This paper reconciles Kingsbury’s dual-tree complex wavelets with Nason and Eckley’s local stationary model. We here establish that the dual-tree wavelets admit an invertible debiasing matrix and that this matrix can be used to invert the covariance relation. We also show that the added directional selectivity of the proposed model adds utility to the standard two-dimensional local stationary model. Non-stationarity detection on random fields is used as a motivating example. Experiments confirm that the dual-tree model can distinguish anisotropic non-stationarities significantly better than the current model.

Index Terms— wavelet, locally stationary, anisotropic

1. INTRODUCTION

First proposed by Nason et al. [1] in the one-dimensional setting, the locally stationary wavelet (LSW) model offers a translation-invariant, statistically well-principled means to capture the local covariance and local spectrum of a large class of, possibly non-stationary, stochastic processes. Extension of the theory to two-dimensional processes was recently pioneered by Eckley et al. [2, 3] and was followed up with methodology applied to non-stationarity detection in image textures [4] and segmentation of imagery into stationary regions [5].

Alternative wavelet approaches to modelling random fields often involve modelling the raw wavelet coefficients directly [6] or passing them as features to machine learning tools [7]. This is in contrast to the LSW where the process is explicitly described with a parameterised, generative model where the parameters are estimated in a statistically rigorous manner to ensure that: the asymptotic bias is understood and corrected; and that consistency is characterised and, where appropriate, obtained. Furthermore, the LSW model offers a natural way to extend the notion of wavelet auto-covariance to the local case—Prop. 1 and Cor. 1 in [2] show respectively that the local auto-covariance converges uniformly to auto-covariance as sample size increases and that one can be derived from the other via explicit linear mappings. Although other wavelet models, such as HMTs [8,9], capture dependency structures between wavelet coefficients, no other wavelet-based model other than the LSW appears to satisfy these fundamental covariance properties. Perhaps the closest model to LSW is [10,11] which was applied to paper and mamographic imagery; however their wavelet variance-based framework is restricted to the smaller class of (intrinsically) stationary fields.

The LSW model is translation invariant thanks to the use of what Nason terms ‘discrete wavelets’, formed by an inverse, undecimated, à trous algorithm using finite impulse response (FIR) wavelet filters. However, the current design currently only accommodates wavelet bases with the standard ‘horizontal’, ‘vertical’, and ‘diagonal’ orientations. Such poor directionality is a serious drawback when the data of interest presents anisotropic behaviour. Some non-stationarity detection experiments in Section 4 illustrate examples of this point.

Motivated by this lack of directionality we here move towards reconciling two significant contributions to the field of wavelet analysis, namely: the LSW model [1,2] and Kingsbury’s dual-tree complex wavelet transform (DTCWT) [8,12–14]. The former evolved from the motivation to build a statistically rigorous notion of localised auto-covariance using wavelets. The latter was motivated by the desire to design a wavelet basis with good shift-invariance which could be implemented in a decimated pyramidal scheme— a consequence of which is that, in two-dimensions, the DTCWT furnishes wavelets with superior directionality than standard real, FIR wavelets. Since the DTCWT is implemented with an FIR filterbank our proposed DTCWT-LSW model also uses ‘discrete wavelets’ in a similar fashion to the LSW case.

By extending some key notions such as the auto-correlation wavelet and the biasing matrix to the complex-valued DTCWT-LSW case we then: (i) establish some mild conditions under which the biasing matrix can be inverted (Theorem 6); (ii) prove that the mapping between local auto-covariance and standard auto-covariance is invertible (Prop. 8); (iii) demonstrate that the added directional selectivity of the proposed model adds utility to the standard two-dimensional local stationary model in that it is more general and, unlike the standard LSW, can be used to detect anisotropic non-stationarities.

Section 2 offers a brief overview of the LSW model and the DTCWT. We propose the DTCWT-LSW model in Section 3, establish some fundamental results, and use a toy example to illustrate that its enhanced directionality represents a more general model space than the standard LSW. Experiments are conducted in Section 4 which demonstrate the superiority of the DTCWT-LSW at non-stationarity detection.
2. BACKGROUND

The (2-d) LSW and DTCWT are here reviewed.

2.1. Locally stationary wavelet model

The local stationary wavelet (LSW) model represents a stochastic process \( X = \sum_{j,\ell} \sum_n w_j^\ell(u) \xi_j^\ell(t) \psi_j^\ell(u \cdot) \) on the finite lattice \([1, T]^2\) as a weighted filtering (with weights \( w \)) of a zero-mean orthonormal increment sequence \( \xi \) and an overcomplete basis comprising ‘undecimated discrete wavelet functions’ \( \psi_j^\ell(\cdot) \in L_2(\mathbb{Z}^2) \) [1, 2]. The wavelet functions are defined/derived by performing the usual inverse decimated wavelet transform (with upsampling) of Delta functions placed following result suggests that this would be far from ideal.

The local stationary wavelet (LSW) model represents a stochastic process \( X = \sum_{j,\ell} \sum_n w_j^\ell(u) \xi_j^\ell(t) \psi_j^\ell(u \cdot) \) on the finite lattice \([1, T]^2\) as a weighted filtering (with weights \( w \)) of a zero-mean orthonormal increment sequence \( \xi \) and an overcomplete basis comprising ‘undecimated discrete wavelet functions’ \( \psi_j^\ell(\cdot) \in L_2(\mathbb{Z}^2) \) [1, 2]. The wavelet functions are defined/derived by performing the usual inverse decimated wavelet transform (with upsampling) of Delta functions placed following result suggests that this would be far from ideal.

The local stationary wavelet (LSW) model represents a stochastic process \( X = \sum_{j,\ell} \sum_n w_j^\ell(u) \xi_j^\ell(t) \psi_j^\ell(u \cdot) \) on the finite lattice \([1, T]^2\) as a weighted filtering (with weights \( w \)) of a zero-mean orthonormal increment sequence \( \xi \) and an overcomplete basis comprising ‘undecimated discrete wavelet functions’ \( \psi_j^\ell(\cdot) \in L_2(\mathbb{Z}^2) \) [1, 2]. The wavelet functions are defined/derived by performing the usual inverse decimated wavelet transform (with upsampling) of Delta functions placed following result suggests that this would be far from ideal.

The local stationary wavelet (LSW) model represents a stochastic process \( X = \sum_{j,\ell} \sum_n w_j^\ell(u) \xi_j^\ell(t) \psi_j^\ell(u \cdot) \) on the finite lattice \([1, T]^2\) as a weighted filtering (with weights \( w \)) of a zero-mean orthonormal increment sequence \( \xi \) and an overcomplete basis comprising ‘undecimated discrete wavelet functions’ \( \psi_j^\ell(\cdot) \in L_2(\mathbb{Z}^2) \) [1, 2]. The wavelet functions are defined/derived by performing the usual inverse decimated wavelet transform (with upsampling) of Delta functions placed following result suggests that this would be far from ideal.

The local stationary wavelet (LSW) model represents a stochastic process \( X = \sum_{j,\ell} \sum_n w_j^\ell(u) \xi_j^\ell(t) \psi_j^\ell(u \cdot) \) on the finite lattice \([1, T]^2\) as a weighted filtering (with weights \( w \)) of a zero-mean orthonormal increment sequence \( \xi \) and an overcomplete basis comprising ‘undecimated discrete wavelet functions’ \( \psi_j^\ell(\cdot) \in L_2(\mathbb{Z}^2) \) [1, 2]. The wavelet functions are defined/derived by performing the usual inverse decimated wavelet transform (with upsampling) of Delta functions placed following result suggests that this would be far from ideal.

The local stationary wavelet (LSW) model represents a stochastic process \( X = \sum_{j,\ell} \sum_n w_j^\ell(u) \xi_j^\ell(t) \psi_j^\ell(u \cdot) \) on the finite lattice \([1, T]^2\) as a weighted filtering (with weights \( w \)) of a zero-mean orthonormal increment sequence \( \xi \) and an overcomplete basis comprising ‘undecimated discrete wavelet functions’ \( \psi_j^\ell(\cdot) \in L_2(\mathbb{Z}^2) \) [1, 2]. The wavelet functions are defined/derived by performing the usual inverse decimated wavelet transform (with upsampling) of Delta functions placed following result suggests that this would be far from ideal.
(approximately) shift invariant but, in two-dimensions, it can also be used to form directionally selective wavelets in the twelve directions \{30k − 15\} \mod 12. In contrast, the DWT only offers space-scale content over ‘two and a half’ directions, namely vertical, horizontal, and a ‘diagonal’ direction which mixes both the +45° and −45° orientations together.

3. THE DTCWT-LSW MODEL

With complex-valued wavelets, the DTCWT-LSW model reads similarly to the real-case, namely

$$X = \sum_\eta w_\eta \xi_\eta \ast \bar{\psi}_\eta,$$

but now the weights \(w_\eta; \mathbb{N}^2 \rightarrow \mathbb{C}\) can be complex-valued. The ACW, Eqn. (1) is hereby generalised to the complex case valued vectors. We hence propose the following definitions as natural generalisations of those found in the LSW framework.

**Definition 3** (Complex ACW). Let the superscript \(^\ast\) denote complex conjugation. Define the ACW associated with the possibly complex-valued wavelet \(\psi\) by

$$\Psi_\eta = \psi_\eta \ast \bar{\psi}_\eta^\ast = \sum_t \psi_\eta(t) \psi_\eta^\ast(-t). \quad (5)$$

**Definition 4** (DTCWT-LSW biasing matrix). Let \(\Psi\) be the complex ACW as above. Then the associated biasing matrix is

$$A_{\eta,\nu} = \langle \Psi_\eta, \Psi_\nu \rangle = \sum_t \Psi_\eta(t)^\ast \Psi_\nu(t). \quad (6)$$

3.1. Properties

With the DTCWT-LSW defined, we now establish some key results which confirm that the biasing matrix of the proposed model can be inverted under very mild conditions.

**Lemma 5.** Let \(\Psi\) be the ACW (Defn. 1). The entries of the biasing matrix (Defn. 4) are positive, real, and symmetric.

**Proof.** By construction, \(A\) is a Gram matrix. Hence, symmetry follows immediately. We then invoke Plancherel to write \(A = \langle \psi_\eta \ast \bar{\psi}_\eta^\ast, \psi_\nu \ast \bar{\psi}_\nu^\ast \rangle = \langle |\psi_\eta^\ast|^2, |\psi_\nu^\ast|^2 \rangle\).

**Theorem 6** (Invertibility of the biasing matrix). Let the biasing matrix \(A\) associated with an appropriately chosen real-valued wavelet \(\psi\) be non-singular. Then, the biasing matrix \(A\) (Defn. 4) associated with the dual-tree wavelet (Eqn. 4) \(\psi = \psi_{0,0} + i \psi_{0,1}\) is non-singular with \(A = 2A^\prime\).

**Proof.** We use the same arguments as the proof Lemma 5 and note that the Fourier magnitudes of the real and imaginary parts of the dual-tree wavelets are identical by design: \(A = \langle \psi_\eta \ast \bar{\psi}_\eta^\ast, \psi_\nu \ast \bar{\psi}_\nu^\ast \rangle = \langle |\psi_\eta^\ast|^2, |\psi_\nu^\ast|^2 \rangle = \langle |\psi_{0,0}^\ast|^2 + |\psi_{0,1}^\ast|^2, |\psi_{0,0}^\ast|^2 + |\psi_{0,1}^\ast|^2 \rangle = 2\langle |\psi_{0,0}^\ast|^2, |\psi_{0,0}^\ast|^2 \rangle\).

**Remark 7.** Thm. 6 also holds for analytic wavelets since the real and imaginary components have equal magnitude.

**Proposition 8** (DTCWT-LWS covariance inversion). Assume that the debiasing matrix \(A\) is non-singular; then

$$S_\eta(z) = \sum_\nu A_{\eta,\nu}^{-1} \sum_t c(t) \Psi_\nu^t(t).$$

**Proof.** As in [1], we follow a proof by verification and then note that \(A\) is both symmetric and, by Lemma 5, real:

$$\sum_\nu A_{\eta,\nu}^{-1} \sum_t S_\eta^t(z) \Psi_\nu^t(t) = \sum_\eta S_\eta(z) \sum_\nu A_{\eta,\nu}^{-1} \sum_t \Psi_\eta^t(t) \Psi_\nu^t(t) = \sum_\eta S_\eta(z) \sum_\nu A_{\eta,\nu}^{-1} \sum_{\eta^\prime} \delta_{\eta,\eta^\prime} = S_\eta(z).$$

3.2. Simulation

If the local wavelet spectrum of a process is known one can readily sample fields directly from the LSW and DTCWT-LSW models. Furthermore, as described in Algorithm 1, one can also simulate fields from data using the estimated LWS \(S\). A telling comparison between LSW and DTCWT-LSW can be seen in Fig. 1. We first simulate some textures using the LSW model. The field ‘dwt, band 1’ is created by setting \(S_\eta = 0\) at all bands except the vertical band at one of the scale levels which is set to 1. Likewise ‘dwt, band 3’ uses only the diagonal band. These fields are then ‘resimulated’ with the DTCWT model using the estimated LWS. The top row of Fig. 1 labels these resimulations as ‘dctw sim of dwt...’. In the second row of Fig. 1, the roles are reversed so that the DTCWT is used to simulate and the DWT is used to resimulate. It should be clear that the DTCWT model can reproduce the behaviour of the DWT model much better than the DWT can reproduce the DTCWT. Clearly, the DWT cannot simulate the range of directionality inherent in the DTCWT-LSW model.

**Algorithm 1** Simulation algorithm, cf. Taylor [4]

for (level, orientation) = \(\eta;\) do

compute LWP \(X_{\eta}^\ast\), LWS \(\hat{S}_\eta\), and \(\mu_\eta := T^{-2} \sum_t \hat{S}_\eta(t)\)

simulate: \(\hat{X} = \sum_\eta \hat{w}_\eta \xi_\eta \ast \bar{\psi}_\eta\)

using \(\hat{w}_\eta(\cdot) = \sqrt{\mu_\eta}\)

end for

4. EXPERIMENTS

Non-stationarity testing was carried out using the bootstrapping scheme proposed in [4], namely for a given field \(X\) we test the null hypothesis that \(S_\eta(z)\) is a constant function for all \(z\) and \(\eta\) using the test statistic \(\tau_X := (JK)^{-1} \sum_\eta \hat{\sigma}^2(\hat{S}_\eta)\) which is the mean, over \(J\) scales and \(K\) subbands, of the LWS sample variance computed over pixels. When compared to bootstrapped samples cf. Alg. 2, this affords a statistically rigorous means to evaluate how likely an observed statistic \(\tau_X\) is compared to bootstrapped values \(\tau_X\), assuming stationarity. One would expect a non-stationary field to register a small
p-value. As such this would indicate that the null hypothesis could be rejected. As in [4], experiments were performed on both stationary and non-stationary fields. Unlike [4], we also include experiments on anisotropic fields.

**Algorithm 2** Stationarity detection, cf. Taylor [4]

```plaintext
for (level, orientation) = η; do
    compute LWS \( \hat{S}_\eta \) from (3) and test statistic \( \tau_{\text{obs}} \)
    compute mean stationary spectrum \( \mu_\eta := T^{-2} \sum_{t} \hat{S}_\eta(t) \)
end for
for \( i \) in 1 : \( B \) bootstraps do
    simulate \( \tilde{X}^i \) using Alg. 1 and compute test statistic \( \tau_{\tilde{X}}^i \)
end for
compute p-value \( p := (B + 1)^{-1}(1 + \#(\tau_{\tilde{X}} \leq \tau_{\tilde{X}}^i)) \)
```

4.1. Stationary processes

Stationarity tests were performed on an isotropic first order spatial moving average field with weight parameter 0.9 (process ‘S2’ in [4]). One hundred realisations of this process were simulated. For each realisation, bootstrapping was performed over one hundred simulations, cf. Alg 2. Both the LWS and DTCWT-LWS model registered p-values greater than 0.05 for all realisations. As such, this experiment corroborates findings in [4] and [18] that the test is conservative.

A second test was performed on anisotropic Gaussian noise, simulated using a simple directional mask oriented at 45° with angular width 30° together with a Fourier synthesis method. The LWS again registered p-values greater than 0.05 for all realisations. The DTCWT-LWS resulted in 2 realisations with p-values greater than 0.05 which is closer to the 5% figure one would expect.

4.2. Non-stationary processes

Further experiments were carried out on non-stationary random fields. An isotropic field was generated where pixels in one half were drawn from unit-variance Gaussian white noise. The other half was formed from a Gaussian random field with an order 1 Bessel-Matérn covariance function (cf. process ‘NS2’ from [4]). This experiment was repeated for three different scale values. Table 1 records the number of p-values below 0.05 that resulted from 100 realisations (i.e. the number of times that the null-hypothesis was correctly rejected). The DTCWT-LSW appears to detect the non-stationarity better than the LSW. The extra redundancy may add some value to the estimation or simulation process but this requires further work to fully explain.

A set of fields with covariance \( \exp(-(t - t')^\top DR_\theta(t - t')) \) were used with \( D = \operatorname{diag}(4, 1) \), where the rotation matrix \( R_\theta \) is piecewise constant with \( \theta = 90° \) for \( (t - t') \in t_0 \) and \( \theta \in \{0°, 20°, 40°, 60°\} \) for \( (t - t') \notin t_0 \) and where \( t_0 \) is a central square region, cf. Fig. 2. Note that these fields are anisotropically non-stationary. The number of instances where the null-hypothesis was correctly rejected is recorded in Table 2. Thanks to its superior directionality, the proposed method offers a clear advantage here. Furthermore, as the orientations of the two regions become more similar the non-stationary detection becomes worse for both methods.

<table>
<thead>
<tr>
<th>Method/scale</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTCWT-LSW</td>
<td>51</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>DWT-LSW</td>
<td>25</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method/θ</th>
<th>0</th>
<th>20</th>
<th>40</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTCWT-LSW</td>
<td>99</td>
<td>99</td>
<td>98</td>
<td>66</td>
</tr>
<tr>
<td>DWT-LSW</td>
<td>9</td>
<td>10</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

5. CONCLUSION

Kingsbury’s dual-tree complex wavelet basis was successfully harmonised with Nason and Eckley’s locally stationary wavelet model. We have established some initial properties of the resulting proposed DTCWT-LSW model and have confirmed the enhanced utility afforded by the extra directionality. The pervasiveness of anisotropic non-stationary random fields in image processing and the elementary description that covariance provides means that there is much scope for further investigation including: quality inspection via machine vision (e.g. texture piling [4, 19]); segmentation [5]; sonar/denoising [20, 21], change detection [22]. It is also anticipated that a more complete range of anisotropy testing machinery (e.g. Thon et al. [11]) would provide greater statistical power.
6. REFERENCES


