An affine Invariant Hyperspectral Texture Descriptor Based Upon Heavy-tailed Distributions and Fourier Analysis [*]

Pattaraporn Khuwuthyakorn[1,2,3] and Antonio Robles-Kelly[2,3] and Jun Zhou[2,3]

[1]Cooperative Research Centre for National Plant Biosecurity, Bruce ACT 2617, Australia
[2]RSISE, Bldg. 115, Australian National University, Canberra ACT 0200, Australia
[3]National ICT Australia (NICTA[†]), Locked Bag 8001, Canberra ACT 2601, Australia

{u4420081, antonio.robles-kelly, jun.zhou}@anu.edu.au

Abstract

In this paper, we address the problem of recovering a hyperspectral texture descriptor. We do this by viewing the wavelength-indexed bands corresponding to the texture in the image as those arising from a stochastic process whose statistics can be captured making use of the relationships between moment generating functions and Fourier kernels. In this manner, we can interpret the probability distribution of the hyper-spectral texture as a heavy-tailed one which can be rendered invariant to affine geometric transformations on the texture plane making use of the spectral power of its Fourier cosine transform. We do this by recovering the affine geometric distortion matrices corresponding to the probability density function for the texture under study. This treatment permits the development of a robust descriptor which has a high information compaction property and can capture the space and wavelength correlation for the spectra in the hyperspectral images. We illustrate the utility of our descriptor for purposes of recognition and provide results on real-world datasets. We also compare our results to those yielded by a number of alternatives.

1. Introduction

The development of image sensor technology has made it possible to capture image data in hundreds of wavelength-resolved images covering a broad spectral range in the visible and near-infrared range. The information-rich representation of the object under study provided by hyperspectral imagery poses significant opportunities and challenges for localisation and classification tasks.

Hyperspectral imaging can be intuitively related to material chemistry so as to employ spectral signatures for purposes of classification and localisation. This hinges on the notion that different materials have a characteristic response as a function of wavelength which can be used so as to provide a description of the object under study. So far, algorithms have been proposed for purposes of recognition and classification based upon spectral imaging [6, 4]. These are often based upon subspace projection methods such as Principal Component Analysis, Linear Discriminant Analysis, Decision Boundary , Projection Pursuit, and kernel methods[10]. All these algorithms treat the raw pixel spectra as input vectors in high dimensional spaces and look for linear or nonlinear mappings to the feature space, often with reduced dimensionality, by optimizing certain criterion, leading to statistically optimal solutions to classification.

Despite effective, the use of higher-level features, such as those based upon texture or shape for purposes of recognition, classification and localisation in hyperspectral imaging are less common in the literature. Moreover, to our knowledge, there are no hyperspectral texture descriptors elsewhere in the literature. In trichromatic imaging, texture has found applications not only as a shape queue [7, 18], but has also attracted broad attention for recognition and classification tasks [13]. Moreover, from the shape recovery perspective, static texture planes can be recovered making use of the structural analysis of predetermined texture primitives [9, 1]. This treatment provides an intuitive geometrical meaning to the task of recovering the parameters governing the pose by making use of methods akin to 3D view geometry.

In this paper, we focus in the development of a hyperspectal texture descriptor. To do this, we use an approach based upon the higher-order statistics of the texture under study. This is somewhat akin to the treatment given to dy-

namic textures in computer vision, where image sequences are modelled via probability density functions which exhibit first and second order time-shift invariant moments [5]. As their static counterparts, dynamic textures have attracted considerable interest for purposes of recognition. [17, 19].

We propose a descriptor based upon the use of statistics and Fourier kernels for purposes of capturing a discriminative and descriptive representation of the hyperspectral textures in the scene. This provides a principled link between statistical approaches, signal processing methods for texture recognition and shape modeling approaches based upon measures of spectral distortion [12, 15]. Hence, this is a means to both, a compact representation of the hyperspectral texture based upon its statistical characterisation and affine invariance via the estimation of the transformation parameters on a locally planar texture plane. Here, we view the hyperspectral texture as arising from a heavy-tailed distribution which can be related to a Fourier transform whose kernel, for purposes of efficiency has been set to a cosine function. This derives into a cosine transform which, when effected with respect to the pixel-coordinates can be related to the affine distortion of the texture. As a result, our descriptor is affine-invariant at the image-level and, at a wavelength level, captures the correlation of the moment generating functions for each of the bands under study.

The paper is organised as follows. In Section 2, we introduce heavy tailed distributions and relate them to Fourier kernels. We also elaborate on the compactness of representation making use of a cosine transform and provide a formulation for the local geometric distortion matrix employed here for purposes of affine invariance. In Section 3, we turn our attention to the computation of the descriptor. In Section 4, we provide further discussion on our descriptor and its affine invariance. In Section 5 we show experimental results making use of our descriptor for hyperspectral texture recognition on two real-world datasets. We also provide comparison to alternatives elsewhere in the literature and provide discussion on the application of the descriptor to other settings. Finally, we provide conclusion in Section 6.

2. Heavy-tailed Distributions

Here, we view hyperspectral textures as stochastic processes whose moment generating functions are invariant with respect to shifts in the image-coordinates. That is, 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 their densities have high dispersion, their probability density functions are governed by further-order moments. These introduces a number of statistical "skewness" variables that allow high variability in terms of wavelength-dependent behaviour. This implies that the spectra in the hyperspectral

texture results in spectral values that can be rather high in terms of their deviation from the texture-spectra mean and variance. This variation, which cannot be ignored in hyperspectral data, is characteristic to the stochastic process governing the texture under study.

Thus, we formulate our descriptor so as to model "rare" stationary wavelength-dependent events on the image plane. This is reminiscent of simulation approaches where importance sampling cannot be effected via an exponential changes of measure 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]. We formulate the density of the pixel-values for the wavelength λ at the pixel u in the image-band I_λ of the texture as random variables $\mathcal{Y}_u$ whose inherent basis $\mathcal{X}_u = \{\chi_u(1), \chi_u(2), \ldots, \chi_u(|\mathcal{X}_u|)\}$ is such that

$$P(\mathcal{Y}_u) = \sum_{k=1}^{|\mathcal{X}_u|} P(\chi_u(k)) \qquad (1)$$

where, $\chi_u(k)$ 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 for the texture under study as arising from a family of heavy-tailed distributions whose variance is not necessarily finite. It is worth noting in passing that, for finite variance, the formalism above implies that $P(\mathcal{Y}_u)$ is normally distributed. Nonetheless, this treatment generalises the stochastic process to a number of independent influences, each of which is captured by the corresponding variable $\chi_u(k)$.

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 recasting it as a function of the variable ς making use of the characteristic function

$$
\begin{aligned}
\psi(\varsigma) \;&=\; \int_{-\infty}^{\infty} \exp(\mathbf{i}\varsigma\mathcal{Y}_u) f(\mathcal{Y}_u) d\mathcal{Y}_u \qquad (2) \\
&=\; \exp\Big(\mathbf{i}u\varsigma - \gamma\,|\varsigma|^\alpha \left(1 + \mathbf{i}\beta\,\mathrm{sign}(\varsigma)\varphi(\varsigma,\alpha)\right)\Big)
\end{aligned}
$$

where $\mathbf{i} = \sqrt{-1}$, 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 $\varphi(.)$ is defined as follows

$$
\phi(\varsigma, \alpha) = \begin{cases} \tan(\alpha\frac{\pi}{2}) & \text{if } \alpha \neq 1 \\ -\frac{\pi}{2}\log|\varsigma| & \text{if } \alpha = 1 \end{cases} \qquad (3)
$$

Note that, for the characteristic function above, $\alpha = 2$ 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 modelling of a large number of distributions that may apply to hyperspectral data.

Note that, so far, we have limited ourselves to the image-plane for a fixed wavelength λ. That is, we have concentrated on the distribution of spectral values accross every wavelength-resolved band in the image. We can extend Equation 2 to the wavelength domain, i.e. the spectra of the texture across a segment of bands, by noting that the equation above is essentially a cross-correlation. Hence, we can write the characteristic function for the texture parameterised with respect to the wavelength λ as follows

$$
\begin{aligned}
\vartheta(t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(\mathbf{i}\lambda\varsigma)\exp(\mathbf{i}\varsigma\mathcal{Y}_u)f(\mathcal{Y}_u)d\mathcal{Y}_u d\varsigma \\
&= \int_{-\infty}^{\infty} \exp(\mathbf{i}\lambda\varsigma)\psi(\varsigma)d\varsigma \qquad (4)
\end{aligned}
$$

which captures the spectral cross-correlation for the characteristic functions for each band.

In this manner, we view the characteristic function for the hyperspectral texture 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 2 corresponds to the image-band domain ς of the texture. Equation 4 is then the wavelength-dependent domain λ. This composition operation also opens-up the possibility of following a two-step process. Firstly, at band-level, the information is represented in a compact fashion and rendered invariant to geometric distortions on the texture plane. In the second step, wavelength-dependent correlation between bands is captured making use of the operation in 4.

2.1. Compactness of Representation

To achieve a compact representation making use of the equations above, we refer to the fundamentals of integral transforms [16]. We can view Equations 2 and 4 as characteristic functions obtained via the integral of the product of the function $g(\eta)$, i.e. $f(\mathcal{Y}_u)$ and $\psi(\varsigma)$, multiplied by a kernel $\mathcal{K}(\omega, \eta)$, which above becomes $\exp(\mathbf{i}\lambda\varsigma)$ and $\exp(\mathbf{i}\varsigma\mathcal{Y}_u)$, respectively. As a result, we have

$$
F(\omega) = \int_{-\infty}^{\infty} g(\eta)\mathcal{K}(\omega, \eta)d\eta \qquad (5)
$$

where $\mathcal{K}(\omega, \eta)$ is a Fourier kernel.

To see the relation between Fourier transforms and the equations in previous sections, we can examine $\psi(\varphi)$ in more detail and write

$$
\begin{aligned}
\log[\psi(\varsigma)] &= \mathbf{i}u\varsigma - \gamma \mid \varsigma \mid^{\alpha} \left(1 + \mathbf{i}\beta \operatorname{sign}(\varsigma)\varphi(\varsigma, \alpha)\right) \\
&= \mathbf{i}u\varsigma - \mid \varsigma \mid^{\alpha} \gamma^{*\alpha} \exp\left(-\mathbf{i}\beta^*\frac{\pi}{2}\vartheta\operatorname{sign}(\varsigma)\right)
\end{aligned}
$$

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

$$
\begin{aligned}
\beta^* &= \frac{2}{\pi\vartheta} \arccos\left(\frac{\cos\left(\alpha\frac{\pi}{2}\right)}{\sqrt{\Omega}}\right) \\
\gamma^* &= \left(\frac{\gamma\sqrt{\Omega}}{\cos\left(\alpha\frac{\pi}{2}\right)}\right)^{\frac{1}{\alpha}} \qquad (6)
\end{aligned}
$$

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

To obtain the kernel for Equation 2, we can use Fourier inversion on the characteristic function and, making use of the shorthands defined above, the PDF may be computed via the equation

$$
\begin{aligned}
f(\mathcal{Y}_u; v, \beta^*, \gamma^*, \alpha) = \frac{1}{\pi\gamma^*} \int_0^{\infty} &\cos\left(\frac{(u - \mathcal{Y}_u)s}{\gamma^*} + s^{\alpha}\sin(\phi)\right) \\
&\exp\left(-s^{\alpha}\sin(\phi)\right)ds
\end{aligned}
$$
(7)

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

This treatment opens-up not only the possibility of effecting functional analysis on the characteristic function making use of the techniques in the Fourier domain, but also allows the use of other Fourier kernels for purposes of compactness and ease of computation. 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

$$
s^{\alpha}sin(\phi) = \omega\eta - A\eta^{\frac{1}{\alpha}} \qquad (8)
$$

Substituting Equation 8 into equation 7, the PDF can be expressed as follows

$$
\begin{aligned}
f(\mathcal{Y}_u; v, \beta^*, \gamma^*, \alpha) &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} \frac{\exp\left(-\omega\eta + A\eta^{\frac{1}{\alpha}}\right)}{\sqrt{2\pi}\gamma^*\alpha\eta^{\left(\frac{\alpha-1}{\alpha}\right)}} \\
&\quad cos(\omega\eta)\, d\eta
\end{aligned}
$$
(9)

where the Fourier kernel becomes

$$
\mathcal{K}(\omega, \eta) = \cos(\omega\eta) \qquad (10)
$$

which can be related, in a straightforward manner, to the Fourier Cosine Transform (FCT) of the form

$$
\begin{aligned}
F(\omega) &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} \frac{\exp\left(-\omega\eta + \frac{(u-\mathcal{Y}_u)}{\gamma^*}\eta^{\frac{1}{\alpha}}\right)}{\sqrt{2\pi}\gamma^*\alpha\eta^{\left(\frac{\alpha-1}{\alpha}\right)}} \\
&\quad cos(\omega\eta)\, d\eta
\end{aligned}
$$
(11)

which is analogous to the expression in Equation 7. 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 texture information in the lower-order Fourier terms, i.e. those for which w is close to the origin. This follows the strong "information compaction" property of FCTs introduced in [14] and assures a good trade-off between discriminability and complexity.

2.2. Invariance to Affine Distortions

Having introduced the notion of the FCT for purposes of representing the PDF of the dynamic texture under study, we now focus on relation between distortions on the texture plane and the Fourier domain. To this end, we follow [3] and relate the Fourier domain distortions to affine transformations on the texture shape. As mentioned earlier, the function $f(\mathcal{Y}_u)$ corresponds to the band-dependent component of the texture and, as a result, its prone to affine distortion. This hinges in the notion that a distortion on the texture plane will affect the geometric factor in the surface reflectance, 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 FCT of the function $f(\mathcal{Y}_u)$ which, as introduced in the previous sections, corresponds to the pixel values for the image-band I_λ in the texture under study is given by

$$F(\xi) = \int_\Gamma f(\mathcal{Y}_u) \cos\left(2\pi\left(\xi^T u\right)\right) du \qquad (12)$$

where $u = [x, y]^T$ is the vector of two-dimensional coordinates for the compact domain $\Gamma \in \Re^2$. It is worth noting in passing that, in practice, the coordinate-vectors $u = [x, y]^T$ 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

$$u' = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c \\ h \end{bmatrix} \qquad (13)$$

This observation is important because we can relate the kernel for the FCT in Equation (12) to the transformed coordinate $u' = [x', y']^T$. Also, note that, for patches centered at keypoints in the image, the texture can be considered devoid of translation. Thus, we can set $f = c = 0$ and write

$$
\begin{aligned}
\xi^T u &= \xi^T \begin{bmatrix} x \\ y \end{bmatrix} \\
&= \begin{bmatrix} \xi_x & \xi_y \end{bmatrix} \begin{bmatrix} a & b \\ d & e \end{bmatrix}^{-1} \begin{bmatrix} x' \\ y' \end{bmatrix} \\
&= \frac{1}{ae - bd} \begin{bmatrix} e\xi_x - d\xi_y & -b\xi_x + a\xi_y \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}
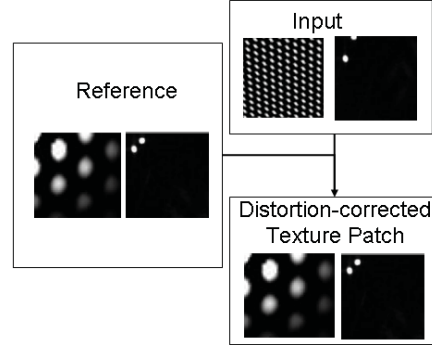\end{aligned}
\qquad (14)
$$

Figure 1. Examples of reference, input and distortion corrected single-band textures. In the panels, the left-hand image shows the single-band textures whereas the right-hand panel shows the power spectrum of the corresponding FCT.

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

Hence, after some algebra, and using the shorthand $\Delta = ae - bd$, we can show that the FCT for the transformed coordinates u' is

$$
\begin{aligned}
F(\xi) = {} & \frac{1}{|\Delta|} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathcal{Y}_{u'}) \cos\left(\frac{2\pi}{\Delta}(e\xi_x - d\xi_y)x'\right. \\
& \left. + (b\xi_x - a\xi_y)y'\right) dx' dy'
\end{aligned}
\qquad (15)
$$

This implies that

$$F(\xi) = \frac{1}{|\Delta|} F(\xi'). \qquad (16)$$

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

$$\xi = \begin{bmatrix} \xi_x \\ \xi_y \end{bmatrix} = \begin{bmatrix} a & d \\ b & e \end{bmatrix} \begin{bmatrix} \xi'_x \\ \xi'_y \end{bmatrix} = \mathbb{T}\xi' \qquad (17)$$

As a result, let the FCT for the affinely transformed pixel $\mathcal{Y}_{u'}$ be given by $\mathbf{U}'$. Analogously, suppose the FCT for the pixel $\mathcal{Y}_u$ be given by $\mathbf{U}$. From Equation 15, 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 Fourier domain for the corresponding FCT. This observation is an important one since it permits achieving invariance to affine transformations on the texture plane via a Fourier-domain distortion correction operation of the form

$$F(\xi) = (\mathbb{T}^T)^{-1} F(\xi') \qquad (18)$$

3. Descriptor Computation

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-FCT for every band in the hyperspectral texture under study. This is equivalent to computing the band-dependent component of the characteristic function $\psi(\varsigma)$. Secondly, we capture the wavelength-dependent behaviour of the dynamic texture by computing the FCT with respect to the spectral domain for the "train" of distortion-invariant FCTs. Thus, 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 FCT for each band in the image. To do this, we make use of the Fourier-domain property in Equation 18 to estimate the distortion matrix with respect to a predefined reference. Here, we make use of the peaks of the power spectrum and express the relation of the FCTs for two texture planes. We have done this following the notion that a blob-like texture composed of a single transcendental function on the texture plane would produce two peaks in the Fourier domain. That is, we have set, as our reference, a moment generating function arising from a texture modelled as a cosine on a texture plane parallel to the camera plane.

Let the peaks of the power spectrum for two locally planar texture 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 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

$$\mathbf{U}_A = (\mathbb{T}_A{}^T)^{-1}\,\mathbf{U}_R \qquad (19)$$

$$\mathbf{U}_B = (\mathbb{T}_B{}^T)^{-1}\,\mathbf{U}_R \qquad (20)$$

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 [15]. Nonetheless, in [15], the normal is recovered explicitly making use of the Fourier transform, whereas here, we employ the cosine 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

$$\mathbf{U}_B = \left(\mathbb{T}_A\mathbb{T}_B^{-1}\right)^T \mathbf{U}_A \qquad (21)$$
$$= \Phi\mathbf{U}_A \qquad (22)$$

where $\Phi = \left(\mathbb{T}_A\mathbb{T}_B^{-1}\right)^T$ 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

$$\Phi = \left(\mathbf{U}_R\mathbf{U}_A{}^{-1}\right)\left(\mathbf{U}_B\mathbf{U}_R^{-1}\right) \qquad (23)$$

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 descriptor of the form

$$\mathbf{V} = [F(I_1)^* \mid F(I_2)^* \mid \ldots \mid F(I_{|\mathbb{I}|})^*] \qquad (24)$$

which is the concatenation of the affine invariant FCTs $F(\cdot)^*$ for the frames in the texture. Moreover, we render the band-level FCT invariant to affine transformations making use of the reference peak matrix $\mathbf{U}_R$ such that the FCT for the frame indexed t is given by

$$F(I_R) = F(I_t)^*\Phi_t^{-1} \qquad (25)$$

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

With the band-level representation $\mathbf{V}$ at hand, we can perform the second FCT computation. This is done by using the discrete analogue of Equation 4. Thus, the kth coefficient for the texture descriptor $\mathcal{G}$ becomes

$$\mathcal{G}_k = F(\mathbf{V}) = \sum_{n=0}^{|\mathbb{I}|-1} F(I_n)^* \cos\left(\frac{\pi}{|\,\mathbb{I}\,|}\left(n + \frac{1}{2}\right)\left(k + \frac{1}{2}\right)\right) \qquad (26)$$

where $|\,\mathcal{G}\,| = |\,\mathbb{I}\,|$.

4. Discussion

In this section, we provide a discussion on the descriptor presented previously. We commence by illustrating the distortion correction operation at the band level in Figure 1. In the panels, we show the reference, corrected and input textures in their spatial and frequency domains. Note that, at input, the texture shows an affine distortion which affects the distribution of the peaks in its power spectrum. The distortion corrected texture patch is in good accordance with the reference.

To further illustrate the effects of affine distortions, in Figure 2, we show a sample texture 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
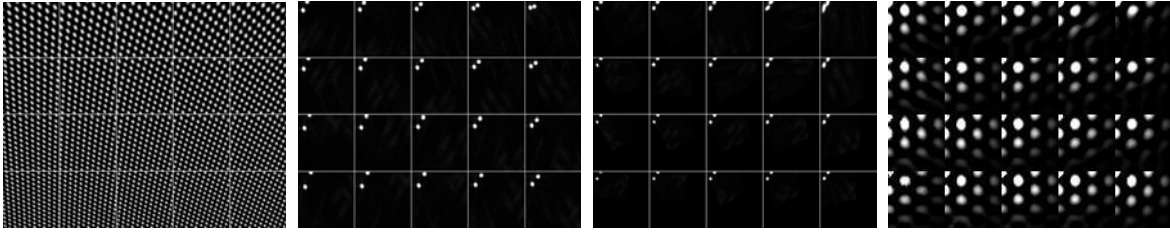
Figure 2. From left-to-right: Affine distortion of a sample single-band texture, FCT of the texture in the left-hand panel, distortion-corrected power spectrums for the FCTs in the second panel, inverse FCTs for the power spectrum in the third panel.

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 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.

Now, we turn our attention to the wavelength dependency in hyperspectral textures. To this end, we illustrate, in Figure 3, the step-sequence of the descriptor computation procedure. We depart from a series of bands in the image and compute the band-by-band 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 texture sizes on the image lattice and N_λ is equivalent to the wavelength range for the dynamic texture sequence. 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 time-indexes of the dynamic textures under study.

5. Experiments

For purposes of illustrating the utility of our descriptor for purposes of texture recognition, we make use of three datasets. The first two of these are hyperspectral imaging ones. The third is the DynTex dataset. The first of the hyperspectral datasets comprises 100 hyperspectral images of Oriental Fruit Moths (OFM). In each view, together with the insects, there is an amount of debris present, which acts as a confounding factor. The second hyperspectral data set is comprised by 50 images of urban scenes in high-oblique views. For the OFM dataset, the images have been acquired

in 10nm-step bands spanning from 380 to 1050nm. In the case of our urban views, these are comprised of 10 bands in regular wavelength intervals of 25nm from 420 to 650nm. Also, note that the third of our datasets is not hyperspectral in nature, but rather a dynamic texture one. This is so as to illustrate how the descriptor presented here may be applied to time-dependent as well as wavelength-resolved textures.

Here, we have used each of these descriptors for purposes of recognition as follows. After selecting a subset of the textures in each dataset for testing, the rest of the image are used as a data-base for purposes of recognition and localisation making use of a k-nearest neighbour classifier. We have organised this section as follows. We commence by presenting our results on hyperspectral data. We then turn our attention to the application of the descriptor presented here to a dynamic texture setting. Along these lines, note that, as mentioned earlier, to our knowledge, there are no hyperspectral texture descriptors available in the literature. As a result, for purposes of comparison, we have turned our attention to the dynamic texture literature and selected the the algorithm of Zhao and Pietikäinen [20] as that to compare against. The reasons for this are twofold. Firstly, this is a method based upon local binary patterns (LBPs), which can be viewed as a vehicle which combines the statistical and structural models of texture analysis. Secondly, from the literature [20], this method provides a margin of advantage over other alternatives.

5.1. Hyperspectral Imagery

In this section, we illustrate the utility of our descriptor for purposes of hyperspectral object localisation via texture recognition. To this end, we have used SIFT [11] keypoints recovered from the average image luminance so as to recover patches on the image plane. We have then used this patches as candidates for purposes of hyperspectral texture recognition. To do this, we have used half of the images in each dataset for purposes of training. The other half was used for testing.

In the case of the OFM dataset, we aim at localising the moths so as to identify them in each view. This is a challenging task since each image portrays areas with specular-
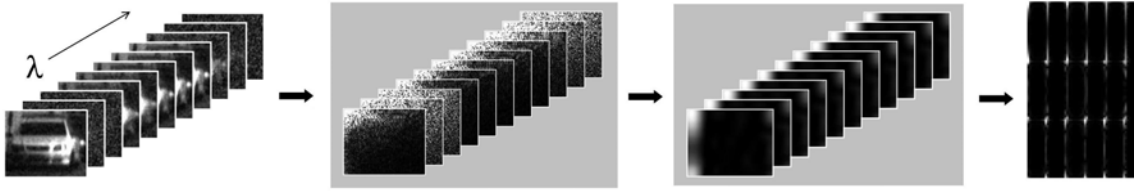
Figure 3. From left-to-right: Hyperspectral texture, the band-wise FCT, the distortion invariant cosine transforms for every band in the image and the final transform, i.e. correlation, with respect to wavelength.

Figure 4. Sample hyper-spectral OFM images used in our experiments.

Figure 5. Sample hyper-spectral urban views used in our car localisation experiments.

Data set	Method	Descriptor Length	Accuracy rate
Urban Scenes	FCT	1000	83.65%
	LBP [20]	4176	69.20%
OFM	FCT	3000	83.09%
	LBP [20]	4176	76.54%

Table 1. Recognition results (%) for the hyper-spectral image data sets with respect to our method (FCT) and the alternative Zhao & Pietikäinen [20]

ities, shadowing, debris and optic distortion. Each image contains eight moths. Figure 4 shows some sample OFM images. So as to deal with photometric artifacts, we have used the illuminant estimation method in [8] for purposes of photometric calibration. For purposes of training, we have used 400 texture patches depicting the moths in the dataset. Thus, the aim is to localise the 400 moths in the 50 remaining testing images.

For the urban-scene dataset, we have labelled the vehicles in each scene. We have done this so as to illustrate the utility of the descriptor for purposes of vehicle localisation in high-oblique hyperspectral imagery of built environments. Sample hyperspectral urban scene images are shown in Figure 5. As in the case of the OFM data, the images present a number of photometric artifacts whose effects in the localisation task have been mitigated through the use of photometric calibration. As before, the illuminant spectrum has been recovered using the method in [8]. With the calibrated images at hand, we have hand-labelled the 387 cars in the imagery. Of these, 189 appeared in the 25 training images and the other 198 were in the testing images. Again, in our experiments, we aim at localising these 198 vehicles within the testing imagery.

In Table 1, we show recognition results for both, our descriptor and the alternative, when applied to the two data sets under study. The accuracy reported in the table corresponds to the percentage of moths or vehicles correctly localised in the testing imagery. Note that, for the two data sets, our method provides a margin of improvement over the alternative. In the table, we also provide details regarding descriptor-lengths. Recall that our descriptor has a length of $N_x \times N_y \times N_\lambda$. Note that a longer descriptor will potentially encode more information. Nonetheless, due to the good information compaction properties of our approach, the complexity of the descriptor presented here is lower than that of the alternative.

5.2. Dynamic Textures

As mentioned earlier, we also present results on the DynTex dataset [1]. This is a dynamic texture depicting light candles, moving plants and trees, smoke plumes, sea waves, etc. Each texture class contains 12 dynamic texture sequences. Following the experimental setting in [20], we have pre-segmented each of the dynamic textures in each class so as to divide them into a number of non-overlapping subsets. Here, we have obtained ten non-overlapping image sequences of random sizes $N_x \times N_y \times N_t$ (here, N_t is equivalent to N_λ in the equations throughout the paper) from each texture sequence and set $k = 6$ for our k-nearest neighbour classifier. We have used the Euclidean distance between the descriptors corresponding to each of the dynamic texture pairs comprised by the query, i.e. testing, and the sequences in the data-base. An example of the segmentation operation on the dataset is shown in Figure 6. This pre-segmentation operation yields 1080 items in the final dataset, half of which we use for testing and the other half we reserve for recognition purposes.

From a more quantitative viewpoint, in Table 2, we show the recognition rates yielded by our descriptor as compared to the method in [20]. In the table, we have shown results for three different descriptor lengths and compare our re-

---

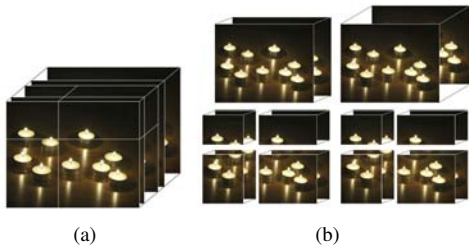[1] The DynTex dataset can be downloaded from http://www.cwi.nl/projects/dyntex/

Figure 6. Example segmentation of a sample DynTex dynamic texture sequence. (a): Original; (b): Segmented sequences

Method	Descriptor Length	Recognition rate
FCT	9000	96.24%
FCT	1000	95.71%
LBP [20]	4176	95.07%

Table 2. Table showing the average recognition results (%) on the DynTex database for our descriptor (FCT) and the alternative (LBP).

sults with those in [20]. It is somewhat expected that, as the length increases, more information can be compacted into the descriptor. This is in accordance with the fact that, for a length of 1000, i.e. $N_x = N_y = N_t = 10$, the recognition rate is slightly lower (approx 1%) than that yielded by the descriptor with length 9000, i.e. $N_x = N_y = 30$ and $N_t = 10$. Nonetheless, our descriptor with length 1000 still provides a margin of advantage over that in [20]. This is a significant observation since the descriptor in [20] has 4176 coefficients. That is, our descriptor outperforms the alternative regardless of being a factor of four less expensive to store.

6. Conclusion

We have presented a novel approach to dynamic texture recognition based upon heavy-tailed distributions. Here, we provide a link to Fourier kernels and, through the use of a cosine transform, we achieve a texture representation that is compact and invariant to spectral distortions. The resulting descriptor is given by the time-correlation for the distortion invariant frame-wise cosine transforms. We have illustrated the utility of the method for purposes of recognition by performing experiments on the DynTex and the MIT temporal texture dataset, where our method outperforms the alternatives.

References

[1] J. Aloimonos and M. Swain. Shape from texture. *Biological Cybernetics*, 58(5):345–360, 1988. 1

[2] S. Asmussen, K. Binswanger, and B. Hojgaard. Rare events simulation for heavy-tailed distributions. *Bernoulli*, 6(2):303–322, 2000. 2

[3] R. Bracewell, K.-Y. Chang, A. Jha, and Y.-H. Wang. Affine theorem for two-dimensional fourier transform. *Electronics Letters*, 29(3):304, Feb. 1993. 4

[4] R. N. Clark, G. A. Swayze, A. Gallagher, N. Gorelick, and F. Kruse. Mapping with imaging spectrometer data using the complete band shape least-squares algorithm simultaneously fit to multiple spectral features from multiple materials. In *Proceedings of the Third Airborne Visible/Infrared Imaging Spectrometer Workshop*, pages 91–28, 1991. 1

[5] G. Doretto, A. Chiuso, Y. Wu, and S. Soatto. Dynamic textures. *Int. Journal on Computer Vision*, 51(2):91–109, 2003. 2

[6] Z. Fu, A. Robles-Kelly, R. T. Tan, and T. Caelli. Invariant object material identification via discriminant learning on absorption features. In *Object Tracking and Classification in and Beyond the Visible Spectrum*, 2006. 1

[7] J. Garding. Direct estimation of shape from texture. In *ISRN KTH*, 1993. 1

[8] C. P. Huynh and A. Robles-Kelly. Optimal solution of the dichromatic model for multispectral photometric invariance. In *S+SSPR 2008*, pages 382–391, 2002. 7

[9] K. Ikeuchi. Shape from regular patterns. *Artificial Intelligence*, 22:49–75, 1984. 1

[10] D. Landgrebe. Hyperspectral image data analysis. *IEEE Signal Process. Mag.*, 19:17–28, 2002. 1

[11] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004. 6

[12] J. Malik and R. Rosenholtz. Computing local surface orientation and shape from texture for curved surfaces. *IJCV*, 23(2):149–168, June 1997. 2

[13] T. Randen and J. H. Husoy. Filtering for texture classification: a comparative study. *Transactions on Pattern Analysis and Machine Intelligence*, 21(4):291–310, 1999. 1

[14] K. R. Rao and P. Yip. *Discrete cosine transform: algorithms, advantages, applications*. Academic Press Professional, Inc., San Diego, CA, USA, 1990. 4

[15] E. Ribeiro and E. Hancock. Shape from periodic texture using the eigenvectors of local affine distortion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(12):1459–1465, 2001. 2, 5

[16] I. N. Sneddon. *Fourier Transforms*. McGraw-Hill, 1951. 3

[17] S. Soatto, G. Doretto, and Y. N. Wu. Dynamic textures. In *ICCV*, pages 439–446, 2001. 2

[18] B. Super and A. Bovik. Shape from texture using local spectral moments. *Transactions on Pattern Analysis and Machine Intelligence*, 17(4):333–343, April 1995. 1

[19] M. Szummer and R. Picard. Temporal texture modeling. *Image Processing, 1996. Proceedings., International Conference on*, 3:823–826 vol.3, Sep 1996. 2

[20] G. Zhao and Pietikäinen. 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, 2007. 6, 7, 8

119