ROBUST TEXTURE FEATURES FOR BLURRED IMAGES USING UNDECIMATED DUAL-TREE COMPLEX WAVELETS

N. Anantrasirichai, J. Burn and David R. Bull

Bristol Vision Institute, University of Bristol, Bristol BS8 1UB, UK

ABSTRACT

This paper presents a new descriptor for texture classification. The descriptor is rotationally invariant and blur insensitive, which provides great benefits for various applications that suffer from out-of-focus content or involve fast moving or shaking cameras. We employ an Undecimated Dual-Tree Complex Wavelet Transform (UDT-CWT) [1] to extract texture features. As the UDT-CWT fully provides local spatial relationship between scales and subband orientations, we can straightforwardly create bit-planes of the images representing local phases of wavelet coefficients. We also discard some of the finest decomposition levels where are most affected by the blur. A histogram of the resulting code words is created and used as features in texture classification. Experimental results show that our approach outperforms existing methods by up to 40% for synthetic blurs and up to 30% for natural video content due to camera motion when walking.

Index Terms— texture, classification, blur, Undecimated DT-CWT, overcomplete wavelet transform

1. INTRODUCTION

Images and videos captured with hand-held or moving cameras often exhibit blur due to incorrect focus or slow shutter speed, e.g. in low light conditions or with fast camera motion. Blurring effects generally alter the spatial and frequency characteristics of the content and this may reduce the performance of a classifier or an image retrieval process, particularly in the areas of texture where strong structures, such as edges, are not distinguishable.

It is well-known that image blurring is an ill-posed problem, despite being simply expressed with a matrix–vector multiplication as in Eq. 1.

$$\mathbf{I}_{obv} = \mathbf{D}\mathbf{I}_{idl} + \varepsilon \quad (1)$$

Here $\mathbf{I}_{obv}$ and $\mathbf{I}_{idl}$ are vectors containing the observed and ideal images, respectively. Matrix $\mathbf{D}$ represents blur, while ε represents noise. Various approaches have attempted to solve this problem by modelling it as a point spread function (PSF) which is generally unknown. $\mathbf{D}$ is considered as a convolution matrix, and deconvolution can be employed with an iterative process to estimate $\mathbf{I}_{idl}$ [2]. Unfortunately the blur estimation

and blind deconvolution processes are highly complex, so are not appropriate in many cases, particularly for real time applications and for large datasets. Therefore, we propose texture features which are robust to image blurring as opposed to attempting to deblur the textures.

Several techniques for classification and recognition have been proposed for blurred images [3,4]. However, they normally include a deconvolution process in their framework. To the best of our knowledge, only one research group has proposed blur invariant features. Ojansivu and Heikkilä have proposed a blur-insensitive texture descriptor using the local phase of Fourier transforms [5,6]. In practice, the neighbouring pixels are highly correlated in real images leading to dependency between the Fourier coefficients. Therefore, their method requires a decorrelation mechanism. We, instead, employ wavelets which have decorrelation (whitening) properties [7], i.e. correlations between two wavelet coefficients decay exponentially with the distance between them. Wavelet transforms decompose an image into a sum of localised functions with various spatial scales. They also lend themselves to low complexity implementation.

Wavelets are an efficient tool for capturing texture information as their decompositions provide information about image structure, strong persistence across scales and compressible properties. As a result, wavelets have been used in many texture-related applications [8,9]. Here, we employ the undecimated dual-tree complex wavelet transform (UDT-CWT) developed by Hill *et al.* [1]. This is an overcomplete version of the DT-CWT [10] which provides shift-invariance and increased directional selectivity over traditional discrete wavelet transform; hence, it has been used for rotation invariant texture classification for a decade [11].

The proposed method extracts texture features from the UDT-CWT domain. We employ the local phase of each decomposition level to create bit-planes. However, we do not exploit all decomposition levels because the frequency properties of the finer levels are changed by blurring more than those of the coarser levels. We therefore do not include some of the finest levels in the stack of bit-planes. The bit-plane representation can be simply created because the UDT-CWT does not suffer from the problem of mis-alignment between features in the image and features in the bases. Our method can also straightforwardly be combined with a wavelet-based

denoising technique [12] leading to robustness for both blur and noise which generally appear in images and videos taken with commercial digital cameras.

The remainder of this paper is organised as follows. The UDT-CWT and the proposed method are described in Section 2 and Section 3, respectively. The performance of the method is evaluated on a set of images in Section 4. Finally, Section 5 presents the conclusions and future work.

2. UNDECIMATED DUAL-TREE COMPLEX WAVELET TRANSFORM (UDT-CWT)

The dual-tree complex wavelet transform (DT-CWT) was first formulated by Kingsbury [13]. It employs two different real discrete wavelet transforms (DWT) to provide the real and imaginary parts of the CWT. Two fully decimated filter bank (FB) trees are produced, one for the odd samples and one for the even samples generated at the first level. The next decomposition levels are processed using low-pass/high-pass filter pairs, $\{h_0[n], h_1[n]\}$ and $\{g_0[n], g_1[n]\}$ for the first and second FBs, respectively, where $n \in \ell^2(\mathbb{Z})$. As a result, the DT-CWT increases directional selectivity over the DWT and is able to distinguish between positive and negative orientations giving six distinct sub-bands at each level, corresponding to $\pm 15^\circ, \pm 45^\circ, \pm 75^\circ$.

Hill *et al.* developed the undecimated dual-tree complex wavelet transform (UDT-CWT) [1], where subsampling of the DT-CWT is avoided by using the upsampled q-step filters which are defined recursively as in Eq. 2.

$$\mathcal{G}_\kappa^{(l+1)}[n] = \mathcal{G}_\kappa^{(l)}[n] \uparrow 2 = \begin{cases} \mathcal{G}_\kappa^{(l)}[\frac{n}{2}] & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd} \end{cases} \quad (2)$$

where $\mathcal{G}_\kappa^{(l+1)}[n]$ is a UDT-CWT filter at level $l+1$ and $\mathcal{G} \in \{h, g\}$, $\kappa \in \{0, 1\}$. $\mathcal{G}_\kappa^{(1)}[n]$ is a non upsampled q-step filter. The undecimated version fully achieves a one-to-one spatial relationship between co-located coefficients in different sub-bands and decomposition levels.

3. ROBUST TEXTURE FEATURES

3.1. Blur image analysis

Ignoring noise and applying the Fourier transform, $\mathcal{F}(\bullet)$, to the convolution model of Eq. 1 ($i_{obv} = d * i_{idl}$, lower case letters are used here for 2D matrices), the magnitude and phase are as shown in Eq. 3.

$$|\mathcal{F}(i_{obv})| = |\mathcal{F}(d)| \cdot |\mathcal{F}(i_{idl})| \quad (3a)$$

$$\angle \mathcal{F}(i_{obv}) = \angle \mathcal{F}(d) + \angle \mathcal{F}(i_{idl}) \quad (3b)$$

As $|\mathcal{F}(d)| \leq 1$, the magnitudes of the blurred image in the frequency domain are always decreased. If the blur PSF d is centrally symmetric, $\angle \mathcal{F}(d) = 0$ for all $|\mathcal{F}(d)| \geq 0$, i.e.

Fig. 1. Magnitudes and phases of wavelet coefficients. Left: level 2. Right: level 4.

$\angle \mathcal{F}(i_{obv}) = \angle \mathcal{F}(i_{idl})$ [14]. In the ideal case, the phase is very robust to the blur. Although the cross-section of the PSF is rectangular in practice, $\angle \mathcal{F}(d)$ is still close to zero at its centre.

DT-CWT coefficients possess similar properties to Fourier coefficients. The complex phases of its coefficients correspond to the angles of dominant directional features in their support regions. When the image is blurred, the phase response is slightly expanded from the centre (e.g. ridge or edge in the image) and the change of the phase is still approximately linearly proportional to the displacement in a direction orthogonal to the subband orientation.

Fig. 1 and Fig. 2 show frequency responses in the wavelet domain of a step function and its blurred version created using a Gaussian blur kernel with a standard deviation of 1. In Fig. 1, the solid and dashed lines are the magnitudes and phases, respectively, whereas they represent the real and imaginary coefficients in Fig. 2. The red signals are generated from the blurred image, while the blue signals are from the sharp image. Fig. 1 shows that the phases are approximately linear around the centre and the phases of sharp and blurred images are not significantly different, while the difference of the magnitudes is noticeable. Fig. 2 reveals that the coarser levels are more robust than the finer levels. The magnitudes of both real and imaginary parts have decreased because of the blur kernel, but the phases have only changed slightly. The coefficients of the finer levels have been affected more; therefore, they should not be included in the descriptors. The finer levels also exhibit higher noise proportional to structure information [15]. This means our proposed method should also be more robust to noise.

3.2. Undecimated complex wavelet-based descriptors

We denote $w_s^l[n]$ as a wavelet coefficient of subband s , $s \in \{1, \dots, 6\}$ at level $l \in \{1, \dots, L\}$, where L is the total number of decomposition levels. We also assume that the real and imaginary parts of the transform's outputs are treated as separate coefficients so that $\mathcal{G}_\kappa$ is a real matrix thereby producing real values for the real and imaginary parts. That is, $\mathbf{w} = \mathcal{G}_\kappa \mathbf{x}$, $w_s^l[n] \in \mathbf{w}$ for a real image $\mathbf{x}$.

In order to generate descriptors, we extract texture information from the wavelet decomposition from level l_0 to level

Fig. 2. Wavelet coefficients of 4 levels of step edge and its blur version

L , where l_0 is the finest included level. Empirically $l_0 = 2$ gives desirable results for classification problem of blur texture images. Since the blur kernel causes the coefficient magnitudes to change more than the phases (as shown in Fig. 2), we convert the wavelet coefficients of subband s to a binary number according to the signs of the wavelet coefficients. The real and imaginary parts are then defined as in Eq. 4.

$$b_{s,l}^{Re}[n] = \begin{cases} 1 & \text{if } \text{Re}(w_s^l[n]) > 0 \\ 0 & \text{if } \text{Re}(w_s^l[n]) \leq 0 \end{cases} \quad (4a)$$

$$b_{s,l}^{Im}[n] = \begin{cases} 1 & \text{if } \text{Im}(w_s^l[n]) > 0 \\ 0 & \text{if } \text{Im}(w_s^l[n]) \leq 0 \end{cases} \quad (4b)$$

Subsequently, these values are combined across levels as defined in Eq. 5 to create a code word for each pixel of the image.

$$q_s[n] = \sum_{l=l_0}^L \left(b_{s,l}^{Re}[n] \cdot 2^{2(l-l_0)} + b_{s,l}^{Im}[n] \cdot 2^{2(l-l_0)+1} \right) \quad (5)$$

For each subband, a histogram of q_s is obtained with $2^{2(L-l_0+1)}$ bins. The frequency of each bin is used as one feature leading to a total of $6 \cdot 2^{2(L-l_0+1)}$ features for all 6 subbands. The histogram of each subband is individually created so that the features overcome asymmetric blur kernels, e.g. motion blur. Note that wavelets also possess decorrelation properties; hence, a whitening transform, as used in the local phase quantization (LPQ) method [5], is not required.

4. RESULTS AND DISCUSSION

We evaluated our texture features using both synthetic blur and real motion blur. In both cases, the classifier was trained using sharp images and then tested with blurred images, allowing assessment of the performance of the descriptors –

specifically how they cope with the frequency alterations caused by the blur kernels. We employed a support vector machine (LIBSVM for MATLAB platform) [16] for the classification process. The results of the classification are compared to those of the local binary pattern (LBP) [17] and rotation invariant LPQ (riLPQ) [5] methods.

4.1. Synthetic blur

We employed the grayscale Outex_TC_00000 dataset [18], which includes 480 images of size 128×128 pixels. There were total of 24 classes, with 20 images per class. The experiments were created with 100 test problems and each result was the mean of these problems. Each problem randomly selected half of the dataset for training – this was pre-defined by [14]. The other half of the images were artificially blurred using i) a Gaussian lowpass filter with several standard deviations σ , ii) the linear motion of a camera with various ranges of movement in 5 directions (5, 23, 45, 75 and 86 degrees). The first case implies an out-of-focus problem, whilst the second case often occurs in general use of hand-held cameras. We tested our proposed method using $L = \{3, 4\}$ and $l_0 = 2$, while using 12 and 36 orientations for the riLPQ method.

Fig. 3 shows the classification results of the Gaussian blur case. The result where $\sigma = 0$ means that the blur kernel was not employed, so the test images were sharp. Our proposed UDT-CWT with $L = 4$ (total 392 features) gave the best classification accuracy and very robust results when the test images were very blurred. The proposed UDT-CWT with $L = 3$ gave worse results but provided better computational cost (total 96 features). The riLPQ with 32 orientations (total 256 features) gave worse results and required more computational time than our UDT-CWT with $L = 3$. The LBP (total 59 features) gave the worst results here, since the blur kernel decreased the difference between the intensity values amongst pixel neighbours.

The classification results of the motion blur case are shown in Fig. 4, which plots the average values of all 5 angles. The UDT-CWT descriptors with $L = 4$ shows significantly better results than the others with up to 40% improvement in classification accuracy for larger blurs. We also tested the DT-CWT method with $l_0 = 1$ and found that the classification accuracy is decreased by 10% at a motion length of 5 pixels. This supports our proposal of discarding the finest decomposition level.

4.2. Real video

We next applied the texture descriptors to sequences which were acquired using a digital camera attached to the forehead while walking. The video resolution was 1920×1080 pixels in RGB format with 30 frames per second. We classified regions appearing in each frame into 3 classes: 1) hard surfaces (tarmac, bricks, tiles, etc.), 2) soft surfaces (grass, soil, sand,

Fig. 3. Average classification accuracy when using Gaussian blurred textures

Fig. 4. Average classification accuracy when using motion blurred textures

etc.), and 3) unwalkable areas (moving objects, fence, plants, etc.). For the training process, the colour images were converted to greyscale and only sharp frames from a number of videos were used. They were segmented into various sizes of regions to generate 1000 samples for each class. The test frames were taken from different regions of the videos than those used for training.

Fig. 5 shows the classification accuracy and the sharpness value of each frame. The sharpness value (green dotted line) was computed from the mean of gradient magnitudes. The plot clearly shows the points where the camera was moving faster (less sharp). The red and blue lines are the results of the proposed UDT-CWT method ($L = 4$) and the riLPQ method (36 orientations), respectively. Our method outperforms riLPQ by up to 30% with the improvement in performance correlated with the degree of blurring. Fig. 6 shows

Fig. 5. Classification accuracy per frame of video captured while walking

Fig. 6. Subjective results of sharp (left) and blur (right) frames. The label 1 (green), label 2 (red) and label 3 (blue) are the areas classified as hard surfaces, soft surfaces and unwalkable areas, respectively.

the results of the sharp and the blur frames computed from the proposed method. There are some incorrect classifications in the blurred image but most areas are predicted correctly (74% accuracy).

5. CONCLUSIONS AND FUTURE WORK

This paper presents a new rotationally invariant and blur insensitive texture descriptor. The texture information is extracted using the undecimated dual-tree complex wavelet transform (UDT-CWT) which provides shift-invariance and directional selectivity. The descriptor is computed from the binary pattern of the phase of the coarse decomposition levels. The proposed features have been tested with synthetically blurred images using Gaussian and motion blur kernels, as well as videos exhibiting real motion blur. The classification results show that our method significantly outperforms the existing methods, particularly for images with high motion blur. In the future, the interlevel relationship of complex wavelets will be included to improve a blur robustness and simultaneously remove noise.

6. REFERENCES

- [1] P. Hill, A. Achim, and D. Bull, "The undecimated dual tree complex wavelet transform and its application to bi-variate image denoising using a cauchy model," in *Image Processing (ICIP), 2012 19th IEEE International Conference on*, Sept 2012, pp. 1205–1208.
- [2] N. Anantrasirichai, J. Burn, and D. R. Bull, "Projective image restoration using sparsity regularization," in *IEEE Int. Conf. Image Processing*, 2013.
- [3] I. Stainvas and N. Intrator, "Blurred face recognition via a hybrid network architecture," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 2, 2000, pp. 805–808.
- [4] P. Vageeswaran, K. Mitra, and R. Chellappa, "Blur and illumination robust face recognition via set-theoretic characterization," *Image Processing, IEEE Transactions on*, vol. 22, no. 4, pp. 1362–1372, April 2013.
- [5] V. Ojansivu, E. Rahtu, and J. Heikkilä, "Rotation invariant local phase quantization for blur insensitive texture analysis," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, Dec 2008, pp. 1–4.
- [6] J. Heikkilä, E. Rahtu, and V. Ojansivu, "Local phase quantization for blur insensitive texture description," in *Local Binary Patterns: New Variants and Applications*. Springer Berlin Heidelberg, 2014, vol. 506.
- [7] A. Tewfik and M. Kim, "Correlation structure of the discrete wavelet coefficients of fractional brownian motion," *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 904–909, Mar 1992.
- [8] S. Livens, P. Scheunders, S. Livens, G. V. D. Wouwer, P. Vautrot, and D. V. Dyck, "Wavelet-based texture analysis," *Int. Journal of Computer Science and Information Management, Special issue on Image Processing (IJC-SIM)*, vol. 1, 1997.
- [9] N. Anantrasirichai, A. Achim, J. E. Morgan, I. Erchova, and L. Nicholson, "SVM-based texture classification in optical coherence tomography," in *International Symposium on Biomedical Imaging*, 2013.
- [10] I. Selesnick, R. Baraniuk, and N. Kingsbury, "The dual-tree complex wavelet transform," *Signal Processing Magazine, IEEE*, vol. 22, no. 6, pp. 123 – 151, Nov. 2005.
- [11] P. Hill, D. Bull, and C. Canagarajah, "Rotationally invariant texture features using the dual-tree complex wavelet transform," in *Image Processing, 2000. Proceedings. 2000 International Conference on*, vol. 3, Sept. 2000, pp. 901–904.
- [12] A. Khare and U. Shanker Tiwary, "A new method for deblurring and denoising of medical images using complex wavelet transform," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, Jan 2005, pp. 1897–1900.
- [13] N. G. Kingsbury, "The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters," *IEEE Digital Signal Processing Workshop*, 1998.
- [14] V. Ojansivu and J. Heikkilä, "Blur insensitive texture classification using local phase quantization," in *Image and Signal Processing*, 2008, vol. 5099.
- [15] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.
- [16] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011.
- [17] T. Ojala, M. Pietikainen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971 –987, Jul. 2002.
- [18] *Outex Texture Database*. Center for Machine Vision Research, University of Oulu, 2013. [Online]. Available: <http://www.outex.oulu.fi/index.php?page=contributed>