

Chapter 1

AFFINE INVARIANT HYPERSPECTRAL IMAGE DESCRIPTORS BASED UPON HARMONIC ANALYSIS

Pattaraporn Khuwuthyakorn^{1,2,3}
Antonio Robles-Kelly^{1,2}
Jun Zhou^{1,2}

¹*RSISE, Bldg.115, Australian National University, Canberra ACT 0200, Australia*

²*National ICT Australia (NICTA) *, Locked Bag 8001, Canberra ACT 2601, Australia*

³*CRC for National Plant Biosecurity (CRCNPB) †, Bruce, ACT, 2617, Australia*

Abstract This chapter focuses on the problem of recovering a hyperspectral image descriptor based upon harmonic analysis. It departs from the use of integral transforms to model hyperspectral images in terms of probability distributions. This provides a link between harmonic analysis and affine geometric transformations between object surface planes in the scene. Moreover, the use of harmonic analysis permits the study of these descriptors in the context of Hilbert spaces. This, in turn, provides a connection to functional analysis to capture the spectral cross-correlation between bands in the image for the generation of a descriptor with a high energy compaction ratio. Thus, descriptors can be computed based upon orthogonal bases capable of capturing the space and wavelength correlation for the spectra in the hyperspectral imagery under study. We illustrate the utility of our descriptor for purposes of object recognition on a hyperspectral image dataset of real-world objects and compare our results to those yielded using an alternative.

Keywords: Hyperspectral Image Descriptor, Harmonic Analysis, Heavy-tailed distributions

*NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

†P. Khuwuthyakorn was partially supported by the CRCNPB grant: CRC60075.

1. Introduction

With the advent and development of new sensor technologies, it is now possible to capture image data in tens or hundreds of wavelength-resolved bands covering a broad spectral range. Compared to traditional monochrome and trichromatic cameras, hyperspectral image sensors provide an information-rich representation of the spectral response for the material under study over a number of wavelengths. This has opened-up great opportunities and posed important challenges due to the high dimensional nature of the spectral data. As a result, many classical algorithms in pattern recognition and machine learning have been naturally borrowed and adapted so as to perform feature extraction and classification [21]. Techniques such as Principle Component Analysis (PCA) [18], Linear Discriminant Analysis (LDA)[13], Projection Pursuit [17] and their kernel versions [11] treat raw pixel spectra as input vectors in a higher-dimensional space, where the dimensionality is given by the number of bands. The idea is to recover statistically optimal solutions to the classification problems by reducing the data dimensionality via a projection of the feature space.

The methods above are often used for purposes of recognition based on individual signatures, which in hyperspectral images, represent single pixels. Nonetheless each signature is generally related to material chemistry, these methods do not take into account the local structure of the images under study. They rather hinge in the notion that different materials have different characteristic responses as a function of wavelengths which can be used to provide descriptions of the target objects. Thus, raw pixels are often treated as input vectors in high dimensional spaces.

In contrast with the pixel-based methods in hyperspectral imaging, the approaches available for content-based image retrieval often take into account the local structure of the scene. These methods often represent images as a bag of features so as to match query images to those in the database by computing distances between distributions of local descriptors. As a result, trichromatic object and image retrieval and classification techniques [7, 26, 37] are often based upon the summarisation of the image dataset using a codebook of visual words [23, 25, 29].

It is surprising that despite the widespread use of higher-level features for recognition and retrieval of monochromatic and trichromatic imagery, local hyperspectral image descriptors are somewhat under-researched. The use of local image descriptors opens-up great opportunities in recognition and classification tasks. Moreover, the multidimensional nature of local image features and descriptors may be combined to improve performance. For instance, Varma and Ray [41] have used a kernel learning approach to learn the trade-off between discriminative power and invariance of image descriptors in classifi-

cation tasks. Other methods, such as the one in [5], rely upon clustering algorithms to provide improved organisation of the codebook. Other alternatives tend to view the visual words as multidimensional data and, making use of unsupervised learning, exploit similarity information in a graph-theoretic setting. Examples of these are the method presented by Sengupta and Boyer [34] and that developed by Shokounfandeh et. al. [36], which employ information-theoretical criteria to hierarchically structure the dataset under study and pattern recognition methods to match the candidates.

Amongst local image descriptors, texture has found applications not only as a shape queue [14, 40], but has also attracted broad attention for recognition and classification tasks [31]. Moreover, from the shape modelling perspective, static texture planes can be recovered making use of the structural analysis of predetermined texture primitives [1, 16, 19]. This treatment provides an intuitive geometrical meaning to the task of recovering the parameters governing the pose of the object by making use of methods akin to 3D view geometry. For dynamic textures, Sheikh, Haering and Shah [35] have developed an algorithm for recovering the affine geometry making use of motion-magnitude constraints. Pteri and Chetevirkov [6] have characterised dynamic textures using features extracted using normal flows. This builds on the comparative study in [12]. Ghanem and Ahuja [15] have used the Fourier phase to capture the global motion within the texture. Rahman and Murshed [30] have estimated optical flow making use of motion patterns for temporal textures. Otsuka *et al.* [27] have used surface motion trajectories derived from multiple frames in a dynamic texture to recover spatiotemporal texture features.

As mentioned earlier, we focus on the problem of recovering a hyperspectral image descriptor by using harmonic functions to model hyperspectral imagery in terms of probability distributions. This is reminiscent of time-dependent textures, whose probability density functions exhibit first and second order moments which are space and time-shift invariant [10]. For instance, in [28], the characterisation of the dynamic texture under study is obtained using the empirical observations of statistical regularities in the image sequence. In [3], statistical learning is used for purposes of synthesising a dynamic texture based upon an input image sequence. Zhao and Pietikäinen [43] have performed recognition tasks using local binary patterns that fit space-time statistics.

The methods above view time-dependent textures as arising from second-order stationary stochastic processes such as moving tree-leaves, sea waves and rising smoke plumes. We, from another point of view, relate hyperspectral image regions to harmonic kernels to capture a discriminative and descriptive representation of the scene. This provides a principled link between statistical approaches, signal processing methods for texture recognition and shape modelling approaches based upon measures of spectral distortion [24]. The method

also provides a link to affine geometric transformations between texture planes and their analysis in the Fourier domain [4].

The chapter is organised as follows. We commence by exploring the link between harmonic analysis and heavy tailed distributions. We then explore the relationship between distortions over locally planar patches on the object surface and the domain induced by an integral transform over a harmonic kernel. We do this so as to achieve invariance to affine transformations on the image plane. With these technical foundations at hand, we proceed to present our hyperspectral image descriptor by incorporating the cross-correlation between bands. This results in a descriptor based upon orthogonal bases with high information compaction properties which can capture the space and wavelength correlation for the spectra in hyperspectral images. Moreover, as we show later on, the choice of bases or kernel is quite general since it applies to harmonic kernels which span a Hilbert space. We conclude the chapter by demonstrating the utility of our descriptor for purposes of object recognition based upon real-world hyperspectral imagery.

2. Heavy-tailed Distributions

As mentioned earlier, we view hyperspectral images as arising from a probability distribution whose observables or occurrences may have long or heavy tails. This implies that the spectra in the image results in values that can be rather high in terms of their deviation from the image-spectra mean and variance. As a result, our formulation can capture high wavelength-dependent variation in the image. This is important, since it allows us to capture information in our descriptor that would otherwise may be cast as the product of outliers. Thus, we formulate our descriptor so as to model “rare” stationary wavelength-dependent events on the image plane.

Moreover, we view the pixel values of the hyperspectral image as arising from stochastic processes whose moment generating functions are invariant with respect to shifts in the image-coordinates. This means that the mean, covariance, kurtosis, etc. for the corresponding joint probability distribution are required to be invariant with respect to changes of location on the image. Due to our use of heavy tailed distributions, these densities may have high dispersion and, thus, their probability density functions are, in general, governed by further-order moments. These introduces a number of statistical “skewness” variables that allow modeling high variability spectral behaviour.

This is reminiscent of simulation approaches where importance sampling cannot be effected via an exponential changes in measurement due to the fact that the moments are not exponential in nature. This applies to distributions such as the log-normal, Weibull with increasing skewness and regularly varying distributions such as Pareto, stable and log-gamma distributions [2]. More

formally, we formulate the density of the pixel-values for the wavelength λ at the pixel u in the image-band I_λ of the image as random variables \mathcal{Y}_u whose inherent basis $\mathcal{X}_u = \{x_u(1), x_u(2), \dots, x_u(|\mathcal{X}_u|)\}$ is such that

$$P(\mathcal{Y}_u) = \sum_{k=1}^{|\mathcal{X}_u|} P(x_u(k)) \quad (1.1)$$

where, $x_u(k)$ are identically distributed variables and, as usual for probability distributions of real-valued variables, we have written $P(\mathcal{Y}_u) = Pr[y \leq \mathcal{Y}_u]$ for all $y \in \mathfrak{R}$.

In other words, we view the pixel values for each band in the image under study as arising from a family of heavy-tailed distributions whose variance is not necessarily finite. It is worth noting that, for finite variance, the formalism above implies that $P(\mathcal{Y}_u)$ is normally distributed and, as a result, our approach is not exclusive to finite variance distributions, but rather this treatment generalises the stochastic process to a number of independent influences, each of which is captured by the corresponding variable $x_u(k)$.

In practice, the Probability Density Function (PDF) $f(\mathcal{Y}_u)$ is not available in close form. As a result, we can re-parameterise the PDF by recasting it as a function of the variable ζ making use of the characteristic function

$$\psi(\zeta) = \int_{-\infty}^{\infty} \exp(\mathbf{i}\zeta\mathcal{Y}_u) f(\mathcal{Y}_u) d\mathcal{Y}_u \quad (1.2)$$

$$= \exp(\mathbf{i}u\zeta - \gamma|\zeta|^\alpha (1 + \mathbf{i}\beta \text{sign}(\zeta)\varphi(\zeta, \alpha))) \quad (1.3)$$

where $i = \sqrt{-1}$, u is, as before, the pixel-index on the image plane and $\gamma \in \mathfrak{R}^+$ are function parameters, $\beta \in [-1, 1]$ and $\alpha \in (0, 2]$ are the skewness and characteristic exponent, respectively, and $\varphi(\cdot)$ is defined as follows

$$\varphi(\zeta, \alpha) = \begin{cases} \tan(\alpha \frac{\pi}{2}) & \text{if } \alpha \neq 1 \\ -\frac{\pi}{2} \log|\zeta| & \text{if } \alpha = 1 \end{cases} \quad (1.4)$$

For the characteristic function above, some values of α correspond to special cases of the distribution. For instance, $\alpha = 2$ implies a normal distribution, $\beta = 0$ and $\alpha = 1$ corresponds to a Cauchy distribution and, for the Levy distribution we have $\alpha = \frac{1}{2}$ and $\beta = 1$. Thus, nonetheless the formalism above can capture a number of cases in exponential families, it is still quite general in nature so as to allow the modeling of a large number of distributions that may apply to hyperspectral data and whose characteristic exponents α are not of those distributions whose tails are exponentially bounded.

So far, we have limited ourselves to the image plane for a fixed wavelength λ . That is, we have, so far, concentrated on the distribution of spectral values across every wavelength-resolved band in the image. Note that, without loss

of generality, we can extend Equation (1.3) to the wavelength domain, i.e. the spectra of the image across a segment of bands.

This is a straightforward task by noting that the equation above can be viewed as the cross-correlation between the function $f(\mathcal{Y}_u)$ and the exponential given by $\exp(i\lambda\zeta)$. Hence, we can write the characteristic function for the image parameterised with respect to the wavelength λ as follows

$$\vartheta(\lambda) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(i\lambda\zeta) \exp(i\zeta\mathcal{Y}_u) f(\mathcal{Y}_u) d\mathcal{Y}_u d\zeta \quad (1.5)$$

$$= \int_{-\infty}^{\infty} \exp(i\lambda\zeta) \psi(\zeta) d\zeta \quad (1.6)$$

where the second line in the equation above corresponds to the substitution of Equation (1.3) into Equation (1.5).

Equation (1.6) captures the spectral cross-correlation for the characteristic functions for each band. In this manner, we view the characteristic function for the hyperspectral image as a heavy-tailed distribution of another set of heavy-tailed PDFs, which correspond to each of the band in the image. This can also be interpreted as a composition of two heavy-tailed distributions, where Equation (1.3) corresponds to the image-band domain ζ of the image and Equation (1.6) is determined by the wavelength-dependent domain λ .

This composition operation suggests a two-step process for the computation of the image descriptor. Firstly, at the band-level, the information can be represented in a compact fashion making use of harmonic analysis and rendered invariant to geometric distortions on the object surface plane. Secondly, the wavelength-dependent correlation between bands can be computed making use of the operation in Equation (1.6).

3. Harmonic Analysis

In this section, we explore the use of harmonic analysis and the fundamentals of integral transforms [38] to provide a means to the computation of our image descriptor. We commence by noting that Equation (1.2) and Equation (1.5) are characteristic functions obtained via the integral of the product of the function $g(\eta)$, i.e. $f(\mathcal{Y}_u)$ and $\psi(\zeta)$, multiplied by a kernel, given by $\exp(i\lambda\zeta)$ and $\exp(i\zeta\mathcal{Y}_u)$, respectively.

To appreciate this more clearly, consider the function given by

$$F(\omega) = \int_{-\infty}^{\infty} g(\eta) K(\omega, \eta) d\eta \quad (1.7)$$

where $K(\omega, \eta)$ is a harmonic kernel of the form

$$K(\omega, \eta) = \sum_{k=1}^{\infty} a_k \phi_k(\omega) \phi_k(\eta) \quad (1.8)$$

where a_k is the k^{th} real scalar corresponding to the harmonic expansion and $\phi_k(\cdot)$ are orthonormal functions such that $\langle \phi_k(\omega), \phi_n(\eta) \rangle = 0 \forall n \neq k$. Moreover, we consider cases in which the functions $\phi_k(\cdot)$ constitute a basis for a Hilbert space [42] and, therefore, the right-hand side of Equation (1.8) is convergent to $K(\omega, \eta)$ as k tends to infinity.

To see the relation between Equation (1.7) and the equations in previous sections, we can examine $\psi(\zeta)$ in more detail and write

$$\log[\psi(\zeta)] = iu\zeta - |\zeta|^\alpha (1 + i\beta \text{sign}(\zeta) \varphi(\zeta, \alpha)) \quad (1.9)$$

$$= iu\zeta - |\zeta|^\alpha \gamma^* \alpha \exp(-i\beta^* \frac{\pi}{2} \vartheta \text{sign}(\zeta)) \quad (1.10)$$

where $\vartheta = 1 - |1 - \alpha|$ and parameters γ^* and β^* are given by

$$\gamma^* = \left(\frac{\gamma \sqrt{\Omega}}{\cos(\alpha \frac{\pi}{2})} \right)^{\frac{1}{\alpha}} \quad (1.11)$$

$$\beta^* = \frac{2}{\pi \vartheta} \arccos \left(\frac{\cos(\alpha \frac{\pi}{2})}{\sqrt{\Omega}} \right) \quad (1.12)$$

and $\Omega = \cos^2(\alpha \frac{\pi}{2}) + \beta^2 \sin^2(\alpha \frac{\pi}{2})$.

To obtain the kernel for Equation (1.7), we can use Fourier inversion on the characteristic function and, making use of the shorthands defined above, the PDF may be computed via this following equation.

$$f(\mathcal{Y}_u; u, \beta^*, \gamma^*, \alpha) = \frac{1}{\pi \gamma^*} \int_0^\infty \cos \left(\frac{(u - \mathcal{Y}_u)s}{\gamma^*} + s^\alpha \sin(\phi) \right) \exp(-s^\alpha \sin(\phi)) ds \quad (1.13)$$

where $\phi = \frac{\beta^* \pi \eta}{2}$.

This treatment not only opens-up the possibility of functional analysis on the characteristic function using the techniques in the Fourier domain, but also allows the use of other harmonic kernels for compactness and ease of computation. This is due to the fact that, we can view the kernel $K(\omega, \eta)$ as the exponential $\exp(-s^\alpha \sin(\phi))$, whereas the function $g(\eta)$ is given by the cosine term. Thus, we can use other harmonic kernels so as to induce a change of basis without any loss of generality. Actually, the expression above can be greatly simplified making use of the shorthands $A = \frac{(u - \mathcal{Y}_u)}{\gamma^*}$, $\eta = s^\alpha$ and $\omega \eta = As + s^\alpha \sin(\phi)$, which yields

$$s^\alpha \sin(\phi) = \omega \eta - A \eta^{\frac{1}{\alpha}} \quad (1.14)$$

Substituting Equation (1.13) with Equation (1.14), the PDF can be expressed as

$$f(\mathcal{Y}_u; u, \beta^*, \gamma^*, \alpha) = \sqrt{\frac{2}{\pi}} \int_0^\infty \frac{\exp(-\omega \eta + A \eta^{\frac{1}{\alpha}})}{\sqrt{2\pi} \gamma^* \alpha \eta^{\frac{\alpha-1}{\alpha}}} \cos(\omega \eta) d\eta \quad (1.15)$$

where the kernel then becomes

$$K(\omega, \eta) = \cos(\omega\eta) \quad (1.16)$$

This can be related, in a straightforward manner, to the Fourier cosine transform (FCT) of the form

$$F(\omega) = \sqrt{\frac{2}{\pi}} \int_0^\infty \frac{\exp(-\omega\eta + \frac{(u-\mathcal{Z}_u)}{\gamma^*} \eta^{\frac{1}{\alpha}})}{\sqrt{2\pi}\gamma^* \alpha \eta^{\frac{\alpha-1}{\alpha}}} \cos(\omega\eta) d\eta \quad (1.17)$$

which is analogous to the expression in Equation (1.13). Nonetheless, the transform above does not have imaginary coefficients. This can be viewed as a representation in the power rather than in the phase spectrum. Moreover, it has the advantage of compacting the spectral information in the lower-order Fourier terms, i.e. those for which ω is close to the origin. This follows the strong ‘‘information compaction’’ property of FCTs introduced in [32] and assures a good trade-off between discriminability and complexity.

It is worth stressing that, due to the harmonic analysis treatment given to the problem in this section, other kernels may be used for purposes of computing other integral transforms [38] spanning Hilbert Spaces. These include wavelets and the Mellin ($K(\omega, \eta) = \eta^{\omega-1}$) and Hankel transforms. In fact, other Kernels may be obtained by performing an appropriate substitution on the term $\cos(\omega\eta)$. Note that, for purposes of our descriptor recovery, we will focus on the use of the cosine transform above. This is due to the information compaction property mentioned earlier and the fact that computational methods for the efficient recovery of the FCT are readily available.

4. Invariance to Affine Distortions

Having introduced the notion of the harmonic analysis and shown how the probability density function can be recovered using a Fourier transform, we now focus on relation between distortions on the object surface plane and the Fourier domain. To this end, we follow [4] and relate the harmonic kernel above to affine transformations on the object locally planar shape. As mentioned earlier, the function $f(\mathcal{Z}_u)$ corresponds to the band-dependent component of the image and, as a result, its prone to affine distortion. This hinges in the notion that a distortion on the object surface will affect the geometric factor for the scene, but not its photometric properties. In other words, the material index of refraction, roughness, etc. remains unchanged, whereas the geometry of the reflective process does vary with respect to affine distortions on the image plane. The corresponding 2D integral transform of the function $f(\mathcal{Z}_u)$ which, as introduced in the previous sections, corresponds to the pixel values

for the image-band I_λ in the image under study is given by

$$F(\xi) = \int_{\Gamma} f(\mathcal{D}_u)K(\xi^T, u)du \quad (1.18)$$

where $u = [x, y]^T$ is the vector of two-dimensional coordinates for the compact domain $\Gamma \in \mathfrak{R}^2$ and, in the case of the FCT, $K(\xi^T, u) = \cos(2\pi(\xi^T u))$.

In practice, the coordinate-vectors u will be given by discrete quantities on the image lattice. For purposes of analysis, we consider the continuous case and note that the affine coordinate transformation can be expressed in matrix notation as follows

$$u' = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c \\ h \end{bmatrix} \quad (1.19)$$

This observation is important because we can relate the kernel for the FCT in Equation (1.18) to the transformed coordinate $u' = [x', y']^T$. Also, note that, for patches centered at keypoints in the image, the locally planar object surface patch can be considered devoid of translation. Thus, we can set $f = c = 0$ and write

$$\xi^T u = \xi^T \begin{bmatrix} x \\ y \end{bmatrix} \quad (1.20)$$

$$= \begin{bmatrix} \xi_x & \xi_y \end{bmatrix} \begin{bmatrix} a & b \\ d & e \end{bmatrix}^{-1} \begin{bmatrix} x' \\ y' \end{bmatrix} \quad (1.21)$$

$$= \frac{1}{ae - bd} \begin{bmatrix} (e\xi_x - d\xi_y) & (-b\xi_x + a\xi_y) \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \quad (1.22)$$

where $\xi = [\xi_x, \xi_y]^T$ is the vector of spectral indexes for the 2D integral transform.

Hence, after some algebra, and using the shorthand $\Delta = (ae - bd)$, we can show that for the coordinates u' , the integral transform is given by

$$F(\xi) = \frac{1}{|\Delta|} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathcal{D}_{u'})K\left(\frac{1}{\Delta}[(e\xi_x - d\xi_y), (b\xi_x - a\xi_y)], [x', y']^T\right) dx' dy' \quad (1.23)$$

This implies that

$$F(\xi) = \frac{1}{|\Delta|} F(\xi') \quad (1.24)$$

where ξ' is the “distorted” analogue of ξ . The distortion matrix \mathbb{T} is such that

$$\xi = \begin{bmatrix} \xi_x \\ \xi_y \end{bmatrix} = \begin{bmatrix} a & d \\ b & e \end{bmatrix} \begin{bmatrix} \xi'_x \\ \xi'_y \end{bmatrix} = \mathbb{T}\xi' \quad (1.25)$$

As a result, from Equation (1.23), we can conclude that the effect of the affine coordinate transformation matrix \mathbb{T} is to produce a distortion equivalent

to $(\mathbb{T}^T)^{-1}$ in the ξ domain for the corresponding integral transform. This observation is an important one since it permits achieving invariance to affine transformations on the locally planar object surface patch. This can be done in practice via a ξ -domain distortion correction operation of the form

$$F(\xi) = (\mathbb{T}^T)^{-1}F(\xi') \quad (1.26)$$

5. Descriptor Construction

With the formalism presented in the previous sections, we now proceed to elaborate further on the descriptor computation. Succinctly, this is a two-step process. Firstly, we compute the affine-invariant 2D integral transform for every band in the hyperspectral image under study. This is equivalent to computing the band-dependent component of the characteristic function $\psi(\zeta)$. Secondly, we capture the wavelength dependent behaviour of the hyperspectral image by computing the cross-correlation with respect to the spectral domain for the set of distortion-invariant integral transforms. By making use of the FCT kernel, in practice, the descriptor becomes an FCT with respect to the band index for the cosine transforms corresponding to wavelength-resolved image in the sequence.

Following the rationale above, we commence by computing the distortion invariant integral transform for each band in the image. To do this, we use Equation (1.26) to estimate the distortion matrix with respect to a predefined reference. Here, we employ the peaks of the power spectrum and express the relation of the integral transforms for two locally planar image patches, i.e. the one corresponding to the reference and that for the object under study. We have done this following the notion that a blob-like shape composed of a single transcendental function on the image plane would produce two peaks in the Fourier domain. That is, we have set, as our reference, a moment generating function arising from a cosine on a plane perpendicular to the camera.

Let the peaks of the power spectrum for two locally planar object patches be given by \mathbf{U}_A and \mathbf{U}_B . Those for the reference are \mathbf{U}_R . As a result, the matrices \mathbf{U}_A , \mathbf{U}_B and \mathbf{U}_R are such that each of their columns correspond to the x - y coordinates for one of the two peaks in the power spectrum. These relations are given by

$$\mathbf{U}_A = (\mathbb{T}_A^T)^{-1}\mathbf{U}_R \quad (1.27)$$

$$\mathbf{U}_B = (\mathbb{T}_B^T)^{-1}\mathbf{U}_R \quad (1.28)$$

Where $\mathbb{T}_A : \mathbf{U}_A \Rightarrow \mathbf{U}_R$ and $\mathbb{T}_B : \mathbf{U}_B \Rightarrow \mathbf{U}_R$ are the affine coordinate transformation matrices of the planar surface patches under consideration.

Note that, this is reminiscent of the shape-from-texture approaches hinging in the use of the Fourier transform for the recovery of the local distortion matrix

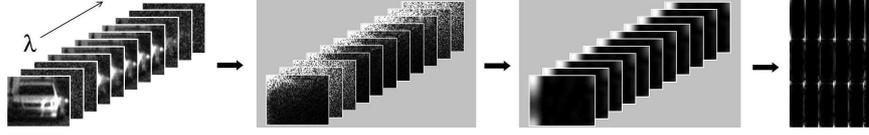


Figure 1.1. From left-to-right: hyperspectral texture, the band-wise FCT, the distortion invariant cosine transforms for every band in the image and the raster scanned 3D matrix \mathbf{V} .

[33]. Nonetheless, in [33], the normal is recovered explicitly making use of the Fourier transform, whereas here we employ the integral transform and aim at relating the FCTs for the two locally planar patches with that of the reference. We can do this making use of the composition operation given by

$$\mathbf{U}_B = (\mathbb{T}_A \mathbb{T}_B^{-1})^T \mathbf{U}_A \quad (1.29)$$

$$= \Phi \mathbf{U}_A \quad (1.30)$$

where $\Phi = (\mathbb{T}_A \mathbb{T}_B^{-1})^T$ is the distortion matrix. This matrix represents the distortion of the power spectrum of \mathbf{U}_A with respect to \mathbf{U}_B .

In practice, note that, if \mathbf{U}_R is known and fixed for every locally planar patch, we can use the shorthands $\mathbb{T}_A^T = \mathbf{U}_R \mathbf{U}_A^{-1}$ and $(\mathbb{T}_B^T)^{-1} = \mathbf{U}_B \mathbf{U}_R^{-1}$ to write

$$\Phi = (\mathbf{U}_R \mathbf{U}_A^{-1})(\mathbf{U}_B \mathbf{U}_R^{-1}) \quad (1.31)$$

Which contrasts with other methods in the fact that, for our descriptor computation, we do not recover the principal components of the local distortion matrix, but rather compute the matrix Φ directly through the expression above. Thus, we can construct a band-level matrix of the form

$$\mathbf{V} = [F(I_1)^* | F(I_2)^* | \dots | F(I_{|\mathbb{I}|})^*] \quad (1.32)$$

which is the concatenation of the affine invariant integral transforms $F(\cdot)^*$ for the band-resolved locally planar object surface patches in the image. Moreover, we render the band-level integral transform invariant to affine transformations making use of the reference peak matrix \mathbf{U}_R such that the transform for the frame indexed t is given by

$$F(I_R) = F(I_t)^* \Phi_t^{-1} \quad (1.33)$$

where Φ_t^{-1} is the matrix which maps the transform for the band corresponding to the wavelength λ to the transform $F(I_R)$ for the reference plane. Here, as mentioned earlier, we have used as reference the power spectrum given by two peaks rotated 45° about the upper left corner of the 2D FCT. The reference FCT is shown in Figure 1.2.

Note that, since we have derived our descriptor based upon the properties of integral transforms and Hilbert spaces, each element of the matrix \mathbf{V} can

be considered as arising from the inner product of a set of orthonormal vectors. Moreover, from a harmonic analysis perspective, the elements of \mathbf{V} are represented in terms of discrete wave functions, over an infinite number of elements [20]. This is analogue to the treatment given to time series in signal processing, where the variance of the signal is described based upon spectral density. Usually, the variance estimations are performed by using Fourier transform methods [39]. Thus, we can make use of the discrete analogue of Equation (1.6) so as to recover the k^{th} coefficient for the image descriptor \mathfrak{G} , which becomes

$$\mathfrak{G}_k = F(\mathbf{V}) = \sum_{n=0}^{|\mathbb{I}|-1} F(I_n) * K\left(\frac{\pi}{|\mathbb{I}|}\left(n + \frac{1}{2}\right), \left(k + \frac{1}{2}\right)\right) \quad (1.34)$$

where $|\mathfrak{G}| = |\mathbb{I}|$ and, for the FCT, the harmonic kernel above becomes

$$K\left(\frac{\pi}{|\mathbb{I}|}\left(n + \frac{1}{2}\right), \left(k + \frac{1}{2}\right)\right) = \cos\left(\frac{\pi}{|\mathbb{I}|}\left(n + \frac{1}{2}\right)\left(k + \frac{1}{2}\right)\right) \quad (1.35)$$

6. Implementation Issues

Now, we turn our attention to the computation of the descriptor and provide further discussion on the previous developments. To this end, we illustrate, in Figure 1.1, the step-sequence of the descriptor computation procedure. We depart from a series of bands in the image and compute the band-by-band

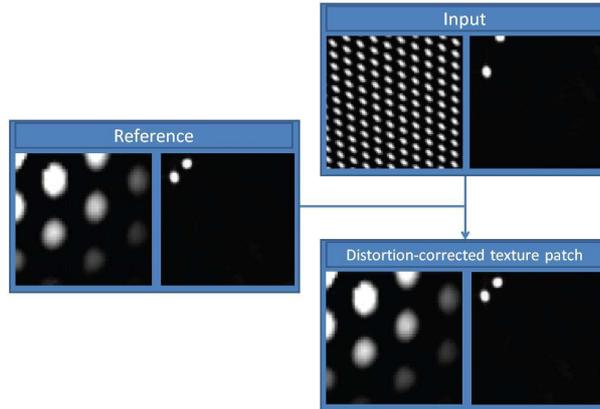


Figure 1.2. Example of reference, input and distortion corrected single-band textures. In the panels, the left-hand image shows the single-band reference texture whereas the right-hand panel shows the power spectrum of the distorted and affine corrected FCT for the texture under study.

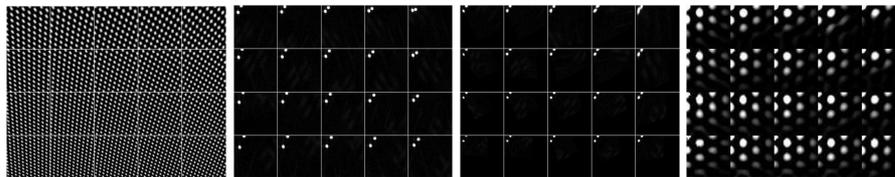


Figure 1.3. From left-to-right: Affine distortion of a sample single-band image; FCT of the image patches in the left-hand panel, distortion-corrected power spectrums for the FCTs in the second panel and inverse FCTs for the power spectrum in the third panel.

FCT. With the band FCTs at hand, we apply the distortion correction approach presented in the previous sections so as to obtain a “power-aligned” series of cosine transforms that can be concatenated into \mathbf{V} . The descriptor is then given by the cosine transform of \mathbf{V} over the wavelength-index. Note that the descriptor will be three-dimensional in nature, with size $N_x \times N_y \times N_\lambda$, where N_x and N_y are the sizes of the locally planar object patches in the image lattice and N_λ is equivalent to the wavelength range for the hyperspectral image bands. In the figure, for purposes of visualisation, we have raster-scanned the descriptor so as to display a 2D matrix whose rows correspond to the wavelength-indexes of the hyperspectral image under study.

We now illustrate the distortion correction operation at the band level in 1.2. In the panels, we show the reference, corrected and input image regions in their spatial and frequency domains. Note that, at input, the textured planes show an affine distortion which affects the distribution of the peaks in its power spectrum.

Moreover, in Figure 1.3, we show a sample textured plane which has been affinely distorted. In the figure, we have divided the distorted input texture into patches that are assumed to be locally planar. We then apply the FCT to each of these patches, represented in the form of a lattice on the input image in the left-hand panel. The corresponding power spectrums are shown in the second column of the figure. Note that, as expected, the affine distortions produce a displacement on the power spectrum peaks. In the third panel, we show the power spectrums after the matrix Φ has been recovered and multiplied so as to



Figure 1.4. Hyperspectral wavelength-resolved bands corresponding to 662nm for six sample objects in our dataset. From left-to-right: plastic dinosaurs and animals, miniature cars, fluffy dolls, plastic blocks, wooden blocks and coins.

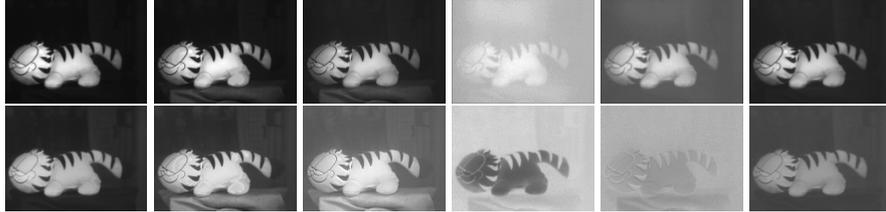


Figure 1.5. From left-to-right: sample hyperspectral images of a fluffy toy at a number of wavelength-resolved bands, i.e. $\lambda = \{550nm, 640nm, 730nm, 820nm, 910nm, 1000nm\}$. The top row shows the bands corresponding to the uncalibrated images and the bottom row shows the calibrated bands.

obtain the corrected FCTs given by $F(\cdot)^*$. The distortion corrected textures in the spatial domain are shown in the right-most panel in the figure. These have been obtained by applying the inverse cosine transform to the power spectrums in the third column. Note that, from both, the corrected power spectrums and the inverse cosine transforms, we can conclude that the correction operation can cope with large degrees of shear in the input texture-plane patches.

7. Experiments

Having presented our image descriptor in the previous sections, we now illustrate its utility for purposes of hyperspectral image categorisation. To this end, we employ a dataset of hyperspectral imagery acquired in-house using an imaging system comprised by an Acousto-Optic Tunable Filter (AOTF) fitted to a firewire camera. The system has been designed to operate in the visible and near infrared (NIR) spectral range.

In our dataset, we have images corresponding to five categories of toys and a set of coins. Each toy sample was acquired over ten views by rotating the object in increments of 10° about its vertical axis whereas coin imagery was captured only in two different views, heads and tails. In our database, there are a total of 62 toys and 32 coins, which, over multiple viewpoints yielded 684 hyper-



Figure 1.6. From left-to-right: 4, 16 and 64-squared image region partitioning of the fluffy toy image.

Level	category	Same Scale		Multiple Scale	
		calibrated %	uncalibrated %	calibrated %	uncalibrated %
4-Region Lattice	animals	97.39	90.32	99.13	99.13
	cars	70.00	77.55	100.0	100.0
	fluffy dolls	83.33	41.49	90.00	96.67
	plastic blocks	80.00	96.24	97.14	97.14
	wooden blocks	96.00	98.74	99.00	99.00
	coins	93.75	87.64	96.88	96.88
	average total	91.23	89.47	97.72	98.54
16-Region Lattice	animals	94.78	98.26	100.0	100.0
	cars	90.00	80.00	100.0	100.0
	fluffy dolls	80.00	93.33	96.67	96.67
	plastic blocks	97.14	94.29	100.0	97.14
	wooden blocks	100.0	100.0	99.00	99.00
	coins	90.63	93.75	96.88	96.88
	average total	94.44	95.91	99.12	98.83
64-Region Lattice	animals	98.26	98.26	97.39	97.39
	cars	96.67	96.67	96.67	100.0
	fluffy dolls	80.00	76.67	90.00	100.0
	plastic blocks	82.86	82.86	97.14	94.29
	wooden blocks	100.0	100.0	100.0	100.0
	coins	90.63	90.63	100.0	96.88
	average total	94.74	94.44	97.66	98.25
Average		93.47	93.27	98.17	98.54

Table 1.1. Image categorisation results as percentage of correctly classified items in the dataset using the nearest neighbour classifier and our descriptor.

spectral images. Each image is comprised of 51 bands for those wavelengths ranging from 550 to 1000 nm over 9nm steps. For purposes of photometric calibration, we have also captured an image of a white Spectralon calibration target so as to recover the power spectrum of the illuminant across the scene. In Figure 1.4, we show the band corresponding to the 662nm-wavelength for five sample toys and a coin in our dataset. In the figure, each object corresponds to one of our six categories.

For our experiments, we have used our descriptors for purposes of recognition as follows. We commence by partitioning the imagery into two sets of equal size. The first of these is used for purposes of training, whereas the rest of the images are used as a testing data-base for purposes of recognition. We do this making use of both, a k -nearest neighbour classifier [8] and a Support Vector Machine (SVM) [9]. For the SVM, we use an RBF kernel whose parameters have been obtained via cross validation.

Note that, to our knowledge, there is no hyperspectral image descriptors available in the literature. Nonetheless, it is worth noting that the wavelength

Level	category	Single Scale		Multiple Scale	
		calibrated %	uncalibrated %	calibrated %	uncalibrated %
4-Region Lattice	animals	97.39	100.0	97.39	100.0
	cars	30.00	93.33	6.67	93.33
	fluffy dolls	88.57	97.14	80.00	97.14
	plastic blocks	56.67	100.0	53.33	100.0
	wooden blocks	52.00	98.00	40.00	98.00
	coins	65.63	96.88	31.25	96.88
	average total	65.04	97.56	51.44	97.56
16-Region Lattice	animals	94.78	99.13	96.52	96.52
	cars	16.67	56.67	3.33	80.00
	fluffy dolls	68.57	94.29	62.86	88.57
	plastic blocks	13.33	70.00	20.00	13.33
	wooden blocks	54.00	100.0	30.00	94.00
	coins	18.75	90.63	3.13	6.25
	average total	44.35	85.12	35.97	63.11
64-Region Lattice	animals	97.39	100.0	94.78	92.17
	cars	0.00	0.00	0.00	3.33
	fluffy dolls	45.71	54.29	51.43	65.71
	plastic blocks	0.00	13.33	0.00	0.00
	wooden blocks	33.00	98.00	28.00	93.00
	coins	0.00	0.00	0.00	0.00
	average total	29.35	44.27	29.04	42.37
Average		46.25	75.65	38.82	67.68

Table 1.2. Image categorisation results as percentage of correctly classified items in the dataset using a nearest neighbour classifier and the LBP-based descriptor in [43].

resolved nature of hyperspectral imagery are reminiscent of the time dependency in dynamic textures, where a pixel in the image can be viewed as a stationary time series. As a result, we compare our results with those yielded using the algorithm in [43]. The reasons for this are twofold. Firstly, this is a dynamic texture descriptor based upon local binary patterns (LBPs), which can be viewed as a local definition of texture and shape in the scene which combines the statistical and structural models of texture analysis. Secondly, from the results reported in [43], this method provides a margin of advantage over other alternatives in the dynamic texture literature. For the descriptors, in the case of the LBP method in [43], we have used a dimensionality of 1938 over the 51 bands in the images. For our descriptor, the dimensionality is 1500.

Since we have photometric calibration data available, in our experiments we have used two sets of imagery. The first of these corresponds to the dataset whose object images are given by the raw imagery. The second of these is given by the images which have been normalised with respect to the illuminant power spectrum. Thus, the first set of images corresponds to those hyperspec-

Level	category	Single Scale		Multiple Scale	
		calibrated %	uncalibrated %	calibrated %	uncalibrated %
4-Region Lattice	animals	97.39	97.39	97.39	99.13
	cars	90.00	100.0	86.67	0.00
	fluffy dolls	90.00	88.57	100.0	85.71
	plastic blocks	100.0	100.0	97.14	96.67
	wooden blocks	99.00	98.00	99.00	99.00
	coins	100.0	96.88	78.13	96.88
	average total	96.07	96.81	93.06	79.56
16-Region Lattice	animals	89.57	100.0	91.30	100.0
	cars	100.0	70.00	96.67	0.00
	fluffy dolls	63.33	62.86	100.0	22.86
	plastic blocks	91.43	100.0	91.43	76.67
	wooden blocks	100.0	100.0	99.00	94.00
	coins	100.0	100.0	71.88	81.25
	average total	90.72	88.81	91.67	62.46
64-Region Lattice	animals	90.43	100.0	94.78	33.04
	cars	80.00	0.00	93.33	26.67
	fluffy dolls	76.67	14.29	90.00	11.43
	plastic blocks	56.67	0.00	82.86	26.67
	wooden blocks	100.0	70.00	92.00	69.00
	coins	53.13	96.88	78.13	90.63
	average total	76.15	46.86	88.52	42.91
Average		87.65	77.49	92.37	47.48

Table 1.3. Image categorisation results as percentage of correctly classified items in the dataset using and SVM with an RBF kernel and our descriptor.

tral data where the classification task is effected upon scene radiance, whereas the latter corresponds to a set of reflectance images. From now on, we denote the radiance-based set as the “uncalibrated” one. We connote the reflectance imagery as “calibrated”. In Figure 1.5, we show sample hyperspectral image bands for a fluffy toy at wavelengths corresponding to 550nm, 640nm, 730nm, 820nm, 910nm, and 1000nm. In the figure, the top row shows the uncalibrated imagery whereas the bottom row shows the calibrated data.

For purposes of recognition, we have computed our descriptors and the alternative making use of an approach reminiscent of the level-1 spatial histogram representation in [22]. This is, we have subdivided the images in a lattice-like fashion into 4, 16 and 64 squared patches of uniform size. In Figure 1.6 we show the 4, 16 and 32-square lattice on the fluffy toy image. As a result, each image in either set, i.e. calibrated or uncalibrated, is comprised by 4, 16 or 64 descriptors. Here, we perform recognition based upon a majority voting scheme, where each of these descriptors is classified at testing time. Further, note that the fact that we have divided each image into 4, 16 and 64 squared

Level	category	Single Scale		Multiple Scale	
		calibrated %	uncalibrated %	calibrated %	uncalibrated %
4-Region Lattice	animals	93.91	98.26	80.87	100.0
	cars	80.00	96.67	20.00	53.33
	fluffy dolls	100.0	100.0	80.00	91.43
	plastic blocks	70.00	100.0	3.33	66.67
	wooden blocks	83.00	100.0	80.00	97.00
	coins	93.75	100.0	6.25	100.00
	average total	86.78	99.15	45.08	84.74
16-Region Lattice	animals	83.48	99.13	82.61	99.13
	cars	31.03	44.83	0.00	3.45
	fluffy dolls	65.71	80.00	42.86	51.43
	plastic blocks	6.67	70.00	0.00	20.00
	wooden blocks	70.00	99.00	70.00	98.00
	coins	28.13	84.38	3.13	93.75
	average total	47.50	79.56	33.10	60.96
64-Region Lattice	animals	79.35	83.48	77.39	88.70
	cars	0.00	0.00	0.00	0.00
	fluffy dolls	19.29	17.14	2.86	5.71
	plastic blocks	0.00	0.00	0.00	0.00
	wooden blocks	61.25	84.00	60.00	67.00
	coins	0.78	0.00	3.13	3.13
	average total	26.78	30.77	23.90	27.42
Average		53.69	69.83	34.03	57.72

Table 1.4. Image categorisation results as percentage of correctly classified items in the dataset using and SVM with an RBF kernel and our the LBP descriptor in [43].

regions provides a means to multiscale descriptor classification. Thus, in our experiments, we have used two majority voting schemes. The first of these limits the classification of descriptors to those at the same scale, i.e. number of squared regions in the image. The second scheme employs all the descriptors computed from multiple scales, i.e. $64 + 16 + 4$ for every image.

In Tables 1.1–1.4 we show the categorisation results for our dataset. In the tables, we show the results, per category and overall average, for the calibrated and uncalibrated data for both classifiers over the two schemes described above, i.e. multiscale and single-scale, when both, our method and the alternative are used to compute the image descriptors for the imagery. From the tables, its clear that our descriptor delivers better categorisation performance consistently for both classifiers. This is ever so important since our descriptor has a lower dimensionality than the alternative. We can attribute this behaviour to the high information compaction of the FCT.

Also, note that for the nearest neighbour classifier, the overall results yielded using our method show no clear trend with respect to the use of reflectance,

i.e. calibrated data, or radiance (uncalibrated imagery). This suggests that our method is robust to illuminant power spectrum variations. In the case of the SVM, the calibrated data with a multiscale approach delivers the best average categorisation results. For the alternative, the nearest neighbour classifier on uncalibrated data yields the best performance. Nonetheless, in average, absolute bests between the two descriptor choices here are 23% apart, being 75.63% for the LBP descriptor and 98.54% for our method. Further, note that for the coins, the alternative can be greatly affected by the effect of specularities at finer scales, i.e. the 64-region lattice. In contrast, our descriptor appears to be devoid of this sort of corruption.

8. Conclusion

In this chapter, we have showed how a local hyperspectral image descriptor can be computed via harmonic analysis. This descriptor is invariant to affine transformations on the corresponding local planar object surface patch. The descriptor is computed using an integral transform whose kernel is harmonic in nature. Affine invariance is then attained by relating the local planar object surface patch to a plane of reference whose orientation is fixed with respect to the camera plane. We have shown how high information compaction in the classifier can be achieved by making use of an FCT. It is worth stressing that the developments in the chapter are quite general and apply to a number of harmonic kernels spanning a Hilbert space. This opens-up the possibility of using other techniques available elsewhere in the literature, such as Mellin transforms, wavelets or Hankel transforms. We have showed the utility of the descriptor for purposes of image categorisation on a dataset of real-world hyperspectral images.

References

- [1] Aloimonos, J., and Swain, M. J. (1988). Shape from patterns: Regularization, *International Journal of Computer Vision*, 2:171–187.
- [2] Asmussen, S., Binswanger, K., and Hojgaard, B. (2000). Rare events simulation for heavy-tailed distributions. *Bernoulli*, 6(2):303–322.
- [3] Bar-Joseph, Z., El-Yaniv, R., Lischinski, D., and Werman, M. (2001). Texture Mixing and Texture Movie Synthesis Using Statistical Learning. *IEEE Transactions on Visualization and Computer Graphics*, 7(2):120-135.
- [4] Bracewell, R. N., Chang, K.-Y., Jha, A. K., and Wang, Y. -H. (1993). Affine theorem for two-dimensional fourier transform. *Electronics Letters*, 29(3):304.
- [5] Chen, Y., Wang, J.Z. and Krovetz, R. (2005) CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning. *IEEE Transactions on Image Processing*, 14(8):1187–1201.
- [6] Chetverikov, D. and Péteri, R. (2005). A brief survey of dynamic texture description and recognition. *Computer Recognition Systems*, pp.17–26.
- [7] Chum, O., Philbin, J., Sivic, J., Isard, M. and Zisserman, A. (2007). Total Recall: Automatic Query Expansion with a Generative Feature Model for Object Retrieval. In *Proceedings of the 11th International Conference on Computer Vision, ICCV 2007*, pp.1–8.
- [8] Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27.
- [9] Cristianini, N., Shawe-Taylor, J. and Elisseeff, A. (2002). On Kernel-Target Alignment. *Advances in Neural Information Processing Systems*, pp.367-373. MIT Press.
- [10] Doretto, G., Chiuso, A., Wu, Y. N., and Soatto, S. (2003). Dynamic textures. *International Journal of Computer Vision*, 51(2):91–109.

- [11] Dundar, M., and Landgrebe, D. (2004). Toward an Optimal Supervised Classifier for the Analysis of Hyperspectral Data. *IEEE Transaction on Geoscience and Remote Sensing*, 42(1):271–277.
- [12] Fazekas, S., and Chetverikov, D. (2005). Normal versus complete flow in dynamic texture recognition: a comparative study. In *Proceedings of the Fourth Int'l Workshop Texture Analysis and Synthesis*, pp.37-42.
- [13] Fukunaga, K. (1990) *Introduction to Statistical Pattern Recognition*. 2nd edition. Academic Press.
- [14] Gårding, J. (1993). Direct Estimation of Shape from Texture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1202–1208.
- [15] Ghanem, B. and Ahuja, N. (2007). Phase Based Modelling of Dynamic Textures. In *Proceedings of the 2007 IEEE International Conference on Computer Vision*, pp.1–8.
- [16] Ikeuchi, K. (1984). Shape from regular patterns. *Artificial Intelligence*, 22(1):49–75.
- [17] Jimenez, L. and Landgrebe, D. (1999). Hyperspectral Data Analysis and Feature Reduction via Projection Pursuit. *IEEE Transaction on Geoscience and Remote Sensing*, 37(6):2653–2667.
- [18] Jolliffe, I.T. (2002). *Principal Component Analysis*. 2nd edition. Springer-Verlag, New York.
- [19] Kanatani, K. and Chou, T. (1989). Shape from texture: general principle. *Artificial Intelligence*, 38(1):1–48.
- [20] Katznelson, Y. (2004). *An introduction to harmonic analysis*, 3rd edition. Cambridge University Press.
- [21] Landgrebe, D. (2002). Hyperspectral image data analysis. In *Signal Processing Magazine*, 19(1):17–28.
- [22] Lazebnik, S., Schmid, C. and Ponce, J. (2006). Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:2169–2178.
- [23] Li, Fei-Fei. and Perona, P. (2005). A Bayesian Hierarchical Model for Learning Natural Scene Categories. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:524–531.

- [24] Malik, J. and Rosenholtz, R. (1997). Computing Local Surface Orientation and Shape from Texture for Curved Surfaces. *International journal of computer vision*, 23(2):149–168.
- [25] Nilsback, M. and Zisserman, A. (2006). A Visual Vocabulary for Flower Classification. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:1447–1454.
- [26] Nistér, D. and Stewénius, H. (2006). Scalable Recognition with a Vocabulary Tree. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2:2161–2168.
- [27] Otsuka, K., Horikoshi, T., Suzuki, S. and Fujii, M. (1988). Feature Extraction of Temporal Texture Based on Spatiotemporal Motion Trajectory. In *Proceedings of the International Conference on Pattern Recognition*, 2:1047.
- [28] Portilla, J. and Simoncelli, E.P. (2000) A Parametric Texture Model based on Joint Statistics of Complex Wavelet Coefficients. *International journal of computer vision*, 40(1):49–71.
- [29] Quelhas, P., Monay, F., Odobez, J., Gatica-Perez, D., Tuytelaars, T., and Van Gool, L. (2005). Modelling scenes with local descriptors and latent aspects. In *Proceedings of the 10th IEEE International Conference on Computer Vision*, pp.883–890.
- [30] Rahman, A. and Murshed, M. (2005). A Robust Optical Flow Estimation Algorithm for Temporal Textures. In *Proceedings of the 2005 International Conference on Information Technology: Coding and Computing (ITCC'05)*, 2:72–76.
- [31] Randen, T. and Husoy, J. H. (1999). Filtering for texture classification: a comparative study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4):291–310.
- [32] Rao, K. R. and Yip, P. (1990). *Discrete Cosine Transform: Algorithms, Advantages, Applications*. Academic Press Professional, Inc., San Diego.
- [33] Ribeiro, E. and Hancock E. R. (2001). Shape from periodic texture using the eigenvectors of local affine distortion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(12):1459–1465.
- [34] Sengupta, K. and Boyer, K.L. (1995). Using Geometric Hashing with Information Theoretic Clustering for Fast Recognition from a Large CAD Modelbase. In *Proceedings of the IEEE International Symposium on Computer Vision*, pp.151-156.

- [35] Sheikh, Y. Haering, N. and Shah, M. (2006). Shape from Dynamic Texture for Planes. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2:2285–2292.
- [36] Shokoufandeh, A., Dickinson, S.J., Siddiqi, K. and Zucker, S.W. (1998). Indexing using a Spectral Encoding of Topological Structure. In *Proceedings of the Computer Vision and Pattern Recognition* 2:491–497.
- [37] Sivic, J. and Zisserman, A. (2003). Video Google: Text Retrieval Approach to Object Matching in Videos. In *Proceedings of the International Conference on Computer Vision*, 2:1470–1477.
- [38] Sneddon, I. N. (1995). *Fourier Transforms*. Dover, New York.
- [39] Stein, EM and Weiss, G. (1971). *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press.
- [40] Super, B.J. and Bovik, A.C. (1995). Shape from texture using local spectral moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(4):333–343.
- [41] Varma, M. and Ray, D. (2007). Learning The Discriminative Power-Invariance Trade-Off. In *Proceedings of the International Conference in Computer Vision (ICCV)*, pp.1–8.
- [42] Young, N. (1988). *An introduction to Hilbert space*. Cambridge University Press.
- [43] Zhao, G. and Pietikäinen, M. (2007). Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):915–928.