16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

STATISTICAL DETECTION OF SALIENT POINTS USING THE DUAL TREE *M*-BAND WAVELET TRANSFORM

W. Ayadi[1,2]*,S. Sevestre-Ghalila*[2,3]*, A. Benazza-Benyahia*[1]

[1] URISA,
Ecole Supérieure des Communications,
Tunisia
walid.ayadi@gmail.com
benazza.amel@supcom.rnu.tn

[2] MAP5, Université Paris 5,
France
[3] U2S, ENIT,
Tunisia
sylvie.sevestre@math-info.univ-paris5.fr

ABSTRACT

In this paper, we develop a novel method to detect salient points in an image from a multiresolution representation. Our contribution is twofold. Firstly, the multiscale representation results from the Dual Tree Wavelet Transform (DTWT) since it enables a great directional selectivity with a reduced redundancy ratio. The second novelty of our work relies on the reliable outliers statistical tests that we apply to detect salient points from the DTWT coefficients. The experiments show the robustness of the approach to noise.

1. INTRODUCTION

During the past years, a great interest has been given to globally describe any image by defining appropriate signatures built on its keypoints or Salient Points (SP). By SP, we mean pixels carrying enough information about the local neighborhood so that they will be distinct from their nearest neighbors. They are also characterized by their robustness to scaling, rotation and illumination changes. Their positions are invariant with respect to geometric and radiometric distortions [1]. Typical examples of SP are blobs, corners and junctions. SP detection is a key issue for various matching problems in computer vision such as image retrieval [2, 3], object recognition [4]. SP detectors can be classified in two categories. The first one employs contours [5]. However, this class of detectors is not always the most suitable one since edge extraction is very sensitive to noise and the contour chaining may not be well in cluttered scenes. In the second class of detectors, the Harris detector [6] is the most employed technique. It consists in locating the local changes in the image by using the first derivative. This detector is robust to image rotation, additive noise and illumination changes but it is not designed for deriving SPs at different scales. Recently, it has been modified to work as a scale-invariant detector [7]. Another solution is to perform the SP detection in a transform domain which is expected to make the problem easier to solve. To this respect, multiscale transforms are suited to this task since they decompose the input image into different scales and orientation that are coherent with the human perception. In [3], a multiresolution framework is retained to detect global as well local image variations and the interest points are assumed to correspond to areas where the local contrast is the highest. However, the lack of shift invariance and the directional selectivity of the wavelet transform has motivated the choice of more sophisticated scale-space transforms [8]. In [9], the dual tree complex wavelet transform is used to generate an "accumulated energy map" that enables the keypoint selection. However, it is worth noting that heuristic criteria have been considered at the detection step. In this paper, we propose a novel multiscale keypoint detection. Our approach departs from the conventional ones: instead of using empirical criteria, SP extraction is based on outliers statistical tests performed at different scales. Furthermore, we will operate in the Dual Tree *M*-band Wavelet Transform (DTWT) since it has been recognized to offer a great directional selectivity and to be nearly shift invariant with a limited redundancy w.r.t. the dual tree complex wavelet transform [10]. The rest of this paper is organized as follows. Section 2 is dedicated to a review of DTWT. In section 3, we describe the statistical behavior of wavelet details coefficients. In Section 4, we briefly describe the most pertinent outliers statistical tests. Then, in Section 5, we explain how we apply them for detecting SP from the resulting DTWT coefficients. In Section 6, experimental results are given and some conclusions are drawn in Section 7.

2. A BRIEF REVIEW OF *M*-BAND DTWT

Multiresolution decomposition allows a multiscale analysis of the details contained in the image corresponding to physical structures. An M-band wavelet transform of $\mathscr{L}^2(\mathbb{R})$, (where $M \in \mathbb{N}^*$) is considered as a very versatile multiscale decomposition. It is characterized by a scaling function $\phi \in \mathscr{L}^2(\mathbb{R})$, and $M-1$ wavelet functions $\psi_m \in \mathscr{L}^2(\mathbb{R})$, $(m = 1,\ldots,M-1)$ that satisfy the following dilation equations:

$$\forall t \in \mathbb{R}, \phi(t) = \sqrt{M} \sum_{k\in\mathbb{Z}} h_0(k)\phi(Mt-k) \qquad (1)$$

$$\psi_m(t) = \sqrt{M} \sum_{k\in\mathbb{Z}} h_m(k)\phi(Mt-k), \qquad (2)$$

where the $h_m \in \ell^2(\mathbb{Z})$. The set $\cup_{m=1}^{M-1}\{M^{-j/2}\psi_m(M^{-j}t-k)\}_{j,k\in\mathbb{Z}}$ is an orthonormal basis of $\mathscr{L}^2(\mathbb{R})$ if the para-unitary condition holds for any couple (m,m') in $\{0,\ldots,M-1\}^2$:

$$\sum_{p=0}^{M-1} \hat{h}_m(\omega + p\frac{2\pi}{M})\hat{h}'_m(\omega + p\frac{2\pi}{M}) = M\delta_{m-m'}, \qquad (3)$$

where $\hat{}$ denotes the Fourier transform. Therefore, h_0 is a low-pass filter, $h_1,\ldots,h_{M-2}$ are band-pass filter and h_{M-1} is a high-pass filter. Consequently, the expansion of any 1D signal $f \in \mathscr{L}^2(\mathbb{R})$ over J resolution levels can be expressed

as:

$$\begin{aligned} f(t) &= \sum_{k\in\mathbb{Z}} a_J(k) M^{J/2}\phi(M^J t-k) \\ &+ \sum_{j\geq J}\sum_{m=1}^{M-1}\sum_{k\in\mathbb{Z}} M^{j/2} d_{m,j}(k)\psi_m(M^j t-k). \end{aligned} \quad (4)$$

The coefficients a_J correspond to a coarse version of f whereas $d_{m,j}$ correspond to details at the scale j in the m-th band. The M-band decomposition is applied to 2D signals of $\mathscr{L}^2(\mathbb{R}^2)$, in a separable manner. At each resolution level, M^2 sub-bands are then generated.

A *dual* M-band multiresolution analysis is derived thanks to a scaling function ϕ^H and mother wavelet functions ψ_m^H, $(m=1,\ldots,M-1)$ that are related to the primal wavelet functions ψ_m^H, $(m=1,\ldots M-1)$ through a Hilbert transform. More precisely, for $m=1,\ldots,M-1$, we obtain the following relationships in the Fourier domain:

$$\widehat{\psi_m^H}(\omega) = -\mathrm{i}sign(w)\widehat{\psi_m}(\omega), \quad |\widehat{\psi_m^H}(\omega)| = \widehat{\psi_m}(\omega) \quad (5)$$

where *sign* is the signum function. In a similar way, it is easy to obtain scaling equations for $\widehat{\psi_m^H}$ involving real-valued sequences $(g_m[k])_{k\in\mathbb{Z}}$. A dual M-band orthogonal wavelet basis of $L^2(\mathbb{R})$ is obtained if also $\hat{g}_m$ satisfies para-unitary conditions. The 2D dual tree M-band wavelets decomposition can be generalized from the 1D dual-tree decomposition in a separable manner for the primal and the dual tree. Furthermore, it is advantageous to add preprocessing and postprocessing steps to the whole DTWT [11]. More precisely, preprocessing enables to extend the formalism of DTWT concerning analog images to the underlying discrete images whereas a linear postprocessing aims at improving the directional selectivity of the decomposition [10]. As a result, the decomposition ouputs postprocessed coefficients of the primal and the dual tree, respectively denoted by $\delta_{j,m}[k,l]$ and $\delta_{j,m}^H[k,l]$ where $m=1,\ldots,M^2-1$ and $[k,l]$ are the spatial indices. Fig. 1 provides a generic block-diagram of the 2D DTWT.

After this brief overview on M-band DTWT, we are interested in studying the statistical properties of the resulting coefficients in order to built efficient SP detectors.

3. A STATISTICAL BEHAVIOUR OF WAVELET DETAILS COEFFICIENTS

In [12], the distribution of the wavelet coefficients could be modelized by a Generalized Gaussian (GG) distribution. The GG distribution depends on three parameters (μ,α,β) where μ, α and β are respectively location, scale and shape parameters. An example of fitting in Fig. 2 shows the histograms of wavelet coefficients of the cameraman image for $M=2$ and $J=1$. As well as many generalized distributions, the GG distribution derives from *power transformation* introduced by Box-Cox [13]. The goal of the power transformations is to preprocess data from any distribution to data that could be modelized by a Gaussian distribution. Moreover, if we assume that the data comes from a GG variable $Z(\mu,\alpha,\beta)$, the power transformation $f(Z)=(\frac{|Z-\mu|}{\alpha})^{\beta/2}$ provides data that could be modelized by a Gaussian distribution.

Most of the reported Outliers Statistical Test (OST) require that the considered data have a Gaussian distribution. Therefore, it seems natural to preprocess the wavelet coefficients at each scale j and each subband m in order to transform their distribution into a Gaussian distribution. More precisely, we map the wavelet coefficient $\delta_{m,j}[k,l]$ at each location (k,l) to

$$\zeta_{m,j}[k,l] = f(\delta_{m,j}[k,l]) = (\frac{|\delta_{m,j}[k,l]-\mu_{m,j}|}{\alpha_{m,j}})^{\beta_{m,j}/2} \quad (6)$$

where $\mu_{m,j}$,$\alpha_{m,j}$ and $\beta_{m,j}$ are the GG density parameters that could be estimated by maximum likelihood estimation [14]. The same mapping is applied to the second postprocessed ouputs coefficients of the primal and the dual tree $\delta_{m,j}^H[k,l]$ to obtain the related $\zeta_{m,j}^H[k,l]$. In Fig. 2, the histograms of ζ of the cameraman image for $M=2$ and $J=1$. The shape of these histograms are close to those of Gaussian data with a distortion at the right due to the power transformation. Indeed, the negative and positive coefficients with significant magnitudes simultaneously contribute to high positive values. In the sequel, we will show that the outliers correspond to the distorted parts of the transformed coefficients histograms.

4. A BRIEF REVIEW OF OUTLIERS STATISTICAL TESTS

After performing the power transformation mapping, it is possible to carry out the detection of the SPs among the wavelet coefficients.

An outlier is defined as an observation generated from a different distribution from the one associated to the whole data sample. In this paper, we assume that SPs can be obtained by locating wavelet coefficients at each orientation and scale that have significant magnitude related to those of their neighbour coefficients. As most of the OST require that the considered data have a Gaussian distribution, we define a wavelet coefficient that have significant magnitude as an outlier of the transformed coefficients ζ and ζ_H introduced in the previous section. Let $\zeta^N=\{\zeta(1),\ldots,\zeta(N)\}$ be the transformed coefficients *sorted* in ascending order of a window where the central coefficient has the maximal value $\zeta(N)$. Resulting from the power transformation of Eq.(6), we can assume that ζ^N is an independent realization of a *Gaussian* random variable. Generally, an outlier corresponds to a measure that has a significantly different magnitude w.r.t. the remaining measures in the data set. If we assume that all the values in ζ^N are non-negative, the maximum value $\zeta(N)$ is considered as a suspect value. Therefore, the problem amounts to decide whether $\zeta(N)$ is an outlier or not and more precisely to a decision problem with the two following hypothesis H_0 and H_1:

$$H_0 : \zeta(N) \sim G, \quad H_1 : \zeta(N) \nsim G \quad (7)$$

where G is the density function of the transformed coefficients. As all statistical tests, the decision is based on a statistic T whose distribution is well known under the hypothesis H_0. If $\zeta(N) \nsim G$, the value t_s of the statistic T computed on the sample leads to a small value of the $p\text{-}value = 1-F(t_s)$ where F is the distribution function of T under H_0. The final decision is obtained by the use of a critical value $t_c = F^{-1}(1-a)$ where a is a significance level of adjusted by the user:

$$\zeta(N) \nsim G \quad \text{if} \quad t_s > t_c = F^{-1}(1-a). \quad (8)$$

The final decision is therefore equivalent to:

$$\zeta(N) \nsim G \quad \text{if} \quad p-value < a. \quad (9)$$

Among the most known OST, more attention is paid to the Dixon test and the Grubbs test. The first one (the Dixon test) is based on the difference of the suspect value $\zeta(N)$ from its nearest one $\zeta(N-1)$ related to the range of all the values of sample [15]. The Dixon test employs an experimental statistic T_{Dixon} called the Q-value defined as follows:

$$T_{Dixon} = \frac{\zeta(N) - \zeta(N-1)}{\zeta(N) - \zeta(1)}. \quad (10)$$

T_{Dixon} is compared to a threshold Q_{crit} which could be viewed as a critical Q-value found in the Dixon table only for $N < 30$. If $T_{Dixon} > Q_{\text{crit}}$, the suspect value can be judged as an outlier and consequently as a suspect salient point. The decision criterion based on the maximum of sample is similar to the classical criteria applied in SP detection in the wavelet transform domain reported in [16][9]: a coefficient $\delta_{j,m}[k,l]$ is retained if $|\delta_{j,m}[k,l]| > \lambda . \max_{(k',l')} |\delta_{j,m}(k',l')|$ where λ is a threshold lower or equal to 1 subjectively adjusted. It is worth pointing out that our contribution consists in aplying a critical threshold deriving from a statistical test.

The Grubbs test evaluates how far the suspect value is from the other sample measures [15]. The distance is computed w.r.t the arithmetic mean of the data $\bar{\zeta}$ normalized by the empirical standard deviation of the sample s. The statistic T_{Grubbs} of this test is therefore:

$$T_{Grubbs} = \frac{|\zeta(N) - \bar{\zeta}|}{s}. \quad (11)$$

The suspect value $\zeta(N)$ is detected as an outlier if $T_{Grubbs} > Z_{\text{crit}}$ with:

$$Z_{\text{crit}} = \frac{(N-1)}{\sqrt{N}} \left(\frac{t^2_{(a/N,N-2)}}{N-2+t^2_{(a/N,N-2)}} \right)^{1/2}, \quad (12)$$

and where $t^2_{(a/N,N-2)}$ denotes the critical value of the Student distribution with $N-2$ freedom degrees and a significance level equal to a/N.

After recalling the basic background on DTWT and OST, we are now able to describe the SP detection method we propose.

5. PROPOSED SP EXTRACTION BY OUTLIERS STATISTICAL TESTS

Here, SPs correspond to wavelet coefficients at each orientation and scale that have significant magnitude related to those of their neighbour coefficients. More precisely, we consider that a SP has significant coefficients in at least two wavelet subbands at the same scale, oriented in different directions. Indeed, for a fixed scale and for each location, if two wavelet coefficients of two different subbands are judged as two outliers, then this position corresponds to SP location for that scale. Therefore, the SPs extraction algorithm is applied in the DTWT domain through OST according to the following steps.

• **Step 1**: **Power transformation**. We start by estimating the parameters of the GG distribution of each subband (m, j). Then, we apply the aforementioned power transformation of Eq. (6) in order to obtain transformed wavelet coefficients that approximately have a Gaussian density.

• **Step 2**: **Outliers detection**. In each subband (m, j) of transformed coefficients, a sliding window of size (n_w, m_w) centred at the current location $[k,l]$ is used. At this point, we test whether a suspect value is an outlier coefficient or not. More precisely, a suspect value is found when the coefficient located at the center of the sliding window takes the maximum value within that window. Giving a significance level a, either Dixon or Grubbs tests is applied in order to judge if the window's center is an outlier related to its neighbours. As a result, we obtain sets of outlier coefficients in all the possible orientations m.

• **Step 3**: **SP detection of the considered scale**. The singularity of SP is exploited by this process: a location is judged as a SP if two outlier coefficients are at least detected at the step 2 in the same position by scanning all the subbands at the same scale j. Consequently, we obtain sets of possible SP in all possible orientations (m, m').

We summarize the detection results at all scales: we locate each SP of the first scale on the image by a circle with a radius proportional to the last scale j at which it was detected. Fig. 3(a) shows an example of SPs detection by using the Grubbs test for $M = 4$, $J = 2$ and $a = 0.1$ on a synthetic and binary image.

A parallel illustration of our SP detection approach is provided by a $p-value$ map based on the existing relationship between outlier decision of the statistical test and its $p-value$. Indeed, we deduce from Eq. (9) that an outlier is a suspect coefficient with small $p-value$ w.r.t the significance level a. Consequently, each subband coefficient is classified into one of the 3 following classes.

• Outlier coefficient: a suspect coefficient such as $p-value < a$.

• Suspect but not an outlier coefficient: a suspect point such as $p-value > a$.

• Not a suspect coefficient: the window's center coefficient is not the maximum of that neighborhood. The statistical test is not applied and then no $p-value$ is provided. We set arbitrarily the $p-value$ to 1 which is the maximum possible value of $p-value$.

As a result, the illustration is performed according to the following steps of the algorithm.

• At step 2: at each subband (m, j), we assign a $p-value$ to each coefficient according to its class (outlier, suspect, not suspect).

• At step 3: at each scale, a SP is detected at a position $[k,l]$ if this position corresponds to outlier coefficients in at least 2 subbands. The related $p-value$ are less than the significance level a. Consequently, their maximum is less than a and we retain this maximum value in the $p-value$ map at location $[k,l]$ at scale j. By convention, we have decided that a SP is detected in the full resolution image if its corresponding coefficients are SP in all the resolution levels.

Fig. 3(b) shows an example of the global $p-value$ map of a test image.

6. EXPERIMENTAL RESULTS

In our experiments, two gray images were used as test images. The first one was artificially created containing junctions of different sizes at different scales. The second one is the "cameraman" image. Our detector is compared to the detector (say detector A) based on the accumulated map of keypoint energies proposed in [9]. For both detectors, we will report the results achieved with the best parameters adjust-

ment. More precisely, we have tested $M = 2$ and $J = 2,3,4$ for the DTCWT related to the detector A. Concerning our detector, the DTWT was tested with $M = 2,3,4$ and $J = 2,3,4$. The best results of detector A are obtained by using the Dual Tree Complex Wavelet Transform (DTCWT) with $M = 2$ and $J = 4$ as proposed in [9] whereas our detector operates with Meyer wavelet transform for the DTWT with $M = 4$ and $J = 2$. We can note that increasing the subband number M improves the detection results since it yields a greater number of spatial analysis directions. Our approach require to tune the size of the sliding window and the significance level of test a. We have employed windows of size 5×5 because the critical Q-value of the Dixon test is tabulated only for samples of size less than 30. Concerning a, its adjustment is different for each OST: Grubbs test requires a small significance level in order to detect nearly all SPs of the image unlike Dixon test. As shown in Fig. 4 and Fig. 5, for the same value of a, our approach with Grubbs test detects SPs more than with Dixon test but we observe more false SPs detection with Grubbs test. In the case of the "cameraman" image, the significance level a must be set to a smaller value than for synthetic images which are noiseless. Even when the window size is increased, SPs detector is not improved for both tested images. Fig. 6 displays the ouptut of detector A. We observe that the results for the synthetic image are quite similar for both detectors. The main difference between the two detectors is the grass region of the cameraman image. Our approach is less sensitive to the presence of noise or micro-texture.

7. CONCLUSION

We presented a new approach to perform multiscale salient points detection based on a dual tree M-bands wavelet decomposition and outliers statistical tests. Junctions of different sizes at various scales can be detected. Compared to the detector proposed in [9], our approach seems to be more robust to noise.

REFERENCES

[1] R. Harrilick and L. Shapiro, "Computer and Robot Vision II," *Addison-Wesley*, 1993.

[2] C. Schmid and R. Mohr, "Local gray value invariants for image retrieval," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530-535, May 1997.

[3] S. Bres, J. M. Jolion, "Detection of interest points for image indexing," *International Conference on Visual Information Systems*, pp. 427-434, June 1999.

[4] D. G. Lowe, "Object recognition from local scale-invariant keypoints," *Proc. of Int. Conf. on Computer Vision*, pp. 1150-1157, 1999.

[5] H. Asada, M. Brady, "The curvature primal sketch," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 2-14, 1986.

[6] C. Harris and M. Stephens, "A combined corner and edge detector," *Alvey Vision Conference*, pp. 147-151, 1988.

[7] K. Mikolaczyk, C. Scmid, "Scale and affine invariant interest point detectors," *International Journal on Computer Vision*, vol. 60, no. 1, pp. 63-86, 2004.

[8] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Journal of Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 234-253, May 2001.

[9] J. Fauqueur, N. G. Kingsbury and R. Anderson, "Multiscale keypoint detection using the dual-tree complex wavelet transform," *IEEE International Conference on Image Processing*, Atlanta, Georgia, USA, Oct. 2006.

[10] C. Chaux, L. Duval and J.-C. Pesquet, "Image analysis using a dual-tree M-band wavelet transform," *IEEE Trans. on Image Processing*, vol. 15, no. 8, pp. 2397-2412, Aug. 2006.

[11] I. W. Selesnick, "The Double-density dual-tree," *IEEE Trans. on Signal Processing*, vol. 52, no. 5, pp. 1304-1314, May 2004.

[12] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. and Machine Intel.*, vol. 11, no. 7, pp. 674-693, Jul. 1989.

[13] E. P. G. Box, D. R. Cox "Analysis of transformations," *Journal of Royal Statisticl Society*, pp. 211-246, 1964.

[14] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. on Image Processing,* vol. 11, no. 2, pp. 146-158, Feb. 2002.

[15] W. J. Dixon, "Analysis of extreme values," *The Annals of Mathematical Statistics*, vol. 21, no. 4, pp. 488-506, 1950.

[16] E. Loupias, N. Sebe, S. Bres and J. M. Jolion, "Wavelet-based salient points for image retrieval," *IEEE International Conference on Image Processing*, Vancouver, Canada, Sept. 2000.

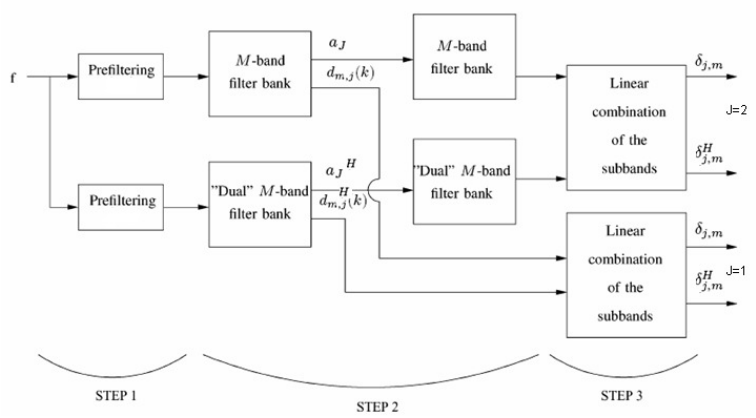
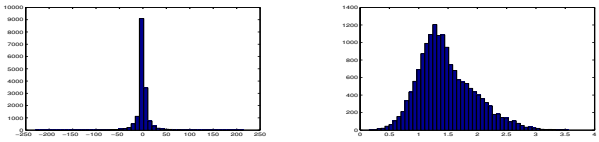
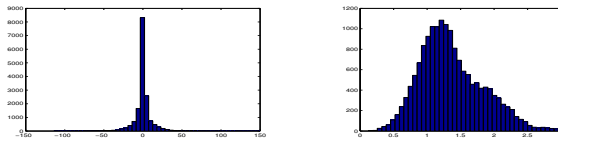


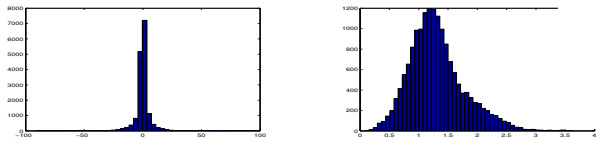
Figure 1: M-band dual-tree decomposition scheme over two resolution levels ($J = 2$)



Horizontal subband $(\widehat{\mu},\widehat{\alpha},\widehat{\beta}) \simeq (0.05, 0.52, 0.87)$



Vertical subband $(\widehat{\mu},\widehat{\alpha},\widehat{\beta}) \simeq (0.14, 0.52, 0.47)$



Diagonal subband $(\widehat{\mu},\widehat{\alpha},\widehat{\beta}) \simeq (0.007, 0.67, 0.51)$

Figure 2: Histograms of wavelet coefficients (left) and transformed wavelet coefficients (right) of the "cameraman" image for $M = 2$ and $J = 1$.

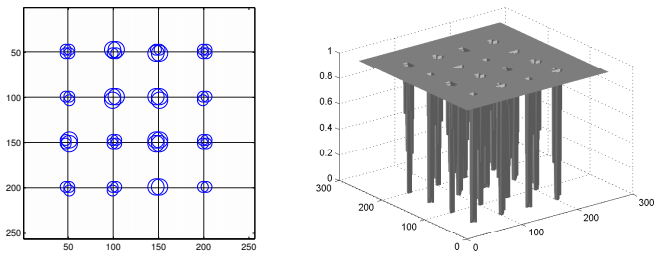


Figure 3: An illustration of SPs detection by using Gru test for $M = 4$, $J = 2$ and $a = 0.1$.(left) Our SPs localiza on the image. (right) The global $p - value$ map of the im

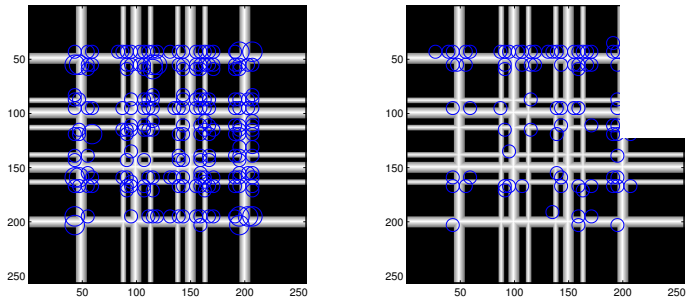


Figure 4: SPs detection with our approach applied on test image. Using Grubbs test $a = 0.05$ (left). Using Dixon test $a = 0.05$ (right).

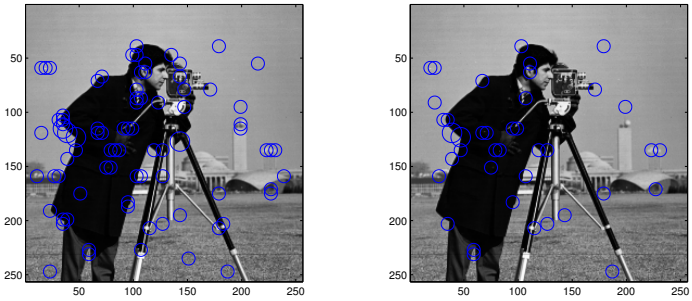


Figure 5: SPs detection with our approach applied on cameraman image. Using Grubbs test $a = 0.01$ (left). Using Dixon test $a = 0.01$ (right).

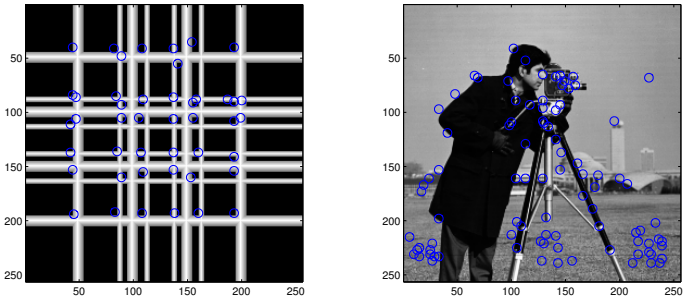


Figure 6: SPs detection with the detector based the accumulated map of keypoint energies applied on test image (left) and cameraman image (right).