_1^{f_2}(\boldsymbol{x})$ highlights the core texture of f_2. This is verified when averaging the response maps[5] over the core region $\boldsymbol{M}_{\text{core}}$ to obtain scalar measurements η_n, where $\eta_1^{f_1} < \eta_1^{f_2}$. Likewise, $h_1^{f_2}(\boldsymbol{x})$ highlights the margin of the tumor in f_2, where $\eta_1^{f_1} < \eta_1^{f_2}$. The texture scales captured by g_2 and g_3 are too large and do not discriminate well between f_1 and f_2, neither for the core nor for the margin of the tumors.

---

[5] The average of the absolute values of $h_n^{f_i}(\boldsymbol{x})$ is computed since LoGs are band-pass functions in the Fourier domain and have zero mean.

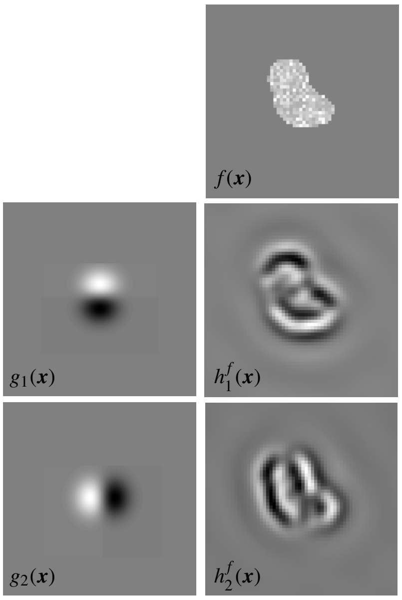
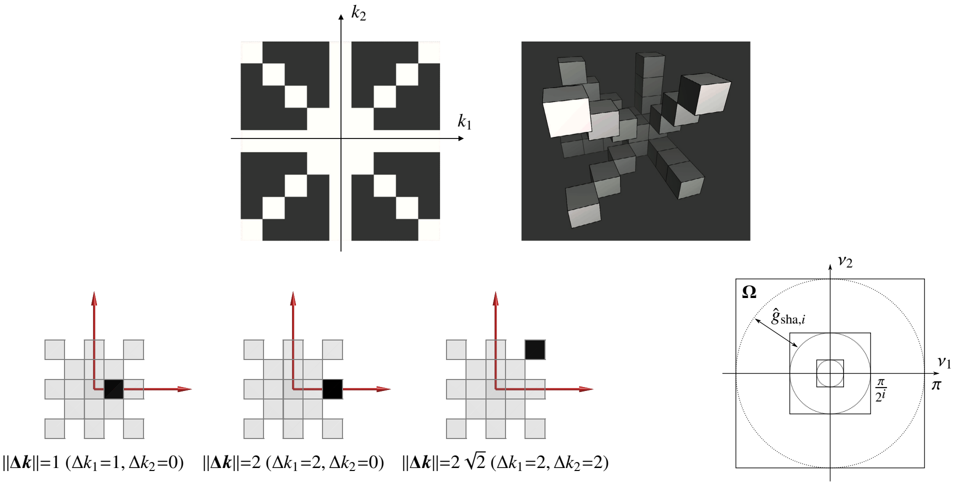


Figure 1.13: Directional response maps of scaled gradient operators along horizontal and vertical axes.

GLCMs [27], HOG [40], LBPs [47], and the ST [8]. Examples of such choices are depicted in Fig. 1.14. Making these choices (*i.e.*, feature handcrafting) raises two issues: How much information is missed or mixed in between two sampled directions or frequency bands? And, among the chosen image scales and directions, which ones are useful for texture discrimination? To address both issues, feature learning approaches relying on either supervised or unsupervised machine learning have recently been proposed to identify context-specific intended texture properties from data. Unsupervised approaches do not require labeled data to design image features and characterize image textons. Notable examples are Principal Component Analysis (PCA) [70], Bags of Visual Words (BoVW) [61] and unsupervised DL [23]. However, although the learned texture operators are data-driven, nothing guarantees that the latter are useful for discriminating a desired texture class from the background or from other types. To overcome this limitation, supervised feature learning approaches can design image features based on a set of labeled data and learning rules. Noteworthy examples are supervised DL [42], learned wavelets [14, 15, 54], and deep CNNs [43, 34]. They are presented in Section 2.3 of Chapter 3. Deep CNNs for texture analysis are thoroughly discussed in Chapters 4, 9 and 10.

The importance of image normalization and operator invariances has already been mentioned in the previous sections and warrants further clarification. They are dis-



$||\boldsymbol{\Delta k}||=1$ $(\Delta k_1=1, \Delta k_2=0)$ $||\boldsymbol{\Delta k}||=2$ $(\Delta k_1=2, \Delta k_2=0)$ $||\boldsymbol{\Delta k}||=2\sqrt{2}$ $(\Delta k_1=2, \Delta k_2=2)$

Figure 1.14: Popular choices of image scales and directions for handcrafted texture operator design. Top: popular 2D and 3D regular sampling of image directions along image axes and diagonals used in GLCMs, RLE, and HOG [16]. Bottom left: 2D sampling of image scales and directions for computing GLCMs. Bottom right: systematic sampling of the Fourier domain $\boldsymbol{\Omega}$ into dyadic bands i with 2D Shannon circularly symmetric wavelets $\hat{g}_{\mathrm{sha},i}(||\boldsymbol{\nu}||)$ [9, 50].

cussed in Sections 3.2 and 3.3.

3.2. Normalization

Image normalization ensures optimal comparisons across data acquisition methods and texture instances. The normalization of pixel values (intensity) is recommended for imaging modalities that do not correspond to absolute physical quantities. Various advanced strategies have been proposed to normalize values and are often modality-specific (*e.g.*, MRI [11], histopathology [41]). Examples of pixel value normalization can be found in Section 2.1 of Chapter 10, and Section 2.4 of Chapter 11. The normalization of image sampling steps $\Delta\xi_d$ $(d = 1, \ldots, D)$ across subjects and dimensions is crucial to ensure accurate comparisons between scales and directions, which is illustrated in Fig. 1.15. Image resampling[6] with identical sampling steps both across image series and image dimensions can be used to normalize image scales and directions. Resampling can be carried out either on the image itself, or on the texture operator. In both cases, care must be taken on the transfer function of the resampling strategy, which can have an important influence on texture properties [65]. Examples of image

---

[6] Classic resampling strategies can be used (*e.g.*, nearest neighbor, multilinear, multicubic).

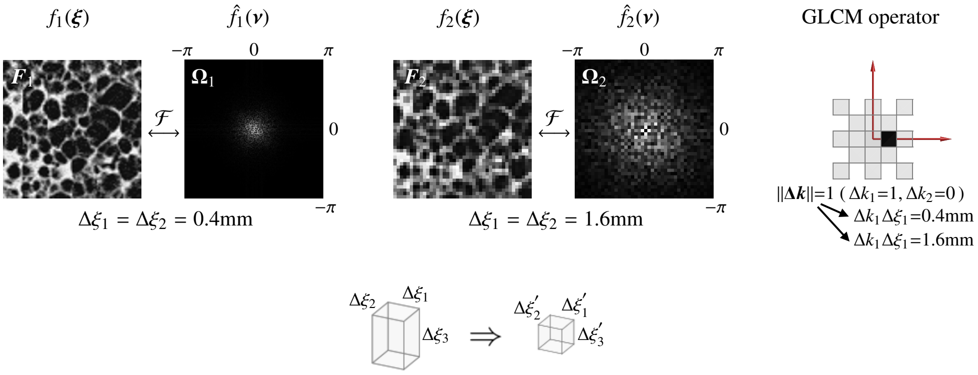


Figure 1.15: Importance of image sampling normalization across image series and image dimensions. Top row: sampling steps $\{\Delta\xi_1, \Delta\xi_2\}$ are four times smaller in the left image than in the center one. This results in a dilated distribution of spatial frequencies in $[-\pi, \pi]$ in the Fourier domain (the moduli of $\hat{f}_i(\boldsymbol{\nu})$ are displayed), because the normalized Nyquist frequencies π are not obtained with the same normalization. It is worth noting that $|\boldsymbol{F}_1| = 4^2 \cdot |\boldsymbol{F}_2|$. Likewise, a GLCM operator parameterized by a horizontal displacement of one pixel ($\|\boldsymbol{\Delta k}\|$=1) corresponds to a physical displacement of either 0.4mm or 1.6mm. Image resampling with a fixed step defined in physical dimensions (*e.g.*, $\Delta\xi_1$=$\Delta\xi_2$=0.4mm) is required to normalize scales across image series. Bottom row: image resampling across dimensions $\{\Delta\xi_1, \Delta\xi_2, \Delta\xi_3\} \Rightarrow \{\Delta\xi_1', \Delta\xi_2', \Delta\xi_3'\}$ to obtain cubic voxels and to ensure an isotropic description of image directions.

resampling for spatial normalization can be found in Section 2.4.1 of Chapter 12, as well as in Sections 4.2.2 and 5.2 of Chapter 4.

3.3. Invariances

An operator or a measurement is *invariant* if its response is insensitive to the transformation of the input image. When the output of the operator is affected by the transformation in the same way the input is, we say that it is *equivariant*.

Invariances of the final texture measures $\boldsymbol{\eta}$ are required to provide robust recognition of all intra-class variants, while preserving inter-class variations (see also Chapter 6 for a machine learning perspective of the problem). Whereas the sources of intra-class variations can be extremely diverse in the context of biomedical tissue (*e.g.*, imaging protocol, subject age, genetics, history), an important subcategory is related to geometric transformations of the tissue architecture. Examples of the effect of such transformations on texture functions are depicted in Fig. 1.16. Assumptions on types of geometric transformations expected in biomedical images are different from those expected in photographic images, because the well-controlled acquisition protocols have little viewpoint variations and allow to control the physical size of pixels (see Section 2.5). As a consequence, while robustness of the texture measures to *affine* transformations (*e.g.*, translation, rotation, scaling) are desired for analyzing photographic images, measures that are robust to a smaller group called *Euclidean* or *rigid* transformations (translation and rotation) are recommended for biomedical image analysis (see Chapter 7). Scale, itself, is a powerful discriminative property in biomedical image analysis.

To better understand the types of geometric invariance that are desirable in the context of a particular image analysis task at hand, we must disentangle invariances of the operator's output (*i.e.*, at the level of the response maps $h_n(\boldsymbol{x}) = \mathcal{G}_n\{f\}(\boldsymbol{x})$) and the ones of the final texture measures $\boldsymbol{\eta}$.

3.3.1. Invariance and equivariance of operators

We will first discuss the desired invariances of texture operators to design BTA approaches that are robust to rigid transformations. Let us consider a general group of geometric transformations $\mathbf{A}$ which can be used for *e.g.*, translation, rotation, scaling, rigid, or affine. In a strict sense, a texture operator $\mathcal{G}_n$ that is invariant to the group $\mathbf{A}$ observes

$$\mathcal{G}_n\{f\}(\,\cdot\,) = \mathcal{G}_n\{\mathcal{A}\{f\}\}(\,\cdot\,) \quad \forall f,\ \forall \mathcal{A} \in \mathbf{A}, \tag{1.6}$$

where $\mathcal{A}$ is a geometric transformation operator implementing a concrete transformation of the group $\mathbf{A}$ (*e.g.*, a 2D translation by the vector $(1,1)$). Eq. (1.6) imposes an extremely strong requirement that would severely limit the ability of operators to extract any useful texture information. For instance if $\mathcal{A}$ is a translation operator as $\mathcal{T}\{f\}(\,\cdot\,) = f(\,\cdot\, - \boldsymbol{x}_0)$, enforcing $\mathcal{G}_n$ to be invariant to $\mathcal{T}$ means that the response map $h_n(\boldsymbol{x})$ (see Eq. 1.4) will not change under the effect of $\mathcal{T}$. Therefore, the spatial positions of the texture “events” will be lost when such a translation-invariant operator is used. This operator will be useless for texture segmentation and difficult to use for texture classification. To preserve the ability of operators to characterize local texture properties, the operators must be *equivariant* to rigid transformations [44], which implies that if the input image $f(\boldsymbol{x})$ is translated and rotated, the response map $h_n(\boldsymbol{x})$ must undergo the same rigid transformation. This means that $\mathcal{G}_n$ must commute with $\mathcal{A}$ as

$$\mathcal{A}\{\mathcal{G}_n\{f\}\}(\,\cdot\,) = \mathcal{G}_n\{\mathcal{A}\{f\}\}(\,\cdot\,) \quad \forall f,\ \forall \mathcal{A} \in \mathbf{A}. \tag{1.7}$$

For some geometric transformations, the origin plays a crucial role. This is the case for rotations and scalings, for which the origin is a fixed point ($\mathcal{A}\{f\}(\mathbf{0}) = f(\mathbf{0})$ for any function f). On images, the position of the origin is arbitrary. For instance, we are not only interested to know the effect of rotations around the origin, but also on rotations centred around any location $\boldsymbol{x}_0$. For an operator $\mathcal{A}$, we denote by $\mathcal{A}_{\boldsymbol{x}_0}$ its shifted version that is centered around $\boldsymbol{x}_0$ instead of $\mathbf{0}$. Mathematically, we have

$\mathcal{A}_{\boldsymbol{x}_0} = \mathcal{T}_{\boldsymbol{x}_0}\mathcal{A}\mathcal{T}_{-\boldsymbol{x}_0}$. We therefore refine the notion of equivariance as follows. We say that a texture operator is *locally equivariant* to the group of transformation $\boldsymbol{A}$ if it commutes with any transformation $\mathcal{A}_{\boldsymbol{x}_0}$ for any $\mathcal{A} \in \boldsymbol{A}$ and any $\boldsymbol{x}_0 \in \mathbb{R}^D$; that is,

$$\mathcal{A}_{\boldsymbol{x}_0}\{\mathcal{G}_n\{f\}\}(\,\cdot\,) = \mathcal{G}_n\{\mathcal{A}_{\boldsymbol{x}_0}\{f\}\}(\,\cdot\,) \quad \forall f,\ \forall \mathcal{A} \in \mathbf{A},\ \forall \boldsymbol{x}_0 \in \mathbb{R}^D. \tag{1.8}$$

We remark that the texture operators introduced in Section 3.1 are precisely the one that are equivariant to translations, due to the sliding property. This has an important consequence: if the texture operator is equivariant to another group of transformations for which the origin plays a central role (such as scalings or rotation), then it is also locally equivariant in the sense of Eq. (1.8).

Strict equivariance is a difficult requirement to achieve in the practical design of operators. Therefore, the equivariance constraint can be approximated to allow robustness of texture measurements to (local) rigid transformations. The desirable approximated operator invariances/equivariances to global/local geometric transformations are also depending on the image analysis task at hand. This is illustrated in Fig. 1.17 where the requirements substantially differ for detecting cars in photographic imagery and for detecting collagen junctions in lung CT. The former requires operator local equivariance to scaling transformations (all cars must be detected regardless of their distance to the viewpoint) and equivariance to translations (important for localizing the positions of the cars), while no equivariance to rotations (important to rule out objects that look like cars but are *e.g.*, upside down). Detecting collagen junctions in lung CT requires operator local equivariance to rotations (all junctions must be detected in spite of their local orientation), equivariance to translations (important for localizing the junctions) but no equivariance to scaling (important to rule out objects that look like junctions but are physically too big or small to be a junction). Since the texture primitives (or textons, see Section 2.4) can have arbitrary local orientations in most biomedical images (see Figures 1.8 and 1.9), it is mostly interesting to design operators that are equivariant to local rotations [53] (see Section 4.1 of Chapter 2).

3.3.2. Invariances of texture measurements

Invariance of texture measurements to geometric transformations are further obtained through the aggregation function. Most of aggregation functions (*e.g.*, integral/summation, max) are adding invariance of the texture measurements over the ROI $\boldsymbol{M}$ in the sense of Eq. (1.6). This is reasonable under the condition that the texture processes are considered stationary (in the relaxed sense, see Section 2.3.1) over $\boldsymbol{M}$. For instance, However, special care must be taken when choosing $\boldsymbol{M}$ to avoid undesirable and destructive side effects of aggregation. This is discussed in Sections 3.2 and 4 of Chapter 2.

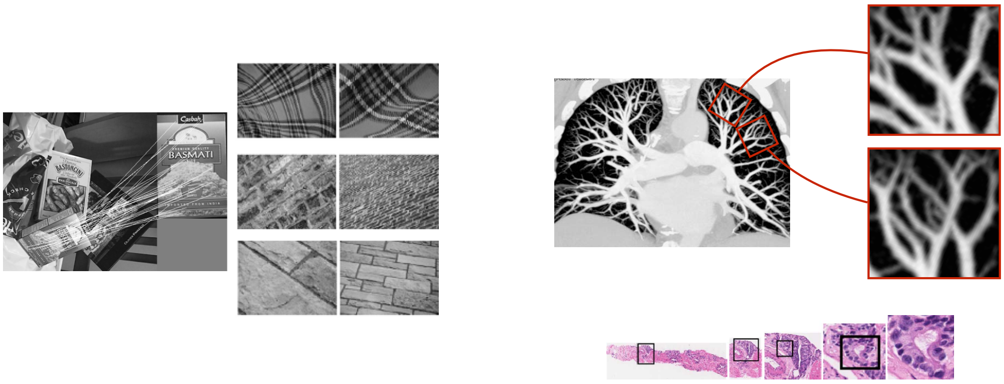


Figure 1.16: Typical geometric transformations of textures encountered in photographic imagery [38] (left) versus biomedical images [25] (right). In most cases, biomedical textures are observed in images with known pixel sizes. As a consequence, texture measures that are robust to translations and rotations (rigid transformations) but sensitive to changes in image scale will yield optimal descriptors. As opposed to photographic image analysis, it is not desirable to enforce any form of scale invariance which truly entails the risk of regrouping patterns of different nature.

Figure 1.17: The requirements in terms of geometric invariance/equivariance of image analysis methods substantially differ depending on the image analysis task at hand. Left: detecting cars in photographic imagery[7] versus right: detecting collagen junctions in lung CT [18].

3.3.3. Non-geometric invariances

Similarly to the general categories of feature design mentioned in Section 3.1, designing texture operators that are invariant or equivariant to certain geometric transformations can be considered as *handcrafted* because it involves prior knowledge of biomedical texture variants. However, more subtle intra-class variations are caused by the diversity of *e.g.*, biology, anatomy, subject age [19]. In this context, texture operators can be trained using machine learning to respond invariantly to subtle intra-

---

[7] `https://www.youtube.com/watch?v=xVwsr9p3irA`, as of March 7 2017.

class variation of the data. Deep CNNs have recently shown to perform very well on learning cascades of non-linear operators and aggregation functions[8] minimizing intra-class variations and maximizing inter-class separability [8, 43]. However, they solely rely on class representations available in the training set, *i.e.*, the collection of texture instances (realizations) with ground truth (labels) available. The latter are difficult to acquire in practice for biomedical images, and data augmentation has been used to include geometric handcrafted invariances to deep CNNs [57] (see Section 2.3.4 of Chapter 3). Transfer learning has also been proposed to reuse deep networks trained on very large and publicly available databases (*e.g.*, ImageNet [13]) for other image recognition tasks [55, 72] (see Section 4.3.3 of Chapter 4). In this particular context, using networks that were trained with images acquired with digital cameras (ImageNet) on biomedical images carries the risk of introducing undesired operator robustness to image scale for instance, discarding a strong discriminative property.

4. Conclusions

In this chapter, we present the foundations of texture processing for biomedical image analysis. We begin with a generic definition for the type of textures encompassing those observed in biomedical imaging (Section 2). We clarify the difference between intensity and texture in ROIs, where the former relies on the statistical distribution of pixel values, and the latter is characterized by the spatial transitions between the pixel values (see Fig. 1.1). The direction, scale and order of these spatial transitions were found to be fundamental properties of biomedical texture and are naturally described in the Fourier domain. From a mathematical point of view, we defined biomedical texture functions as realizations of intricate and non-stationary stochastic processes. When compared to general photographic image analysis, the acquisition devices and protocols in biomedical imaging yield data with well controlled and standardized fundamental properties such as pixel size and intensity, as well as image orientation and sample/patient position. Texture analysis challenges are therefore specific to the domain of biomedical imaging and requires adequate methods for obtaining optimal results.

Second, we introduced a general problem formulation for BTA in Section 3. It essentially consisted of considering that any biomedical texture analysis approach can be characterized by a series of local texture operators and regional aggregation functions. Operators can be handcrafted to highlight desired properties of the input texture function such as spatial scales in a well defined frequency band, image gradients along horizontal directions, co-occurrences, local binary patterns or circular frequencies. They

---

[8] The forward function of image operators results from the composition of linear and slightly non-linear operations (*e.g.*, rectified linear unit, ReLU).

can also be learned from data to yield optimal texture representation for reconstruction or discrimination. Image normalization and operator invariance/equivariance to rigid transformations were found to be fundamental for BTA. Aggregation functions are required to summarize the responses of operators over ROIs. It allows obtaining collections of scalar-valued texture measurements that can be used as quantitative imaging biomarkers. The latter can be further combined with other -omics and patient data to allow precision and personalized medicine by predicting diagnosis, treatment response, as well as to enable biomedical discovery (see Chapter 8).

The challenges of multi-scale and multi-directional biomedical texture analysis are further developed in Chapter 2. A qualitative comparison of popular approaches in terms of the proposed general problem formulation introduced in this chapter is discussed in Chapter 3.

Acknowledgments

This work was supported by the Swiss National Science Foundation (under grant PZ00P2_154891), and the CIBM.

Bibliography

1. A. Ahmadian and A. Mostafa. An efficient texture classification algorithm using Gabor wavelet. In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 1, pages 930–933, September 2003.
2. Katherine P. Andriole, Jeremy M. Wolfe, and Ramin Khorasani. Optimizing Analysis, Visualization and Navigation of Large Image Data Sets: One 5000-Section CT Scan can ruin your whole day. *Radiology*, 259(2):346–362, May 2011.
3. Zelena Anne Aziz, Athol U Wells, David M Hansell, G A Bain, Susan J Copley, Sujal R Desai, Stephen M Ellis, Fergus V Gleeson, Sisa Grubnic, Andrew G Nicholson, S P Padley, K S Pointon, J H Reynolds, R J Robertson, and Michael B Rubens. HRCT diagnosis of diffuse parenchymal lung disease: inter-observer variation. *Thorax*, 59(6):506–511, jun 2004.
4. B Julesz and and R A Schumer. Early Visual Perception. *Annual Review of Psychology*, 32(1):575–627, 1981.
5. C. Blakemore and F. W. Campbell. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology*, 203(1):237–260, 1969.
6. Malcolm Bourne. *Food texture and viscosity: concept and measurement*. Academic press, 2002.
7. A.C. Bovik, M. Clark, and W.S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):55–73, 1990.
8. Joan Bruna and Stephane G. Mallat. Invariant Scattering Convolution Networks . *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1872–1886, 2013.
9. Nicolas Chenouard and Michael Unser. 3D steerable wavelets and monogenic analysis for bioimaging. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 2132–2135, April 2011.
10. Pol Cirujeda, Yashin Dicente Cid, Henning Müller, Daniel Rubin, Todd A Aguilera, Billy W Loo Jr., Maximilian Diehn, Xavier Binefa, and Adrien Depeursinge. A 3–D Riesz–Covariance Texture Model for Prediction of Nodule Recurrence in Lung CT. *IEEE Transactions on Medical Imaging*, 2016.
11. G. Collewet, M. Strzelecki, and F. Mariette. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magnetic Resonance Imaging*, 22(1):81–91, 2004.

12. Fergus Davnall, Connie S. P. Yip, Gunnar Ljungqvist, Mariyah Selmi, Francesca Ng, Bal Sanghera, Balaji Ganeshan, Kenneth A. Miles, Gary J. Cook, and Vicky Goh. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? *Insights into Imaging*, 3(6):573–589, 2012.
13. Jia Deng, Wei Dong, R Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition*, CVPR 2009, pages 248–255, 2009.
14. Adrien Depeursinge, Antonio Foncubierta-Rodríguez, Dimitri Van De Ville, and Henning Müller. Multiscale lung texture signature learning using the Riesz transform. In *Medical Image Computing and Computer–Assisted Intervention MICCAI 2012*, volume 7512 of *Lecture Notes in Computer Science*, pages 517–524. Springer Berlin / Heidelberg, October 2012.
15. Adrien Depeursinge, Antonio Foncubierta-Rodríguez, Dimitri Van De Ville, and Henning Müller. Rotation–covariant texture learning using steerable Riesz wavelets. *IEEE Transactions on Image Processing*, 23(2):898–908, February 2014.
16. Adrien Depeursinge, Antonio Foncubierta-Rodríguez, Dimitri Van De Ville, and Henning Müller. Three-dimensional solid texture analysis and retrieval in biomedical imaging: review and opportunities. *Medical Image Analysis*, 18(1):176–196, 2014.
17. Adrien Depeursinge, Camille Kurtz, Christopher F Beaulieu, Sandy Napel, and Daniel L Rubin. Predicting visual semantic descriptive terms from radiological image data: Preliminary results with liver lesions in CT. *IEEE Transactions on Medical Imaging*, 33(8):1–8, August 2014.
18. Adrien Depeursinge, Zsuzsanna Püspöki, John-Paul Ward, and Michael Unser. Steerable wavelet machines (SWM): Learning moving frames for texture classification. *IEEE Transactions on Image Processing*, 26(4):1626–1636, 2017.
19. Adrien Depeursinge, Daniel Racoceanu, Jimison Iavindrasana, Gilles Cohen, Alexandra Platon, Pierre-Alexandre Poletti, and Henning Müller. Fusing visual and clinical information for lung tissue classification in high–resolution computed tomography. *Artificial Intelligence in Medicine*, 50(1):13–21, September 2010.
20. Thomas M Deserno. *Fundamentals of Biomedical Image Processing*. Springer, 2011.
21. A Dumas, M Brigitte, M F Moreau, F Chrtien, M F Basl, and D Chappard. Bone mass and microarchitecture of irradiated and bone marrow-transplanted mice: influences of the donor strain. *Osteoporosis International*, 20(3):435–443, 2009.
22. Mary M Galloway. Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, 4(2):172–179, 1975.
23. Mehrdad J Gangeh, Ali Ghodsi, and Mohamed S Kamel. Dictionary Learning in Texture Classification. In *Proceedings of the 8th international conference on Image analysis and recognition - Volume Part I*, pages 335–343, 2011.
24. Donald F. Gleason, George T. Mellinger, and The Veretans Administration Cooperative Urological Research Group. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *The Journal of Urology*, 167(2):953–958, 2002.
25. Metin N. Gurcan, Laura E. Boucheron, Ali Can, Anant Madabhushi, Nasir M. Rajpoot, and Bulent Yener. Histopathological image analysis: A review. *IEEE Reviews in Biomedical Engineering*, 2:147–171, 2009.
26. Mark A. Haidekker. *Texture Analysis*, pages 236–275. John Wiley & Sons, Inc., 2010.
27. Robert M Haralick. Statistical and Structural Approaches to Texture. *Proceedings of the IEEE*, 67(5):786–804, may 1979.
28. Aapo Hyvärinen, Jarmo Hurri, and Patrik O. Hoyer. Energy Correlations and Topographic Organization. In *Natural Image Statistics: A Probabilistic Approach to Early Computational Vision*, pages 239–261. Springer London, London, 2009.
29. B Julesz, E N Gilbert, L A Shepp, and H L Frisch. Inability of Humans to Discriminate between Visual Textures That Agree in Second-Order Statistics—Revisited. *Perception*, 2(4):391–405, 1973.
30. Bela Julesz. Visual Pattern Discrimination. *IRE Transactions on Information Theory*, 8(2):84–92, feb 1962.
31. Bela Julesz. A theory of preattentive texture discrimination based on first-order statistics of textons. *Biological Cybernetics*, 41(2):131–138, 1981.

32. Bela Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, 290(5802):91–97, mar 1981.
33. Elizabeth A. Kidd and Perry W. Grigsby. Intratumoral Metabolic Heterogeneity of Cervical Cancer. *Clinical Cancer Research*, 14(16):5236–5241, 2008.
34. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
35. Virendra Kumar, Yuhua Gu, Satrajit Basu, Anders Berglund, Steven A. Eschrich, Matthew B. Schabath, Kenneth Forster, Hugo J. W. L. Aerts, Andre Dekker, David Fenstermacher, Dmitry B. Goldgof, Lawrence O. Hall, Philippe Lambin, Yoganand Balagurunathan, Robert A. Gatenby, and Robert J. Gillies. Radiomics: the process and the challenges. *Magnetic Resonance Imaging*, 30(9):1234–1248, 2012.
36. Curtis P. Langlotz. RadLex: A new method for indexing online educational materials. *Radiographics*, 26(6):1595–1597, 2006.
37. Elizabeth Lazarus, Martha B Mainiero, Barbara Schepps, Susan L Koelliker, and Linda S Livingston. BI-RADS Lexicon for US and Mammography: Interobserver Variability and Positive Predictive Value. *Radiology*, 239(2):385–391, 2006.
38. Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. A Sparse Texture Representation Using Local Affine Regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1265–1278, aug 2005.
39. Weili Lin, Karthikeyan Kuppusamy, E. Mark Haacke, and Harold Burton. Functional MRI in human somatosensory cortex activated by touching textured surfaces. *Journal of Magnetic Resonance Imaging*, 6(4):565–572, jul 1996.
40. David G Lowe. Distinctive Image Features from Scale–Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
41. Marc Macenko, Marc Niethammer, J S Marron, David Borland, John T Woosley, Xiaojun Guan, Charles Schmitt, and Nancy E Thomas. A Method for Normalizing Histology Slides for Quantitative Analysis. In *Proceedings of the Sixth IEEE International Conference on Symposium on Biomedical Imaging: From Nano to Macro*, ISBI’09, pages 1107–1110, Piscataway, NJ, USA, 2009. IEEE Press.
42. Julien Mairal, Francis Bach, J Ponce, Guillermo Sapiro, and Andrew Zisserman. Supervised dictionary learning. *Advances in Neural Information Processing Systems*, pages 1033–1040, 2008.
43. Stéphane Mallat. Understanding deep convolutional networks. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), mar 2016.
44. Diego Marcos, Michele Volpi, Nikos Komodakis, and Devis Tuia. Rotation equivariant vector field networks. 2016.
45. David P McGovern, Kevin S Walsh, Jason Bell, and Fiona N Newell. Individual differences in context-dependent effects reveal common mechanisms underlying the direction aftereffect and direction repulsion. *Vision Research*, pages –, 2016.
46. Timo Ojala, Topi Mäenpää, Matti Pietikäinen, Jaakko Viertola, Juha Kyllönen, and Sami Huovinen. Outex – new framework for empirical evaluation of texture analysis algorithms. In *16th International Conference on Pattern Recognition*, volume 1 of *ICPR*, pages 701–706. IEEE Computer Society, aug 2002.
47. Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. Multiresolution gray–scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, jul 2002.
48. Fanny Orlhac, Christophe Nioche, Michael Soussan, and Irene Buvat. Understanding changes in tumor textural indices in PET: a comparison between visual assessment and index values in simulated and patient data. *Journal of Nuclear Medicine*, October 2016.
49. Fanny Orlhac, Michaël Soussan, Jacques-Antoine Maisonobe, Camilo A Garcia, Bruno Vanderlinden, and Irène Buvat. Tumor Texture Analysis in 18F-FDG PET: Relationships Between Texture Parameters, Histogram Indices, Standardized Uptake Values, Metabolic Volumes, and Total Lesion Glycolysis. *Journal of Nuclear Medicine*, 55(3):414–422, 2014.

50. Manos Papadakis, G Gogoshin, Ioannis A Kakadiaris, Donald J Kouri, and David K Hoffman. Nonseparable radial frame multiresolution analysis in multidimensions and isotropic fast wavelet algorithms. In *Proc. SPIE Wavelets: Applications in Signal and Image Processing X*, volume 5207, pages 631–642, aug 2003.
51. Maria Petrou. Texture in biomedical images. In Thomas M. Deserno, editor, *Biomedical Image Processing*, pages 157–176. Springer-Verlag Berlin Heidelberg, 2011.
52. Maria Petrou and Pedro García Sevilla. *Image processing: dealing with texture*. Wiley, 2006.
53. Zsuzsanna Püspöki, Martin Storath, Daniel Sage, and Michael Unser. Transforms and Operators for Directional Bioimage Analysis: A Survey. In Winnok H De Vos, Sebastian Munck, and Jean-Pierre Timmermans, editors, *Focus on Bio-Image Informatics*, volume 219, pages 69–93. Springer International Publishing, Cham, 2016.
54. Gwénolé Quellec, Mathieu Lamard, Pierre Marie Josselin, Guy Cazuguel, Béatrice Cochener, and Christian Roux. Optimal wavelet transform for the detection of microaneurysms in retina photographs. *IEEE transactions on medical imaging*, 27(9):1230–1241, 2008.
55. Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. CNN features off-the-shelf: An astounding baseline for recognition. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 512–519, 2014.
56. P. J. Robinson. Radiology’s Achilles’ heel: error and variation in the interpretation of the Rontgen image. *British Journal of Radiology*, 70(839):1085–1098, 1997.
57. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Nassir Navab, Joachim Hornegger, William M Wells, and Alejandro F Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2015*, volume 9351 of *Lecture Notes in Computer Science*, pages 234–241. Springer International Publishing, 2015.
58. Jacob Scharcanski. Stochastic texture analysis for monitoring stochastic processes in industry. *Pattern Recognition Letters*, 26(11):1701–1709, 2005.
59. W Schiff and E Foulke. *Tactual Perception: A Sourcebook*. Paperback Re-issue. Cambridge University Press, 1982.
60. Laurent Sifre and Stéphane Mallat. Rigid-Motion Scattering for Texture Classification. *Submitted to International Journal of Computer Vision*, abs/1403.1:1–19, 2014.
61. J Sivic, B C Russell, A A Efros, A Zisserman, and W T Freeman. Discovering objects and their location in images. In *Tenth IEEE International Conference on Computer Vision, 2005. ICCV 2005.*, volume 1, pages 370–377, oct 2005.
62. Ronald M. Summers. Texture analysis in radiology: Does the emperor have no clothes? *Abdominal Radiology*, 2016.
63. Alina Surmacka Szczesniak. Texture is a sensory property. *Food Quality and Preference*, 13(4):215–225, 2002.
64. Bart M ter Haar Romeny. Multi–Scale and Multi–Orientation Medical Image Analysis. In Thomas M Deserno, editor, *Biomedical Image Processing*, pages 177–196. Springer-Verlag Berlin Heidelberg, 2011.
65. P Thévenaz, T Blu, and M Unser. Image Interpolation and Resampling. In I N Bankman, editor, *Handbook of Medical Imaging, Processing and Analysis*, chapter 25, pages 393–420. Academic Press, San Diego CA, USA, 2000.
66. Guillaume Thibault, Bernard Fertil, Claire Navarro, Sandrine Pereira, Pierre Cau, Nicolas Levy, Jean Sequeira, and Jean-luc Mari. Texture Indexes and Gray Level Size Zone Matrix Application to Cell Nuclei Classification. *Pattern Recognition and Information Processing*, pages 140–145, 2009.
67. Georgia D. Tourassi. Journey toward Computer–aided Diagnosis: Role of Image Texture Analysis. *Radiology*, 213(2):317–320, jul 1999.
68. Georgia D. Tourassi, Sophie Voisin, Vincent Paquit, and Elizabeth A Krupinski. Investigating the link between radiologists’ gaze, diagnostic decision, and image content. *Journal of the American Medical Informatics Association*, 20(6):1067–1075, nov 2013.
69. Manik Varma and Andrew Zisserman. A Statistical Approach to Texture Classification from Single Images. *International Journal of Computer Vision*, 62(1-2):61–81, 2005.

70. C Vonesch, F Stauber, and M Unser. Steerable PCA for rotation-invariant image recognition. *SIAM Journal on Imaging Sciences*, 8(3):1857–1873, 2015.
71. T. Aisling Whitaker, Cristina Simões-Franklin, and Fiona N. Newell. Vision and touch: Independent or integrated systems for the perception of texture? *Brain Research*, 1242:59–72, 2008.
72. Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems 27 (Proceedings of NIPS)*, 27:1–9, 2014.
73. Song-Chun Zhu, Cheng-En Guo, Yizhou Wang, and Zijian Xu. What are Textons? *International Journal of Computer Vision — Special Issue on Texture Analysis and Synthesis*, 62(1–2):121–143, apr 2005.