*Journal of Signal and Information Processing*, **2011, 2, 105-111**
doi:10.4236/jsip.2011.22014 Published Online May 2011 (http://www.SciRP.org/journal/jsip)

# Tracking of Non-Rigid Object in Complex Wavelet Domain

**Om Prakash[1,2], Ashish Khare[1]**

[1]Department of Electronics & Communication, University of Allahabad, Allahabad, India; [2]Centre of Computer Education, University of Allahabad, Allahabad, India.
Email: au.omprakash@gmail.com, ashishkhare@hotmail.com

Received December 13th, 2010; revised March 18th, 2011; accepted April 25th, 2011.

## ABSTRACT

*In this paper we have proposed an object tracking method using Dual Tree Complex Wavelet Transform (DTCxWT). The proposed method is capable of tracking the moving object in video sequences. The object is assumed to be deformable under limit i.e. it may change its shape from one frame to another. The basic idea in the proposed method is to decompose the image into two components: a two dimensional motion and a two dimensional shape change. The motion component is factored out while the shape is explicitly represented by storing a sequence of two dimensional models. Each model corresponds to each image frame. The proposed method performs well when the change in the shape in the consecutive frames is small however the 2-D motion in consecutive frames may be large. The proposed algorithm is capable of handling the partial as well as full occlusion of the object.*

**Keywords**: *Object Tracking, Dual Tree Complex Wavelet Transform, Model Based Tracking, Biorthogonal Filters*

## 1. Introduction

In computer vision [1] applications, one of the challenging problems is the tracking of objects in video, in cluttered environment [2,3]. Object tracking and its applications are found in many diverse areas including medical imaging [4,5], recognition and interpretation of objects in a scene [6], video surveillance, target detection and identification [4] etc.

Various spatial and frequency domain techniques are used for these applications, but recent trend is to use wavelet transforms for such applications [7]. However, the form of wavelet used in these techniques has some major drawbacks [8]. To overcome those drawbacks some new form of wavelet transforms like complex wavelet transform [9-14] has been used, which give better results. The major limitations of Discrete Wavelet Transform (DWT) are its shift sensitivity, poor directionality and absence of phase information. These limitations are removed or at least reduced by using complex wavelet Transforms. Dual tree complex wavelet transform (DTCxWT) is one of the commonly used complex wavelet transforms. Although Dual Tree Complex Wavelet Transform (DTCxWT) suffers from the high computational cost due to the redundancy of coefficients but is free from shift variance [15] and directional selectivity [11] problems which are useful in segmentation and tracking of objects in different scenes. Many researchers have used DTCxWT for tracking of rigid objects. Although in some cases Daubechies Complex Wavelet Transform has also been used [13]. Perfect reconstruction is one of the desirable properties for construction of filters. Kingsbury [16,17] has introduced a complex wavelet transform called DTCxWT which gives Perfect Reconstruction and other properties like shift-invariance and directional selectivity.

Rest of the paper is organized as follows: Section 2 of this paper gives overview of complex wavelets and construction of dual tree complex wavelet coefficients, Section 3 describes the proposed method and its implementation while the Section 4 explains the results which are tested and reported for different videos. At last, the Section 5 of this paper describes the conclusions and future scope.

## 2. Complex Wavelet Transform

Although different researchers have shown the magical results after using real valued wavelet transforms but real valued wavelet transforms suffer from the serious disadvantages like shift sensitivity and poor directionality. One of the solutions to overcome these shortcomings is the use of Complex Wavelet Transform.

Kingsbury introduced a more computationally efficient approach to shift invariance [15], the Dual-Tree Complex Wavelet Transform (DTCxWT). Furthermore the DTCxWT gives better directional selectivity when filtering multi-dimensional signals. In summary, it has the following properties:

- Approximate shift invariance;
- Good directional selectivity in 2-dimensions (2-D) with Gabor-like filters (also true for higher dimensionality, *m*-D);
- Perfect reconstruction (PR) using short linear-phase filters;
- Limited redundancy, independent of the number of scales, 2:1 for 1-D (2*m*:1 for *m*-D);
- Efficient order-*N* computation – only twice the simple DWT for 1-D (2*m* times for *m*-D).

Kingsbury [11] observed that the approximate shift invariance with a *real* DWT can also be achieved by doubling the sampling rate at each level of the tree. For this to work, the samples must be evenly spaced. One way to double all the sampling rates in a conventional wavelet tree, such as Tree *a* of **Figure 1**, is to eliminate the down-sampling by 2 after the level 1 filters, $H_{0a}$ and $H_{1a}$. This is equivalent to having two parallel fully- decimated trees, *a* and *b* in **Figure 1**, provided that the delays of filters $H_{0b}$ and $H_{1b}$ are one sample offset from the delays of $H_{0a}$ and $H_{1a}$, which ensures that the level 1 down samplers in tree *b* pick the opposite samples to those in tree *a*. Then it is found that to get uniform intervals between samples from the two trees *below* level 1, the filters in one tree must provide delays that are *half a sample* different (at each filter's input rate) from those in the opposite tree. For linear phase filters, this requires *odd-length* filters in one tree and *even-length* filters in the other. Greater symmetry between the two trees occurs if each tree uses odd and even filters alternately from level to level, it has also been shown that the positions of the wavelet basis functions when the filters are arranged to be odd and even as in **Figure 1**.

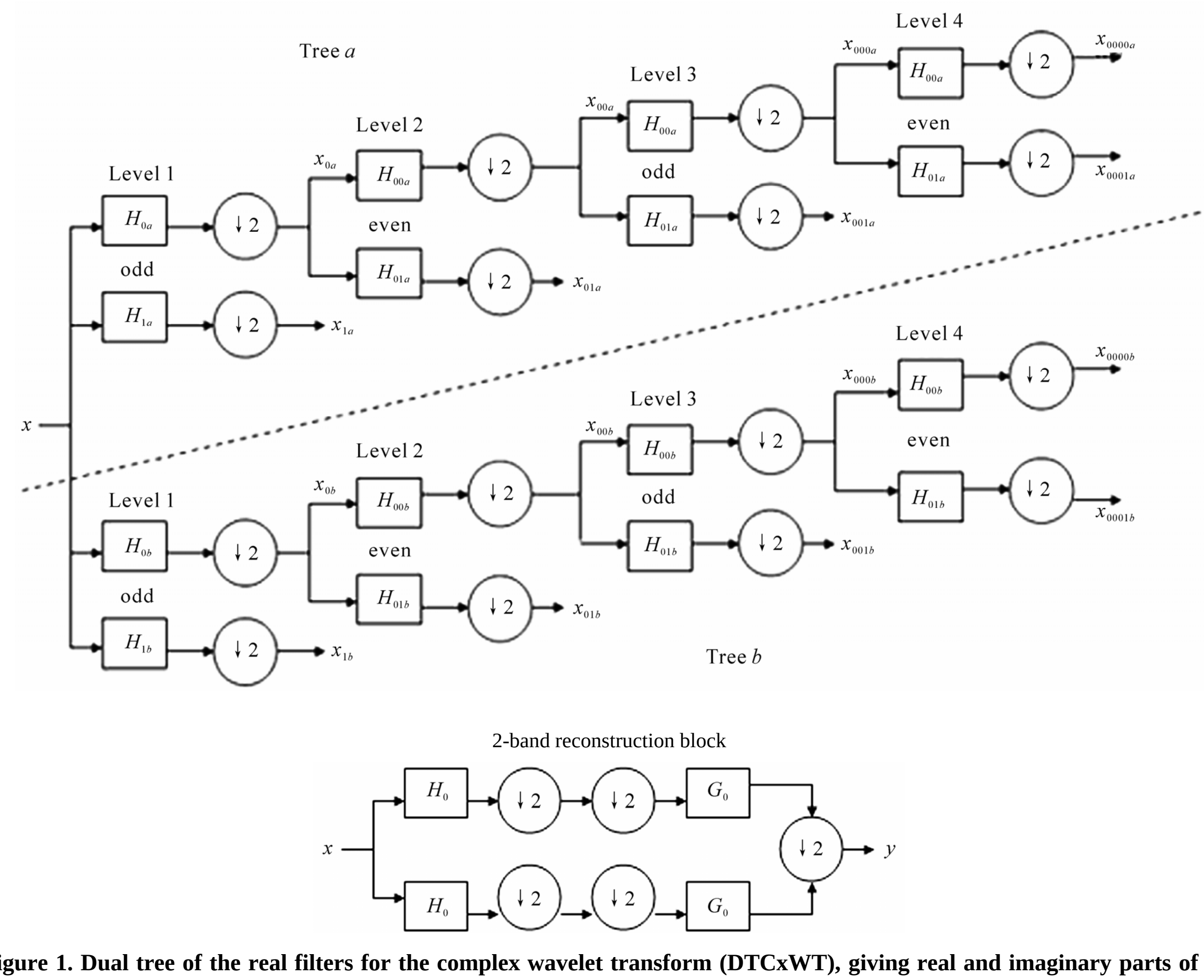



**Figure 1. Dual tree of the real filters for the complex wavelet transform (DTCxWT), giving real and imaginary parts of the complex coefficients.**

To invert the DTCxWT, each tree in **Figure 1** is inverted separately using biorthogonal filters $G$, designed for perfect reconstruction with the corresponding analysis filters $H$ in the 2-band reconstruction block, shown in **Figure 1**. Finally the two tree outputs are averaged in order to obtain an approximately shift invariant system. This system is a wavelet *frame* with redundancy two; and if the filters are designed such that the analysis and reconstruction filters have very similar frequency responses, then it is an almost tight frame, which means that energy is approximately preserved when signals are transformed into the DTCxWT domain. The basis functions were obtained by injecting unit pulses separately into the inverse DTCxWT at each scale in turn. The real and imaginary parts were obtained by injecting the unit pulses into trees $a$ and $b$ in turn.

## 3. The Proposed Object Tracking Approach

In the proposed method an object represented as a sequence of binary images, are corresponding to the each frame of the input image sequence. Each frame of the model corresponds to a set of pixels in a given region of the corresponding image frame. A model $M_t$ is assumed to consist of $m$ points (*i.e.* in the binary representation of $M_t$ there are $m$ non-zero pixels). *i.e.* the image feature at any instant of time consists of the model at that time. These sets of pixels represent the shape of the object at particular frame.

The main ideas of our proposed approach are as below:

1) The 3D moving object is divided into two parts: a 2D shape change and a 2D motion. The change in the shape should be small while there is no restriction on the motion.

2) Change in 2D shape between successive image frames is captured using 2D geometric models.

3) Perform 2D model matching using minimum Hausdorff distance [18] which can be defined, for any two point sets $P$ and $Q$, as

$$H(P,Q) = \max\left(h(P,Q), h(Q,P)\right)$$

where,

$$h(P,Q) = \max_{p\in P} \min_{q\in Q} \left\| p - q \right\|,$$

$\|.\|$ denotes Euclidean distance. This Hausdorff distance simply measures the proximity of points in the two sets which represents the difference between fixed point sets whereas we require the difference between *shapes* of point sets. Therefore, for some transformation group $G$, minimum difference between two shapes w.r.t. the group action is given by

$$D_G(P,Q) = \min_{g\in G} H\left(g(P), Q\right)$$

In other words, the distance between two shapes is minimum difference between them under all possible transformations of one shape w.r.t. the other. $D_G(P,Q)$ satisfies metric properties (identity, symmetry and triangle inequality)

4) If $D_G(P,Q)$ is zero then two shapes are same otherwise change in shape is measured.

5) Huttenlocher, *et al.* [18] has defined the partial distance or Rank order, which is used in tracking the 2-D shape change, as

$$h_K(P,Q) = \underset{p\in P, q\in Q}{K^{th}} \min \left\| p - q \right\|$$

This minimum directed partial Hausdorff distance [18] is used to find the current location of the object. This tracks the 2-D motion of the object. Also, the distance between image and transformed model is used to select those set of pixels which are the part of the next model, gives us change in the 2-D shape of the object.

### 3.1. Algorithm for Tracking of Non-Rigid Objects

1) Initialize the model $M_{t,}$ which is assumed to consist of m points (*i.e.* in the binary representation of $M_t$ there are m non-zero pixels), move it to the next time frame $I_{t+1}$.

2) Find the new model $M_{t+1}$, from $M_t$ and $I_{t+1}$.

3) Locate the object in the new image frame by computing $d$ as

$$\begin{aligned} d &= \min_{g\in G} h_K\left(g(M_t), I_{t+1}\right) \\ &= \min_{g\in G} \underset{p\in M_t, q\in I_{t+1}}{K^{th}} \min \left\| g(p) - q \right\| \end{aligned}$$

computed value of $d$ identifies the transformation $g^* \in G$ of $M_t$ which minimizes the rank order *i.e.* it identifies the best "position", $g^*$, of $M_t$ in the image frame $I_{t+1}$.

### 3.2. Finding the Model's New Location

1) Identify all possible locations of the model $M_t$ in the next time frame $I_{t+1}$.

2) Compute the set of transformations of $M_t$, say $X$ such that the partial Hausdorff distance is not larger than some value $\tau$ *i.e.*

$$X = \left\{ x : h_K\left(M_t \oplus x, I_{t+1}\right) \le \tau \right\}$$

where $\oplus$ is the Minkowski sum notation. The $X$ can be computed as

- Compute distance transform $D_{t+1}$ of $I_{t+1}$ which is an array specifying each location in the image the distance to the nearest non-zero pixel of $I_{t+1}$.
- Compute directed Hausdorff distance, $h_K(M_t \oplus x, I_{t+1})$ for a given translation $x$ of $M_t$.

This distance is the $K^{th}$ largest of the m valued $D_{t+1}$

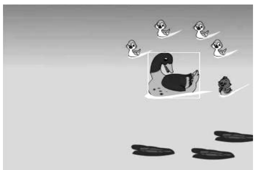
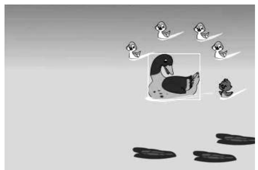
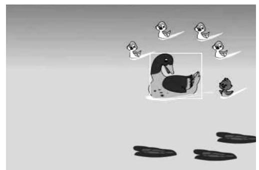
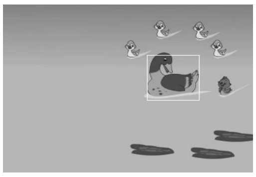
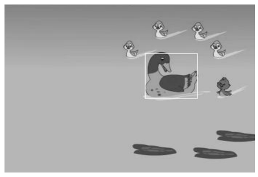
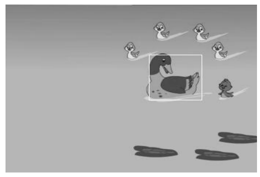
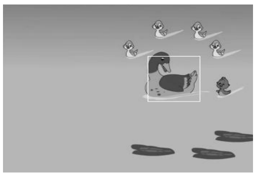
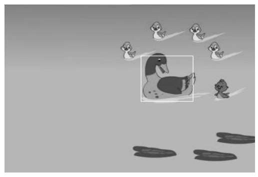
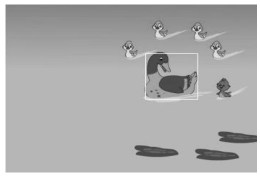
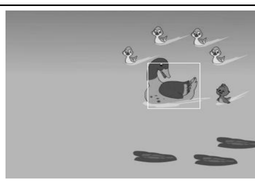
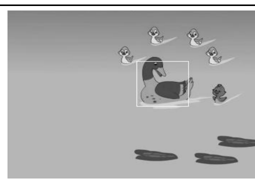
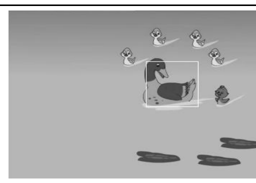
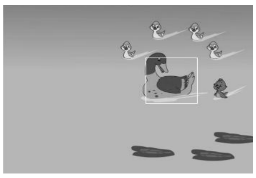
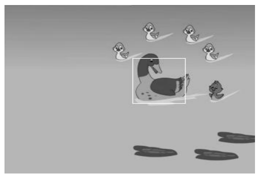
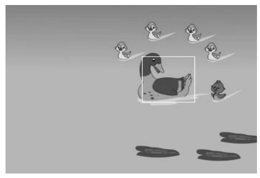
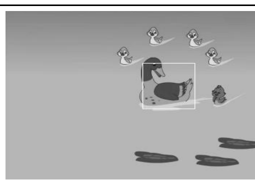
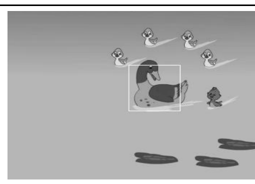
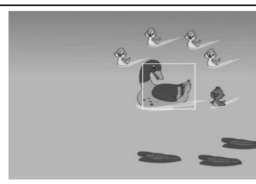
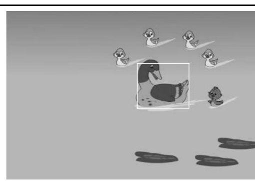
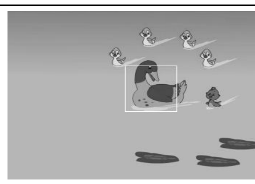
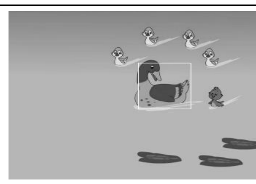
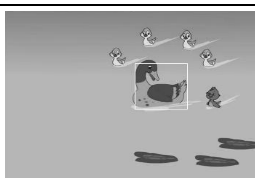
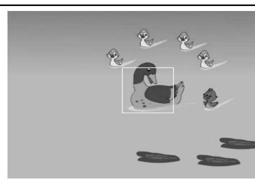
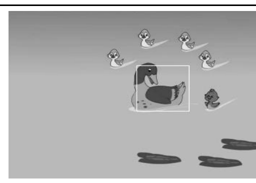
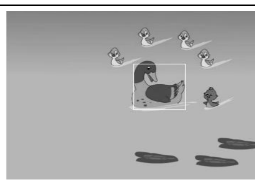
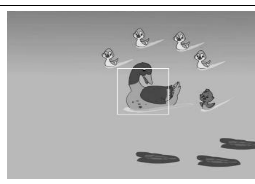
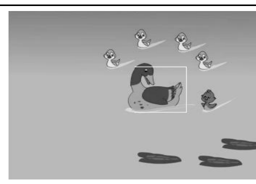
| Frame No. | Spatial domain | Real Wavelet Domain | Complex Wavelet Domain |
|---|---|---|---|
| Frame-1 |  |  |  |
| Frame-5 |  |  |  |
| Frame-10 |  |  |  |
| Frame-15 |  |  |  |
| Frame-20 |  |  |  |
| Frame-25 |  |  |  |
| Frame-30 |  |  |  |
| Frame-35 |  |  |  |
| Frame-40 |  |  |  |

**Figure 2. Tracking for duck video.**

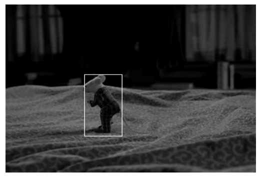
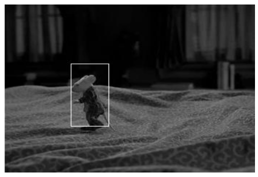
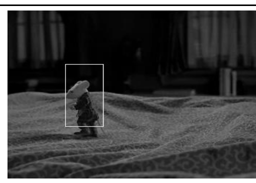
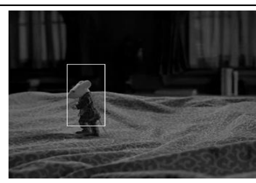
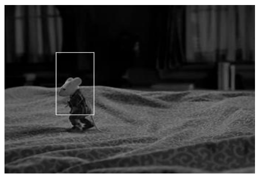
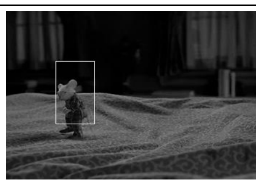
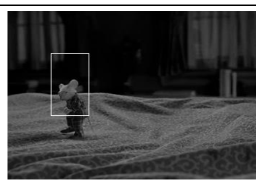
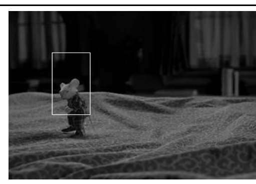
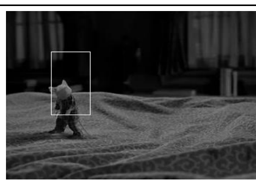
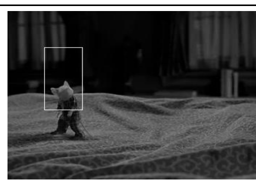
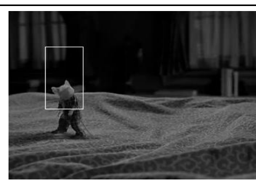
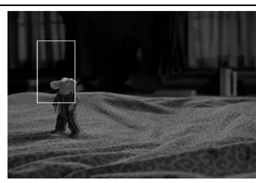
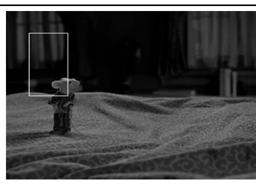
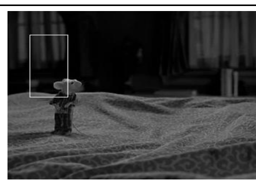
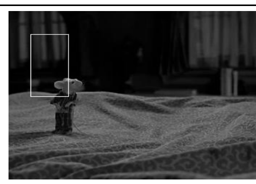
| Frame No. | Spatial domain | Real Wavelet Domain | Complex Wavelet Domain |
|---|---|---|---|
| Frame-1 |  |  |  |
| Frame-5 |  |  |  |
| Frame-10 |  |  |  |
| Frame-15 |  |  |  |
| Frame-20 |  |  |  |
| Frame-25 |  |  |  |
| Frame-30 |  |  |  |
| Frame-35 |  |  |  |
| Frame-40 |  |  |  |

**Figure 3. Tracking for stuart video.**

specified by the non-zero points of $M_t + x$. that is for a given translation $x$, for each non-zero pixel $p$ of $M_t$ find the location $p + x$ of $D_{t+1}$. The $K^{th}$ largest of these m values found so for is the partial directed Hausdorff distance $h_K(M_t + x, I_{t+1})$.

### 3.3. Updating the Model

The new model, $M_{t+1}$ constructed from finding the translation $x$ of $M_t$ with respect to the $I_{t+1}$ by selectively choosing the non zero pixels of $I_{t+1}$ which falls within a distance of $\delta$ of non zero pixels of $M_t \oplus x$. This can be obtained by dialating $M_t$ by a disc radius $\delta$, shifting this by $x$, and then computing the logical AND of $I_{t+1}$ with the dialated and translated model.

## 4. Experimental Results

The initial model, corresponding to the first frame of the image sequence is taken. The user needs to take a rectangle in the first frame containing this initial model. Then image is processed to select the subset of the edge pixels in this rectangle. This is done by assuming that the camera will not move in first two consecutive frames and using this fact to filter the first image frame based on the second image frame. Those edge pixels in the user selected rectangle that moved between the first and the second frame are used as initial model. Thus the only input from the user is the rectangle in the first frame of the image sequence that contains the moving object.

The proposed method of tracking of non-rigid object is applied on various videos and the results are compared by computing the dual tree complex wavelet coefficient to that by real and spatial domain coefficients. It has been observed that:

- In spatial domain, tracking is more accurate but high execution time
- In real wavelet domain, tracking is not as good as in the spatial domain but the execution time is approximately 4 times less
- Tracking in Complex wavelet domain is not only fast as compared to real wavelet domain but also it is comparable to spatial domain.

Also, we observed that in the spatial domain and real wavelet domain the tracking work well up to the 15 frames and thereafter tracker is not able to track the object accurately while as in complex wavelet domain it is more accurate and efficient. Here the results are only mentioned up to the 40 frames. The comparative results for two representative videos are presented: in spatial domain, real wavelet domain and dual tree complex wavelet domain in **Figure 2** and **Figure 3**.

## 5. Conclusions and Future Scope

The experiments are performed on various moving videos and non-rigid objects are tracked in spatial domain, real wavelet domain and in complex wavelet domain. Comparing the tracking results in three domains, we found that the tracking in spatial domain is very accurate because the operations are performed pixel-wise basis but the computational time in spatial domain is high. It has been observed that in real wavelet domain, the computational complexity is low (approximately 4 times less) but results of tracking are very poor. Tracking in complex wavelet domain is fast as well as it produce good results as compared to real wavelet domain. Here in complex wavelet domain, both real as well as imaginary components are taken which provide us better segmentation and tracking as compared to that of real wavelets. The proposed tracking algorithm can be extended for tracking of multiple objects in video sequences.

## 6. Acknowledgements

The authors are thankful to the Department of Science & Technology, New Delhi and the University Grants Commission (UGC), New Delhi, India for providing research grant vide its grant nos. SF/FTP/ETA-023/2009 and 36-246/2008(SR) for research project.

## REFERENCES

[1] A. K. Jain, “Fundamentals of Digital Image Processing,” Prentice Hall of India Pvt. Ltd., New Delhi, 2001.

[2] S. Nigam and A. Khare, “Curvelet Transform Based Object Tracking,” *Proceedings of IEEE International Conference on Computer and Communication Technologies*, Allahabad, 17-19 September 2010, pp. 230-235.

[3] M. Khare, T. Patnaik and A. Khare, “Dual Tree Complex Wavelet Transform Based Video Object Tracking,” *Communications in Computer and Information Science*, Vol. 101, No. 2, 2010, pp. 281-286.

[4] H. Goszczynska, “A Method for Densitometric Analysis of Moving Object Tracking in Medical Images,” *Machine Graphics & Vision International Journal*, Vol. 17, No. 1, 2008.

[5] Z. M. Budimlija, M. Leckpammer, D. Popiolek, F. Fogt, M. Forenderick and R.Bieber, “Forensic Applications of Capture Laser Microdissections: Use in DNA-Based Parentage Testing and Plateform Validation,” *Croatian Medical Journal*, Vol. 46, No. 4, 2005, pp. 549-555.

[6] R. Anderson, N. Kingsbury and J. Fauqueur, “Coarse Level Object Recognition Using Interlevel Products of Complex Wavelets,” *Proceedings of IEEE Conference on Image Processing*, Genoa, September 2005.

[7] N. G. Kingsbury and J. F. A. Magarey, “Wavelet Transforms in Image Processing,” *Proceedings of* 1*st European Conference on Signal Analysis and Prediction*, Prague, 24-27 June 1997, pp. 23-34.

[8] I. W. Selesnick, R. G. Baraniuk and N. G. Kingsbury, “The Dual-Tree Complex Wavelet Transform,” *IEEE*

*Signal Processing Magazine*, Vol. 22, No. 6, 2005, pp. 123-151. doi:10.1109/MSP.2005.1550194

[9] F. C. A. Fernandes, R. L. C. Spaendonck and C. S. Burrus, "A New Framework for Complex Wavelet Transform," *IEEE Transactions on Signal Processing*, Vol. 51, No. 7, 2003, pp. 1825-1837. doi:10.1109/TSP.2003.812841

[10] A. A. Bharath and J. Ng, "A Steerable Complex Wavelet Construction and Its Applications to Image Denoising," *IEEE Transactions on Image Processing*, Vol. 14, No. 7, 2005, pp. 948-959. doi:10.1109/TIP.2005.849295

[11] N. G. Kingsbury, "Shift Invariant Properties of the Dual-Tree Complex Wavelet Transform," *Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing*, Phoenix, Vol. 3, 16-19 March 1999.

[12] N. Kingsbury, "Rotation-Invariant Local Feature Matching with Complex Wavelets," *Proceedings of European Conference on Signal Processing*, Florence, 4-8 September 2006.

[13] A. Khare and U. S. Tiwary, "Daubechies Complex Wavelet Transform Based Moving Object Tracking," *IEEE Symposium on Computational Intelligence in Image and Signal Processing*, Honolulu, 1-5 April 2007, pp. 36-40.

[14] N. G. Kingsbury, "Image Processing with Complex Wavelets," *Philosophical Transactions of Royal Society London A*, Special Issue for the Discussion Meeting on "Wavelets: The Key to Intermittent Information?" Vol. 357, No. 1760, 1999, pp. 2543-2560.

[15] N. G. Kingsbury, "The Dual-Tree Complex Wavelet Transform: A New Technique for Shift Invariance and Directional Filters," *Proceedings of* 8*th IEEE DSP Worksho*p, Bryce Canyon, August 1998.

[16] N. G. Kingsbury, "Complex Wavelets for Shift Invariant Analysis and Filtering of Signals," *Applied and Computational Harmonic Analysis*, Vol. 10, No. 3, 2001, pp. 234-253. doi:10.1006/acha.2000.0343

[17] N. G. Kingsbury, "The Dual-Tree Complex Wavelet Transform: A New Efficient Tool for Image Restoration and Enhancement," *The* 9*th European Signal Processing Conference*, Rhodes, September 1998.

[18] D. P. Huttenlocher, J. J. Noh and W. J. Rucklidge, "Tracking Non-Rigid Objects in Complex Scenes," *Proceedings of* 4*th International Conference on Computer Vision*, Berlin, 11-14 May 1993, pp. 93-101.