

An Algorithm to Estimate Scale Weights of Complex Wavelets for Effective Feature Extraction in Aerial Images

Shubha Bhat*, Ramesh Babu D.R., Krishnan Rangarajan, and Ramakrishna K.A.

Dayananda Sagar College of Engineering, Bengaluru-560 078, India

*E-mail: shubha22@gmail.com

ABSTRACT

Research on vision-based navigation of unmanned air vehicles (UAVs) has been in focus in the recent years, both in military and civilian application domains. Vision-based navigation uses features as cues to match successive images and compute the location of the UAV. The method proposed in this paper is an extension to the existing work to automatically compute scale weights of dual-tree complex wavelet transform (DTCWT) coefficients to improve the performance in terms of detection of features as against discrete wavelets or its Fourier counterparts. Existing DTCWT coefficients give improved time-shifted sensitivity, better directional selectivity with local phase information and limited redundancy. However, the drawback of this method is that the results depend on appropriate selection of scale weights α and β , which are different for different scenes. The existing technique incorporates a rule of thumb-based recommended values of α and β , irrespective of the scene which generates too many key-points, making the registration process computationally intensive. As a real-time application, the need of the problem is to extract less number of strong features. Thus, an algorithmic approach to compute optimal values of scale weights considering precision and accuracy is proposed. The method is tested on various synthetic, simulated, and aerial images having different transformations and in the presence of noise between successive images. It is observed that DTCWT descriptor performs the best and gives better results than scale invariant feature transform and Speeded up robust features.

Keywords: Dual-tree complex wavelet transform, feature detection, scale weights, precision-factor, recall-factor, F-measure

1. INTRODUCTION

An unmanned air vehicle¹ is used both for military and civilian applications. The mission of the UAV considered here is to navigate autonomously from an initial position to a final position using known way points² or landmarks on the terrain using camera as the sensor and precisely estimate its position and orientation during its flight. The cameras capture the imagery of the underlying terrain containing the features and these features act as cues for navigation from one point to the other with an initial known position of the vehicle as the reference. The location of the UAV is computed using the relative location of the identical features in successive frames. In this manner with camera as the only sensor, it is possible to estimate the vehicle's position and orientation even in the absence of information from other sensors, and hence, this type of navigation is called feature-based navigation¹.

To accomplish this, registration of image frames is carried out which involves detection, extraction, and matching of features between successive frames. The most popular ones are the Harris Corner Detector³, and recently, scale-invariant feature transform (SIFT⁴), speeded up robust features (SURF⁵) and its variants.

Lowe⁴ in 1999 proposed a revolutionary technique namely, SIFT descriptor, that is invariant to image translations and rotations, to scale changes (blur), and robust-to-illumination

changes. It is also robust-to-orientation changes of the viewpoint up to 60 degree. Based on the Scale Space theory, the SIFT procedure simulates all Gaussian blurs and normalises local patches around scale-covariant image key points that are Laplacian extremes. But SIFT is computationally-intensive, and is not completely invariant to image rotation. Also, experimental evaluation shows SIFT seems to be more robust for target recognition, however, it presents large number of false positives over all its matches, thus reducing the accuracy.

Bay⁵, *et al.* proposed SURF descriptor, which is a fast performing scale and partially rotation-invariant interest point descriptor. The important speed gain is due to the integration of images, which drastically reduce the number of operations for small box convolutions, independent of the chosen scale. Even without any dedicated optimisation, an almost real-time computation without loss in performance is possible.

The type of feature detector however depends on the type of feature one needs to detect based on the application. Features are categorised as point features (features such as corners of geometrical structures or localised points of interest) and geometric features (features such as edges, straight lines, curves, profile or contours of structures). In the current work, point features particularly corners, junctions and blobs, are used. These features, are more prominent in an urban terrain and these also exhibit the properties needed by the application

like invariance to rotation, translation, scale and perspective changes. The current work is an extension to the framework⁶ wherein the scale weights are automatically chosen to eventually enhance the detection and matching steps.

The wavelet theory, in general, provides a multi-scaling model that allows in choosing the appropriate scale in accordance with the scale-space representation. The complex variant of the wavelet transform, the DTCWT⁷ considered here, has several advantages over discrete wavelet transforms, as explained further in Section 2. The Gabor filters (constructions similar to wavelets), on the other hand, have been used in certain applications to obtain scale and rotation invariance. These are robust to shifts, however these are non-decimated due to which they are time consuming.

2. DUAL-TREE COMPLEX WAVELET TRANSFORM-BASED SCALE-SPACE REPRESENTATION

Dual-Tree Complex Wavelet Transforms⁷ (DTCWT) use a complex-valued filtering approach that decomposes the complex signals into real and imaginary parts in the transformed domain. The real and imaginary coefficients are used to compute amplitude and phase information that is needed to accurately describe the energy of oscillating functions. The DTCWT is better over than Discrete Wavelet Transforms (DWT) as it offers directional selectivity (six orientation subbands) and improved time-shifted sensitivity.

Feature extraction using DTCWT is based on decimated dyadic decomposition (image down-sampling with a factor of 2 at each level) of the given image as in Fig. 1(a), at several scales and six sub-bands, as in Fig. 1(b) for each scale to obtain complex wavelet energy coefficients. These energy coefficients are accumulated to obtain the so-called ‘accumulated energy map’, the peak of which is the ‘keypoint’ location, as shown in Fig. 1(c). The prominence of the keypoint is determined using the strength of the feature. A sharp peak determine a salient feature with high energy content and a wide peak determines a coarser feature with less energy content of the keypoint. In Fig. 1(c), it can be observed that there are few sharp peaks indicating their presence at multiple scales, and few wide peaks indicating their presence at coarser scales. However in certain images, only wider peaks may be present which indicate the existence of features at coarser scales requiring a precise scale.

For example in Fig. 1(a), the aerial image contains a terrain with different geometrical features. The peak in the surface plot of Fig. 1(d) indicates a sharp feature in the image. This feature is the right intersection formed by the ‘H’ shaped feature in the image. However for other features, a wide peak is observed, indicating their presence at coarse scales.

Julien⁷, *et al.* proposed an energy model using DTCWT which is expressed as a product of the magnitude of wavelet coefficients $\rho_1, \rho_2, \rho_3, \rho_4, \rho_5$, and ρ_6 at all six sub-bands pertaining to six orientations ($15^\circ, 45^\circ, 75^\circ, 105^\circ, 135^\circ$, and 165°) as shown in Fig. 1(b). It is measured at each scale as per Eqn. (1).

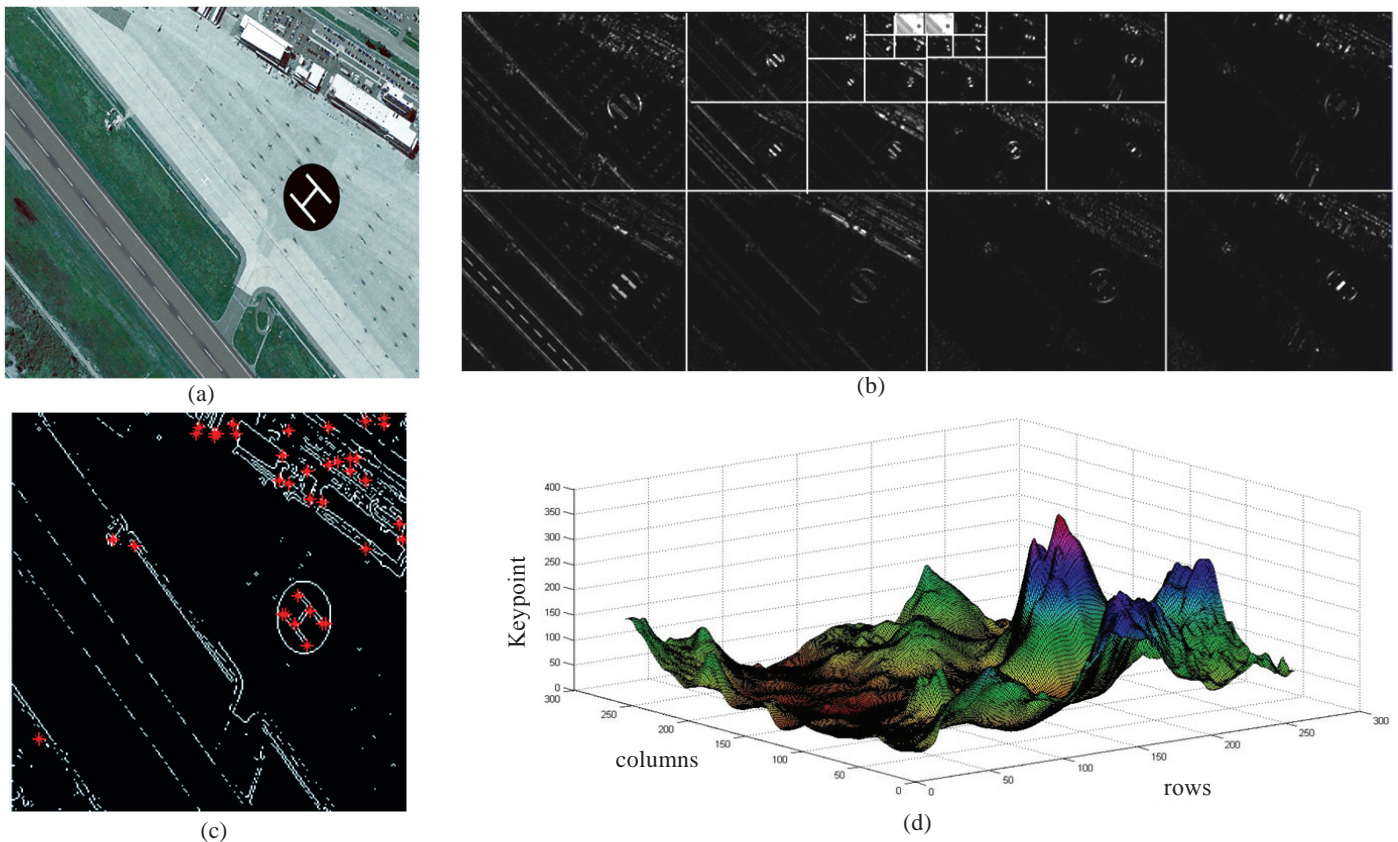


Figure 1. (a) A stable, jitter-free image captured from a bottom looking camera of the UAV; (b) Equivalent of the image in the dual-tree complex wavelet domain. The figure shows four scales of decomposition and six sub-bands for six orientations in each scale; (c) Keypoints detected; and (d) Surface plot of keypoints detected as energy coefficients on image.

$$E(C) = \alpha^s \left(\prod_1^6 \rho_k \right)^\beta \tag{1}$$

In the Eqn. (1), $E(C)$ represent the wavelet energy coefficients, α and β are the scale weights, ρ_k is the magnitude of the wavelet coefficients, and s is the scale parameter ranging between $1 \dots n$, with typical $n=5$. The parameters α and β are dimensionless quantities and indicate fine and coarse tuning parameters, respectively. The parameter α is the tunable gain to the scale responses. This parameter establishes those feature point responses that have improved time-shifted sensitivity. While the parameter β , is the power scaling parameter. This determines where the features exist. These parameters control the scale, and ultimately the accuracy of the detection step which is very crucial in matching features in successive images.

In the research by Fauqueur, Julien⁷, *et al.* the values for α and β have been fixed. However it is observed that α and β need to be changed from image-to-image for better detection and localisation of features, which is time-consuming and a tedious process. It is observed from experiments performed on various image sets that α and β are well-defined in a certain confidence interval, viz., $[0.1, 0.2]$ for β with a step size of 0.01 and $[0, 1]$ for α with a step size of 0.1. This confidence interval is established from experiments conducted on different types of real and synthetic images. The features detected outside this interval indicate image borders, degenerated image structures, etc., and hence, need to be discarded. This optimum range for α and β is determined based on the corresponding number of keypoints detected in this interval, as illustrated in Figs. 2(a) and 2(b) for the sample aerial image in Fig. 1(a).

From the plot in Fig. 2(a), it can be observed that for a noise-free image, no keypoints are detected for $\beta < 0.1$ while too many keypoints are detected for $\beta > 0.2$, and thereafter, the number of keypoints increase exponentially. Heuristically the number of keypoints is restricted to the order of 1000s. In the presence of noise, [Fig. 2(b)], the number of keypoints detected in the range between $[0.1 \ 0.2]$ are 100 and still sufficient to carry out the correspondence step. However it is noteworthy to observe that in both the plots, the number of keypoints detected, include both the true positives and the false positives. To ensure maximum number of true positives, the authors have proposed a method to choose the appropriate α and β so as to obtain a high precision-factor on the detected keypoints. It is also ensured that a minimum of six true positives are sufficient to estimate the transformation between the successive images⁸. Similar experiments were also carried to determine the optimum range for α .

It is observed from the experiments that β provides an approximate region over which the features exist. It needs to be chosen appropriately for every image depending on various factors, eg, the content of the image, noise levels, etc. In certain images, rich features may not exist, therefore the values of β should be selected accordingly to detect optimum number of features, and thereby match with the successive images. Once the value of β is set, the fine-tunable gain α is tuned so that the feature responses which have improved time-shifted sensitivity, remain while the others are eliminated, reducing the number of redundant keypoints. The proposed method automatically

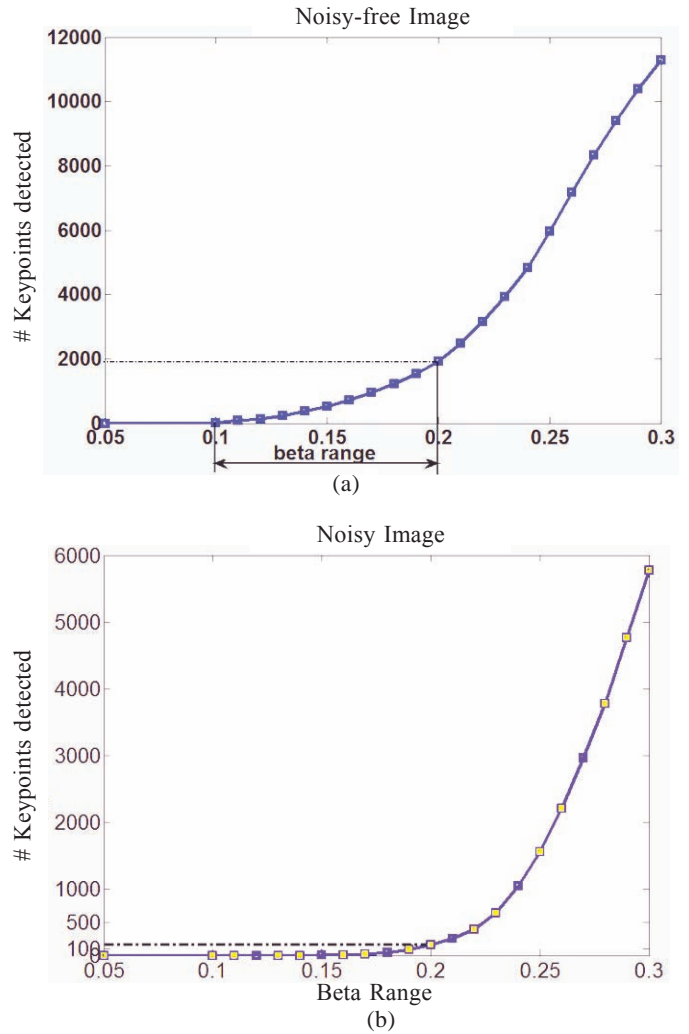


Figure 2. (a) Plot of beta wrt the no. of keypoints to establish the range of beta for noise-free image (b) Similar plot for a noisy image.

selects α and β based on the precision-factor and recall-factor. This eliminates the need to manually select the scale weights for every image.

3. PROPOSED ALGORITHM TO ESTIMATE SCALE WEIGHT

The proposed algorithm extracts features using appropriate scale weights α and β based on better detection and localisation of features captured through precision and recall factors⁹ for the chosen scale weights α and β .

The precision-factor is a measure of the true positives (tp) detected among the total keypoints detected, as shown in Eqn. (2). This is considered to be an important factor for the application with a minimum of six true positives sufficient to carry out the matching between the successive frames. Ideally, the number of true positives must be equal to the total keypoints detected, which means that all the false positives are rejected successfully.

$$precision - factor(P) = \frac{tp}{tp + fp} \tag{2}$$

The recall-factor on the other hand considers the false negatives (fn) or the missed features, as shown in Eqn. (3).

Ideally, no true positive should be missed. However, it is not essential for the application to determine all true positives or not miss any keypoint; nevertheless, a tolerable recall-factor should always be guaranteed.

$$\text{recall-factor}(R) = \sum \frac{tp}{tp + fn} \quad (3)$$

A perfect precision-factor of 1.0 means that every feature detected is correctly classified, but it doesn't say anything about whether all true positives were detected. Similarly, a perfect recall-factor of 1.0 means all true positives were detected, but doesn't say anything about how many remaining keypoints were classified incorrectly. To combine the effect of both precision-factor and recall-factor, F -measure or balanced F -score is also used, which relates the two with harmonic mean, as given in Eqn. (4). A perfect F -measure score of 1.0 indicates a high precision factor and high recall-factor, which means that all the features are correctly classified. However a high precision factor and a low recall-factor or a low precision factor and a high recall-factor indicates a low F -measure score.

$$F\text{-measure} = 2 * \frac{P * R}{P + R} \quad (4)$$

Here P is the Precision factor and R is the Recall factor.

Algorithm: Automatic scale selection

Input: Image frame of size: ($x * y$)

Output: Optimal scale weights

Method:

Start: For each image,

1. Pre-process the image (remove any noise and jitter caused due to vehicle movement using existing techniques).
2. Equalise the image histogram (adjust the illumination levels of the image using histogram equalisation technique).
3. Transform the image from spatial domain to complex wavelet domain using DTCWT.
4. For each scale $s = 1, 2, \dots, m$ scales (until the coarsest scale has resolution of atleast 7×7), do
 - a. Initialise $\alpha \leftarrow 1$,
 - b. for $\beta = 0.1$ to 0.2 with step size: 0.01 , do
 - i. Compute the energy coefficient for every pixel of the image for all 6 sub-bands.
 - ii. Detect the keypoints based on a threshold on energy coefficients with the preset α .
 - iii. Calculate recall-factor (from Eqn. (3)).
 - iv. Repeat steps i to iii till $\beta=0.2$ is reached.
 - v. Select best β with high recall-factor among all scales.

return best β

- c. Set β to the value obtained from step 4b.
- d. for $\alpha = 0$ to 1 with stepsize: 0.1 , do
 - i. Detect the keypoints based on a threshold on energy coefficients for the selected β .
 - ii. Calculate precision-factor, recall-factor and F -measure for all α (from Eqns. (2) (3) and (4)).
 - iii. Plot precision-factor, recall-factor and F -measure
 - iv. Repeat steps i to iii till $\alpha=1$ is reached.
 - v. The region in the plot at which a high precision-

factor and a tolerable recall-factor is obtained (with true positives detected at least in the order of 10^3 's or above) till the solution point where all the three factors have a common value (intersect) can be selected as the best α .

return best α

e. Return the corresponding scale weights α and β .

5. The keypoints generated for the selected scale weights are stored to match with the keypoints from successive images.
6. REPEAT steps 1-5 for different images.

Stop

4. EXPERIMENTS

The experiments carried out to select the appropriate scale weights for DTCWT descriptors is described. A perfectly vertical downward-looking camera with negligible pitch and roll angles between consecutive images is assumed. The computation of pitch and roll angles demands the need of 3-D information of features on the terrain to map these to the 2-D points on the image. The datasets for the experiments are synthetically generated, simulator generated (classified data with ground truth) and real image sequences^{10,11} with different affine transformations (translation, rotation, and scaling).

Experiments were performed with different values of α and β at every scale to measure the location of the detected keypoints. It is obvious that the energy of each pixel at any given scale is the inner product of all the sub-bands raised to the power of a factor β to detect and localise the features of interest. Experiments are conducted with varying values of α to detect features of interest, with a fixed value to β obtained from the previous step. Precision-factor, recall-factor and F -measure scores are calculated at each run for every scale factor. Those values of α and β are selected whose precision-factor, recall-factor, and F -measure scores lie in an optimum range as shown in Fig. 3(b). In Fig. 3(a), α is fixed to 1 and β is varied until all the true positives are detected or until the recall-factor of 1 is reached. It is observed that $\beta = 0.13$ is the optimum value obtained for recall-factor of 1. With this fixed value of β , α is varied and the precision-factor, recall-factor, and F -measure score is calculated.

A plot between precision-factor, recall-factor, and F -measure score against α is shown in Fig. 3(b). The region in the plot where a high precision-factor and a tolerable recall-factor (with true positives detected at least in the order of 10^3 's or above) until the point where all the three factors have a common value (intersect) is considered as the best α .

5. RESULTS AND DISCUSSION

The precision-factor, recall-factor, and F -measure scores are calculated and verified using the ground truth from the simulator and hand-labeled points in synthetic images. This is done to verify the correctness of the algorithm. The scale weight selection method is illustrated with a sample noisy synthetic test case. Figure 4(a) is the noiseless image and Gaussian noise with fixed mean and increasing variance or sigma has been added to this image to obtain a noisy image as shown in Fig. 4(b).

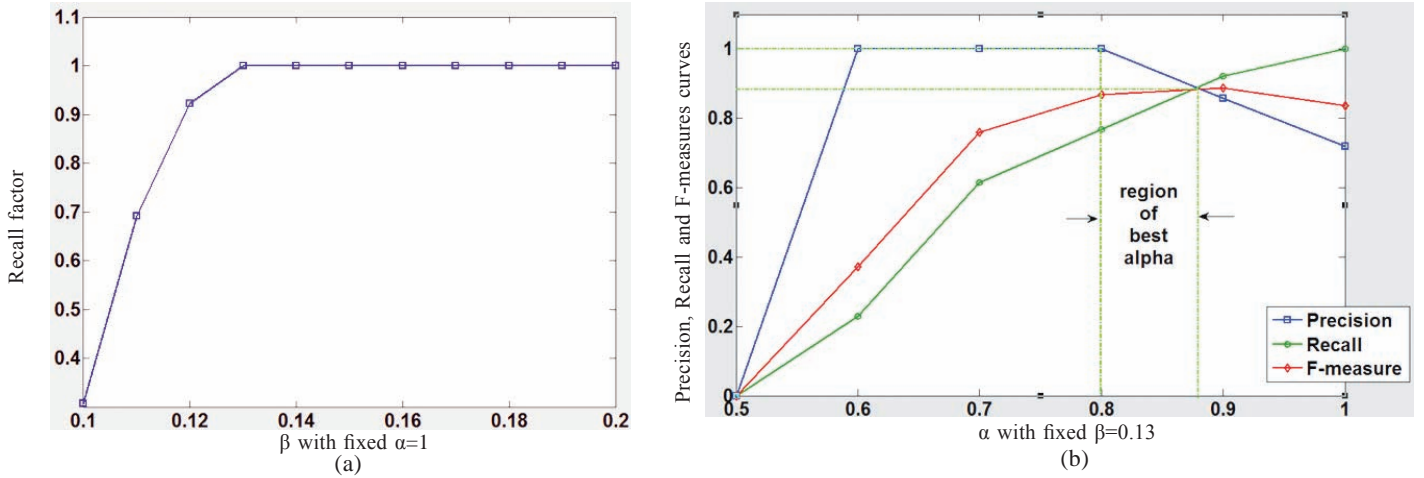
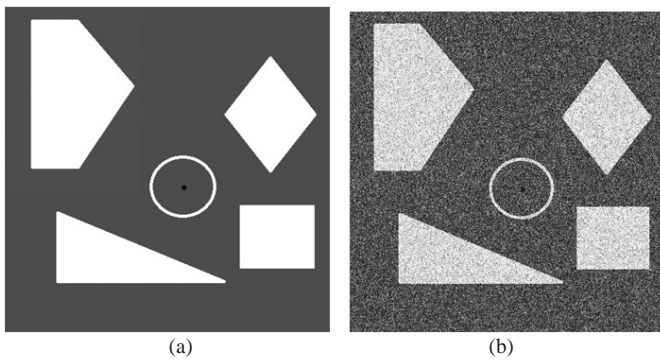


Figure 3. Recall-factor curves for (a) fixed $\alpha=1$, varying β ; (b) Fixed $\beta=0.13$, varying α for a synthetic image.

As explained in Section 4, the influence of α and β on increasing sigma of the Gaussian noise is studied. Similar strategy is applied to obtain the best α and β values by plotting



the precision-factor, recall-factor, and F -measure score. The variance level of the noise is increased from 0.01 to 0.1 (or equivalently sigma from 0.1 to 0.3) at fixed mean.

The intersection of the precision-factor and recall-factor curves in Fig. 4(d) (shown in pink dotted line) gives a low value for precision-factor and recall-factor due to the presence of noise or false positives. However from Fig. 4(c), it is observed that, at a value of $\beta = 0.18$ (for this example), the curve reaches a point beyond which, the number of false positives begin to increase considerably. This value of β gives a better precision-factor and recall-factor than the intersection point (0.19, shown as light green dotted line) and hence β is chosen to be 0.18.

A similar test is performed for varying values of α with fixed $\beta = 0.18$ obtained from the previous step. Figure 4(e)

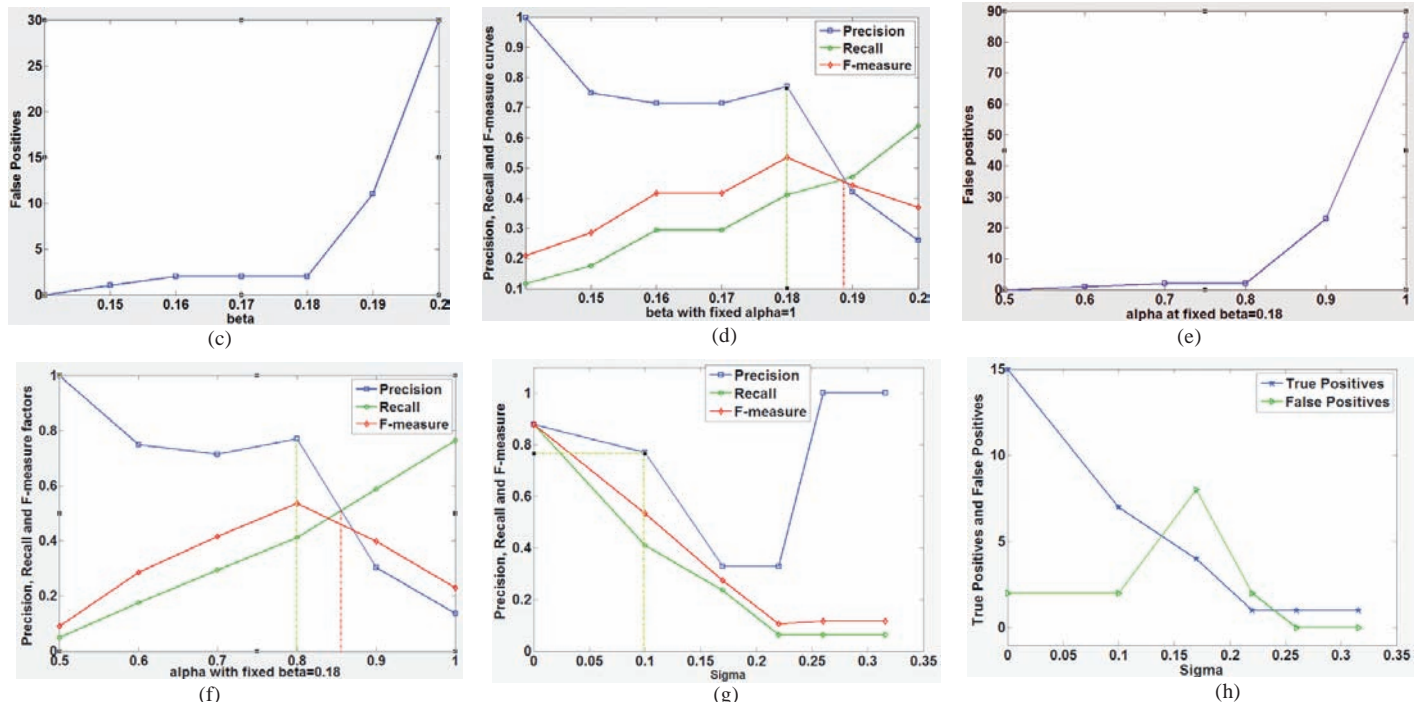


Figure 4. (a) Synthetic image without noise; (b) Noisy image, sigma=0.316; (c) β plotted against false positives at fixed α ; (d) β plotted against precision-factor and recall-factor; (e) α plotted against false positives at fixed β ; (f) α plotted against precision-factor and recall-factor; (g) Effect of noise on precision-factor and recall-factor for various values of sigma; and (h) Effect of noise on true positives and false positives.

indicates the best $\alpha=0.8$ and Fig. 4(f) (light green dotted line) indicates the corresponding best precision-factor and recall-factor with the same justification for α as in the case for β .

The Gaussian sigma level of the noise is increased each time to obtain the precision-factor, recall-factor, and F -measure curves, as shown in Figs. 4(g) and 4(h). The algorithm works well until sigma $\sigma \leq 0.1$ (shown as light green dotted line), beyond this, the precision-factor and recall-factor fall significantly. It can be observed from the plot in Fig. 4(g) that for sigma value = 0.26 and above, the precision-factor reaches the highest value of 1, however the recall-factor or the number of true positives at that particular point, as shown in Fig. 4(h), also needs to be monitored. A very low recall-factor and a high precision-factor indicate that very few true positives have been detected precisely, as shown in Fig. 4(h).

The tradeoff between the precision-factor and the recall-factor should be such that a high precision-factor should always be taken into account primarily while a tolerable compromise with the recall-factor should also be ascertained. In the experiments with different datasets, it has been observed that, a precision-factor > 0.8 and a recall-factor > 0.6 ensure the accuracy and robustness of the detector.

Similar experiments were performed on real images to detect features from a terrain containing dense geometrical structures. Figure 5(a) shows an aerial image containing a terrain with geometrical features. In this case and in general for real images, both the detection and matching step are performed to measure the precision-factor and recall-factor. The true positives, false positives, and false negatives were measured based on the correct matches, miss-matches and missed matches obtained after the matching step. Figures 5(b), 5(c), and 5(d) show the keypoints detected by the proposed, SIFT and SURF methods, respectively.

Performance comparison¹² is carried out for the proposed method against SIFT and SURF descriptors for precision-factor, recall-factor, and F -measure is done for the example aerial image, as shown in Fig. 6. The results, depicted in Fig. 6 reveal that the features have a precision-factor of 86.6 per cent for the proposed method as against 5.4 per cent and 8.3 per cent for SIFT and SURF, respectively. The F -measure is 71.37 per cent for the proposed method as against 10.31 per cent and 14.37 per cent for SIFT and SURF, respectively. The recall-factor of all the three techniques is also more or less close to each other. It is therefore evident that DTCWT-based proposed method has a better precision-factor and comparable recall-factor as against SIFT and SURF. For the application considered here, precision-factor is the most important factor than recall-factor. Recall-factor decides whether all the features are detected or not, and is essential for applications like target recognition¹³. However, a reasonable recall-factor ($>70\%$) should always be ensured which also reflects on the number of true positives detected.

Further experiments on real images (6 databases each with 30 images) were carried out on all the three descriptors. The experimental procedure involved detecting features and generating descriptors, feature descriptor matching, computing transformation model and observing the relative time computation for the three descriptors. Results have been tabulated as shown in Table 1 on the real datasets^{10,11}.

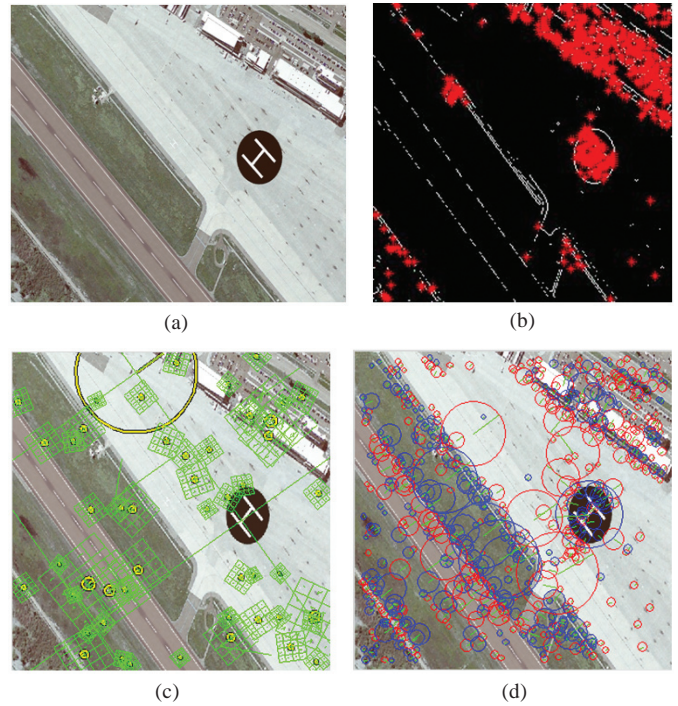


Figure 5. (a) Aerial image; (b) Keypoints detected using the proposed DTCWT-based method; (c) SIFT keypoints; (d) SURF keypoints.

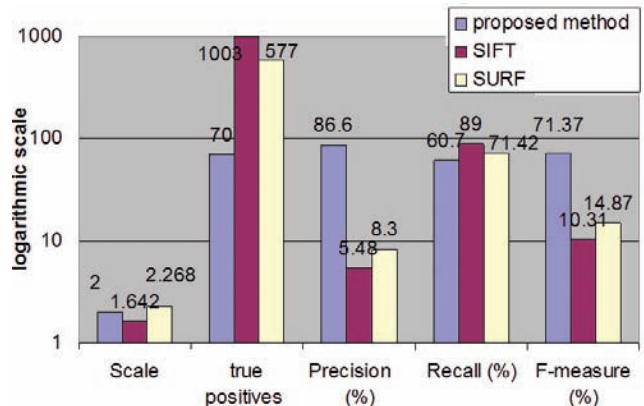


Figure 6. Comparison of proposed DTCWT-based method, SIFT and SURF descriptors.

These datasets contain images taken over urban as well as forest regions. The complexity involved in forest region is that the features are not very prominent as in the case of urban geometrical structures.

It is seen that DTCWT is best suited in terms of accuracy when compared to SIFT or SURF. An average error rate of 7.61 per cent was obtained using the DTCWT descriptor for measuring yaw angle which is within the acceptable limits¹, whereas SIFT and SURF descriptors gave an error rate beyond acceptable range.

However wrt to time computation, DTCWT performed slower than SURF and SIFT as illustrated in Table 1. Nevertheless in case of real-time deployment, a hardware implementation¹⁵ can bridge the performance gaps. Moreover, it is reported¹⁴ that DTCWT is already being used as a descriptor for real-time tracking in aerial images.

Table 1. Comparison of SIFT, SURF, and DTCWT for estimating the transformation model

Dataset and Frame Nos.	Transformation model ground truth (degree)		Computed transformation model (degree)		
	Yaw		DTCWT	SIFT	SURF
(SENSEFLY) 4461 and 4462	5.03°		5.24°	3.44°	-3.5°
(SENSEFLY) 44 60 and 4461	25.62°		25.7°	61.3°	-13.6°
(SENSEFLY) 4456 and 4457	-7.4°		-17.2°	-10.42°	13.3°
(VOLVO) 2509 and 2511	-18.30		-20.70	-20.50	-13.60
(VOLVO) 2491 and 2492	15.30		12.70	17.20	9.520
(VOLVO) 2484 and 2485	10.30		9.160	20.30	5.170
Average error percentage (%) - Accuracy			7.61%	32.32%	72.6%
Average time computation (s)			38.23	34.99	25.33

6. CONCLUSION

In this paper, automated scale weight selection technique for feature detection using DTCWT has been presented. The method is applied on images captured from an unmanned air vehicle navigating at a specified altitude. The scale-space representation presented here is based on the multi-scaling concept of the DTCWT. It is evident from experiments that the choice of appropriate scale weight parameters α and β is a crucial factor in detecting the features to register the images and estimate the position and orientation of the vehicle. The proposed method provides a solution to automatically choose α and β based on the precision-factor and recall-factor.

The proposed method is compared with popular descriptors to measure its performance in terms of precision-factor and recall-factor. It is observed that the proposed method has a higher precision-factor and tolerable recall-factor as compared to SIFT and SURF. The experiments showed that SIFT generated too many keypoints making it computationally intensive and thereby slow. While SURF detected and matched the keypoints faster than SIFT and worked well for real-time images, it was confined for a small range of rotation angles and it was not robust against rotation. DTCWT descriptor worked well for rotation angles between 0° and 90°, though slower than SURF. Though time, performance of the algorithm was recorded, it was not considered as a major factor in comparison as faster implementation can make all the three descriptors satisfy the time performance requirement.

It is further intended to evaluate the use of scale-selection technique with the selected DTCWT descriptor to improve accuracy and time as a part of the future work.

REFERENCES

- Kong, W.Y.; Egan, G.K. & Cornall, T. Feature-based navigation for UAVs. *In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006. doi: 10.1109/iros.2006.281640
- Kehoe, J.J.; Causey, Ryan; Abdulrahim, Mujahid & Lind, Rick. Waypoint navigation for micro air vehicle using vision-based attitude estimation. Technical report, University of Florida. 2004. doi: 10.2514/6.2005-6400
- Harris, C. & Stephens, M. A combined edge and corner detector. *In Alvey Vision Conference*, 1988, pp. 147-151. doi:10.5244/c.2.23
- Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 2004, **60**(2), 91-110. doi:10.1023/B:VISI.0000029664.99615.94
- Bay, Herbert; Tuytelaars, Tinne & Gool, Luc Van. SURF: speeded up robust features. *In Computer Vision ECCV, Computer Vision and Image Understanding*, 2008, **110**(3), 346-359. doi:10.1016/j.cviu.2007.09.014
- Bhat, Shubha; Malagi, Vindhya P.; Babu, Ramesh D.R.; Ramakrishna, K.A. & Ravishankar, M. Scale weight selection for feature extraction using complex wavelets: A framework. *In Proceedings IEEE International Conference on Image Information Processing*, 2011. doi:10.1109/iciip.2011.6108960
- Fauqueur, Julien; Kingsbury, Nick & Anderson, Ryan. Multiscale keypoint detection using the dual-tree complex. *In Proceedings of IEEE Conference on Image Processing*, 2006. doi:10.1109/icip.2006.312656
- Ganapathy, S. Decomposition of transformation matrices for robot vision. *In Proceedings of International Conference on Robotics and Automation*, 1984. pp. 130-139. doi: 10.1109/robot.1984.1087163
- Fawcett, Tom. An introduction to ROC analysis. *Pattern Recog. Lett.*, 2006, **27**, 861-874. doi:10.1016/j.patrec.2005.10.010
- Sensefly database: <http://www.sensefly.com/support/download.html>
- VOLVO database: ftp://aerialrobotics.eu/2012-0430%20VOLVO/_proc/
- Mikolajczyk, K. & Schmid, C. A performance evaluation of local descriptors. *In Proceedings of 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03)*. 2003. doi: 10.1109/cvpr.2003.1211478
- Khotanzad, A. & Hong, Y.H. Rotation invariant image recognition using features selected via systematic method. *Pattern Recognit.*, on 1990, **23**(10), 1089-1101. doi:10.1016/0031-3203(90)90005-6
- Pang, S.K.; Nelson, J.D.B.; Godsill, S.J. & Kingsbury, N. Video Tracking using dual-tree wavelet polar matching and Rao-Blackwellised particle filter. *EURASIP J. Adv. Signal Process.*, 2009, Article ID 620404, 2009, 13. doi:10.1155/2009/620404
- Ghantous, Milad & Bayoumi, Magdy. MIRF: A multi-modal image registration and fusion module based on DTCWT. *J. Signal Process. Sys.*, 2013, **71**(1), 41-55. doi:10.1007/s11265-012-0679-1

CONTRIBUTORS



Ms Shubha Bhat received her B.E.(Elect. and Electronics Engg.) and M.S.(Comp. Sci. Engg.) in from Mangalore University, India and Technical University of Eindhoven, Netherlands, in 2000 and 2006, respectively. She is currently pursuing her PhD in Computer Vision and Image Processing from Visvesvaraya Technological University, Belgaum, India. She is currently working as an assistant professor at Dayananda Sagar college of Engineering, Bangalore. Her current research interests include : Computer vision, image processing and pattern recognition.



Dr Ramesh Babu D.R. received his B.E (Instrumentation Technology) and M.tech. (Industrial Technology) from University of Mysore, India in 1996 and 1999, respectively. He received his PhD in computer science – image processing from University of Mysore, India in 2004. He is currently the Head and Professor of Computer Science and Engineering Department at Dayananda Sagar College of Engineering, Bangalore, India. His current research interests include : Medical image processing, computer vision and pattern recognition.



Dr Krishnan Rangarajan received his BE (hons) (Mechanical engineering) and MTech (Comp. Sci. Engg.) from REC Tiruchirapalli, India and IIT Delhi, India in 1983 and 1985, respectively. He received his PhD in computer science from University of Central Florida, Orlando, USA in 1990. He is currently a Professor in Computer Science and Engineering Department at Dayananda Sagar College of Engineering, Bangalore, India. His current research interests include : Object tracking in the field of computer vision and pattern recognition.



Prof. K.A. Ramakrishna received his MS (Comp. Sci. Engg.) from IIT Madras, India and MSc (Mathematics) from Bangalore University, India, in 1984 and 1970, respectively. He worked as Scientist ‘G’ Group director at Aeronautical development Establishment, Bangalore. He is currently a Professor in Computer Science and Engineering Department at Dayananda Sagar College of Engineering, Bangalore, India. His current research interests include : Computer vision and pattern recognition.