# Covid Image Classification using Wavelet Feature Vectors and NN

Stella Vetova
*Department Informatics*
*Technical University of Sofia*
Sofia, Bulgaria
vetova.bas@gmail.com

*Abstract*— In the following paper an algorithm for Covid images classification is presented. The algorithm classifies images into positive and negative group on the base of wavelet feature vectors and neuron network applied. To assess the proposed algorithm, experiments are conducted on the accuracy, precision and recall parameters. The results are analyzed showing the advantages of the algorithm and its appropriate application.

***Keywords—wavelet feature vectors, wavelets, wavelet coefficients, neural network, image decomposition, image classification***

## I. Introduction (*Heading 1*)

Genetic and image tests are the first methods for Covid-19 disease diagnosis which was reported by the World Health Organization a year ago. For the purpose of Covid image processing and classification two main approaches, which show significant results, and are widely used. The first one is image classification performed on the base of image features extraction generating the feature vectors of the images and similarity computation afterwards. The second approach is based on deep learning using artificial neural networks. Image features vectors can be generated on the base of primitive. Discrete Wavelet Transform [1], Gabor Wavelet Transform [2], [3], [4], Dual-Tree Complex Wavelet Transform [5], Contourlet Transform [6], [7], Curvelet Transform [8], 3D Fourier Transform [9] are most often used used for texture extraction and orientation recognition. On the other hand, to generate feature vectors by color widely used are color histogram [10], [11], [12], color moments [13] and correlogram [14]. For the cases of shape recognition typical techniques are: Harris Corner Detector [15], Barnard Detector [16], Contour-based shape features [17], [18], [19], Hough Transform [20]. To compute similarity in the algorithms and systems for image classification distance measures is applied. Such measures are Manhattan distance, Mahalanobis distance, Canberra distance, Euclidean distance, etc.

Another technique for image classification is artificial neural networks (ANN). ANN are designed on the model of natural neural networks and the performance of functions of the human brain. The neural network (NN) is made up of a finite number of neurons connected to each other in a certain order and model to perform a specific task. This approach is used to classify data on the base of extracted feature vectors in advance performing deep learning to learn the neural network for a new classification task [21], [22].

In addition, neural networks are designed to solve a wide range of tasks such as human face detection, user image classification, biological data identification [23], object recognition, disease assessment based on medical images for diagnosis and appropriate treatment. Gancheva and Georgiev [24] propose a system for learning and knowledge extraction including patterns from a collection of input data or past experience. Generally, the system processes knowledge in three phases: features extraction and training the system for identifying specific patterns; detecting and classifying a possible pattern; execution of Machine Learning (ML) algorithm to determine the most appropriate model to represent the behavior or the pattern of the data.

On the other hand, the problem of healthcare data security and privacy remains one of the hottest issues facing medical science. It was reported that in 2016 the increase of hacking attacks achieved 320% [25]. In their quest to find solution of cyberattacks and threads, the scientists and software developers propose strategies and methods for data protection against phishing attacks [26], spam, malicious links [27] and network and software-based attacks [28], such as authentification, encryption, data masking, monitoring and auditing, etc.

The following paper presents an algorithm for Covid image classification using image feature vectors and NN. It is organized as follows. Section 2 indroduces the Dual-Tree Complex Wavelet Transform used for image feature vectors generation. Section 3 presents the proposed algorithm for the case study of Covid image classification and experimental results. Section 4 concludes the paper.

## II. The Dual-Tree Complex Wavelet Transform

### A. *Selecting a Template (Heading 2)*

In 1998 Nick Kingsbury introduced an effective Complex Wavelet Transform (CWT) method called the Dual-Tree Complex Wavelet Transform (DT CWT). It is CWT based on complex valued scaling function and complex-valued wavelet:

$$\Psi_c(t) = \Psi_r(t) + j\Psi_i(t) \qquad (1)$$

where $\Psi_r(t)$ - real and even part, $j\Psi_i(t)$ - imaginary and odd part, $\Psi_c(t)$ - analytic signal;

Kingsbury's idea was to develop a transform which produces analytic signal on the analogy of Fourier transform and which possesses the following properties:

- smooth non-oscillating magnitude;
- nearly shift-invariant magnitude;

- significantly reduced aliasing effect;
- directional wavelets in higher dimensions;

For the DT CWT realization Kingsbury uses two Discrete Wavelet Transforms (DWTs) performed on two different binary wavelet trees A and B (Fig.1) for each. Thus he designs the real and the imaginary part of DT CWT to produce the analytic signal.

Fig. 1 illustrates graphically the 1-D DT CWT analysis filter bank (FB) structure.

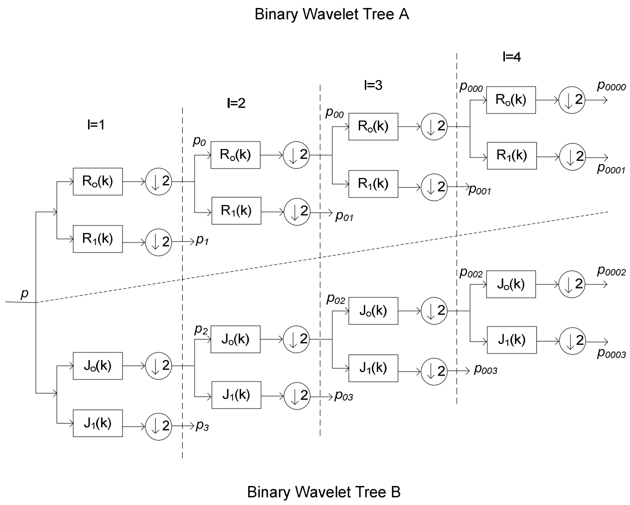


Fig.1. The 1-D DT CWT analysis filter bank (FB) structure

The input signal $p$ is decomposed by the lowpass filters $R_0(k)$ for the real part and $J_0(k)$ for the imaginary one and decimated by 2:1 generating its lowpass components $p_0$ and $p_2$. The highpass filters $R_1(k)$ and $J_1(k)$ and decimation by 2:1 produce the highpass components $p_1$ and $p_3$ of the signal $p$. This process continues as far as required for levels $l=1, 2, 3, 4$. The final result of the decomposition of $p$ is: $p_1$, $p_{01}$, $p_{001}$, $p_{0001}$ for the real part and $p_3$, $p_{03}$, $p_{003}$, $p_{0003}$ for the imaginary part. The filters used for DT CWT are chosen to be linear-phase satisfying the Perfect Reconstruction (PR) condition [14] and are joined so that the final result of the transform is approximately analytic:

$$\Psi(t) := \Psi_R(t) + j\Psi_J(t) \tag{2}$$

where: $\Psi_R(t)$, $\Psi_J(t)$ - wavelets generated by the two DWTs.

In addition, both low-pass filters $R_0(k)$ and $J_0(k)$ have to be designed to possess a property so as the corresponding wavelets to form an approximate Hilbert transform pair:

$$\Psi_J(t) \approx H\{\Psi_R(t)\} \tag{3}$$

where:

$$\Psi_R(t) = \sqrt{2}\sum_k R_1(k)\Phi_R(t) \tag{4}$$

$$\Phi_R(t) = \sqrt{2}\sum_k R_0(k)\Phi_R(t) \tag{5}$$

For this goal one of the two lowpass filters has to be nearly half-sample shift to the other:

$$J_0(k) \approx R_0(k-0.5) \Rightarrow \Psi_J(t) \approx H\{\Psi_R(t)\} \tag{6}$$

This half-sample delay leads to nearly shift-invariant wavelet transform.

Besides one-dimensional application, DT CWT may be used for two dimensional tasks through 2-D DT CWT relying on the M-D dual-tree wavelets properties to be approximately analytic and oriented. Thus, it is suitable for edge and surface detection in image processing. The process of filtering is performed by two different groups of filters providing two 2-D separable DWTs and six subbands: two HL, two LH, and two HH subbands.

2-D DT CWT finds application in image segmentation [29], motion estimation [30], texture analysis and synthesis [31], feature extraction [32], [33].

## III. Proposed Algorithm and Experimental Results

### A. *Image Feature Extraction Using DT CWT*

The proposed wavelet features-based algorithm for biomedical image classification is graphically illustrated in fig. 2. It is divided into two stages performed directly one after the other. The first stage concerns the process of image decomposition and wavelet feature vectors generation and consists of seven phases. In the first one the greyscale test biomedical image database is loaded (Fig.3). In the second one its compound images are preprocessed being resized into 256x256px (Fig.4). On this base, the wavelet coefficient generation is performed using 2D DT CWT [34, 35, 36, 37, 38, 39, 40, 41] at level four (l=4) on the whole image. At this phase the image feature vectors are constructed and at phase four they are stored in a database designed for this purpose. On the other hand, stage 2 concerns the process of user query-image submission and its processing. To achieve accurate result, the phase sequence of stage 1 is needed to be followed. Thus, there is no difference between the first three phases of the two stages. Phase four is needed only for the process of stage 1. The generated wavelet features form the feature vectors of the biomedical images and is submitted to the Neural Network used for training performance. Image classification and class affiliation is computed on the base of the query-image feature vector and the training model. At the end of the process the classification result is displayed.

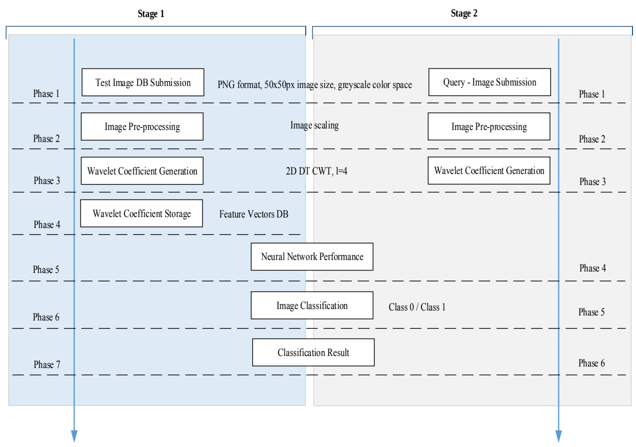


Fig.2. Flowchart of the algorithm for Covid image classification

ISSN: 2692-5079

Fig.3. Source image

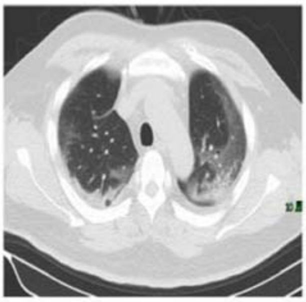

Fig.4. Resized image

## *B. Experimental results*

For the goal of the experiments conducted, Covid test image database is used. It contins 758 Covid images in grayscale color space. The images are classified into two groups: Covid positive as illustrated in Fig. 5 and Covid negative in Fig. 6 with 361 and 397 images respectively. They are distinguished for JPEG and PNG format.

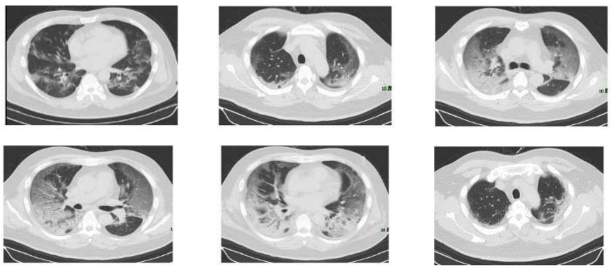

Fig. 5. Covid positive test image database

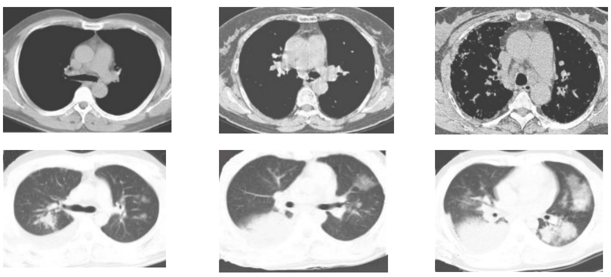

Fig. 6. Covid negative test image database

The software for mathematical and engineering computation Matlab is used to design the wavelet based algorithm and to conduct the research experiments.

In order to evaluate and analyze the performed algorithm for the case study of biomedical image classification, performance metrics are needed to be employed to determine the degree of relevance between the query-images and the classified ones. For the purposes of the presented work the conducted experiments are evaluated using classification accuracy, precision and recall measures defined in Table 1:

TABLE I. EVALUATION METRICS

| **Evaluation Metric** | **Definition** |
|---|---|
| **Accuracy** | $Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)}$ |
| **Precision** | $Precision = \frac{TP}{(TP+FP)}$ |
| **Recall** | $Recall = \frac{TP}{(TP+FN)}$ |

where:

- TP – True Positives – a class of images defining the cases when the actual output fits the prediction parameter and both have YES value;
- TN – True Negatives – a class of images defining the cases when the actual output fits the prediction and both have NO value;
- FP – False Positives – a class of images defining the cases when the predicted observations are evaluated as negatives (NO value) and the actual output is positive (YES value). Thus, there is no fit between both parameters.
  - FN – False Negatives – a class of images defining the cases when the predicted observations evaluated as positives (YES value) and the actual output is negative (NO value) with no fit between both parameters.

Since the accuracy measure is appropriate to be applied for the cases when the images in the test database are equal number of false positives and false negatives for a complete evaluation precision and recall measures are calculated as well. Precision is used to compute the ratio of the number of positive observations to the number of predicted positive observations in total showing the percentage of the relevant result. In addition, recall is the measure used to evaluate the ratio of the number of positive observations to the total number of TP and FN determining the percentage of the actual positives.

The results of the performed five experiments are listed in Table 2. The average value is presented in the cells of the last row of the table. On the base of the results the proposed wavelet-based algorithm with NN classification demonstrates high results for the evaluation metrics accuracy, precision and recall. With regard to the first parameter the algorithm shows 95.2% in determining the affiliation of an image to category 0 or category 1. Furthermore, in terms of the precision the result exceeds 93%. In addition, the recall result reaches 87.8% which leads to the conclusion that the use of NN is appropriate for the cases when a false negative is high.

TABLE II. RESULTS OF THE PROPOSED WAVELET-BASED ALGORITHM

| Experiment | Accuracy (%) | Precision (%) | Recall (%) |
|---|---|---|---|
| Exp. 1 | 95% | 84% | 86% |
| Exp. 2 | 93% | 96% | 92% |
| Exp. 3 | 98% | 99% | 81% |
| Exp. 4 | 96% | 100% | 87% |
| Exp. 5 | 94% | 87% | 93% |
| Total | 95,2% | 93.2% | 87.8% |

Figures 7, 8 and 9 graphically present the obtained results for experiments 1, 2 and 3 for wavelet feature-based algorithm using neural network.

**Submitted classification image** **Classiffied Image**

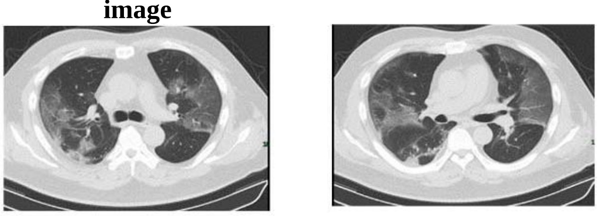

Fig. 7. Covid image classification result for experiment 1

**Submitted classification image** **Classiffied Image**

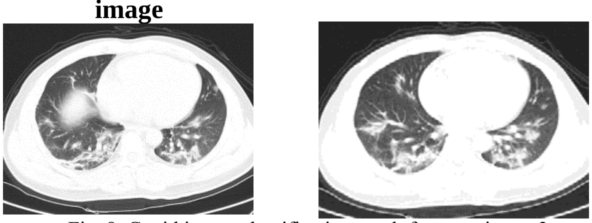

Fig. 8. Covid image classification result for experiment 2

**Submitted classification image** **Classiffied Image**

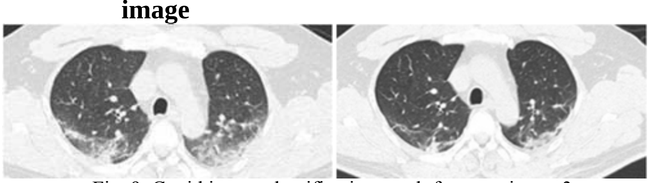

Fig. 9. Covid image classification result for experiment 3

## IV. CONCLUSION

The report presents an algorithm for classification of Covid images into positive and negative class. To this end, wavelet-based feature vectors are generated on the images from the test image database and the query-image and are used for the classification process using NN. For a complete and accurate assessment of the proposed algorithm research experiments are conducted on parameters accuracy, precision and recall. According to the results obtained, the algorithm reaches high result values for the three parameters and is recommended for the cases when false negative is high.

Identify applicable funding agency here. If none, delete this text box.

## ACKNOWLEDGMENT

Primary funding for the presented work was provided by the National Science Fund, Ministry of Education and Science, Republic of Bulgaria under contract KP-06-N37/24, research project "Innovative Platform for Intelligent Management and Analysis of Big Data Streams Supporting Biomedical Scientific Research".

## REFERENCES

[1] J. Krishna, T. Sirisha, V. Madhavi, "Image Retrieval based on Texture using Various Discrete Wavelet Transform Sub Bands", National Conference on Emerging Trends in Computing, pp.22-24, 2017.

[2] B. S. Manjunath, W. Y. Ma, "Texture Features for Browsing and Retrieval of Image Data," IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 837-842, Aug. 1996.

[3] Hong Shao, Jun Ji, Yan Kang, Hong Zhao, "Application Research of Homogeneous Texture Descriptor in Content-based Image Retrieval," International Conference Information Engineering and Computer Science (ICIECS), pp. 1-4, 2009.

[4] M. Kokare, P. K. Biswas, B. N. Chatterji, "Rotation-Invariant Texture Image Retrieval Using Rotated Complex Wavelet Filters," IEEE Transactions Systems, Man and Cybernetics- Part B: Cybernetics, vol. 36, no. 6, pp. 1273-1282, Dec. 2006.

[5] B. Liao, F. Peng, "Rotation-Invariant Texture Features Extraction using Dual Tree Complex Wavelet Transform," International Conference Information, Networking and Automation, vol.1, pp. 361-364, 2010.

[6] A. Mosleh, F. Zargari, R. Azizi, "Texture Image Retrieval Using Contourlet Transform," International Symposium Signal, Circuits and Systems (ISSCS), pp.1-4, 2009.

[7] Xin-Wu Chen, Guang-Li Yu, Jun-Bin Gong, "Contourlet-1.3 texture Image Retrieval System," Proceedings International Conference Wavelet Analysis and Pattern Recognition, pp.49-54, July 2010.

[8] I. J. Sumana, Md. M. Islam, D. Zhang, G. Lu, "Content Based Image Retrieval Using Curvelet Transform," 10th Workshop Multimedia Signal Processing, pp.11-16, 2008.

[9] B. S. Bama, S. Raju, "Fourier Based Rotation Invariant Texture Features for Content Based Image Retrieval," National Conference Communications (NCC), pp. 1-5, 2010.

[10] J. Hu, A. Mojsilovic, "Optimal Color Composition Matching of Images," Proceedings 15th International Conference on Pattern Recognition, DOI:10.1109/ICPR.2000.902862, pp. 47-50, Sept. 2000.

[11] N S T Sai, R C Patil, "New Feature Vector for Image Retrieval: Sum of Value of Histogram Bins," International Conference Advances in Computing, Control and Telecommunication Technologies, pp.550-554, 2009.

[12] W. Rasheed, G. Kang, J. Kang, J. Chun, J. Park, "Sum of Values of Local Histograms for Image Retrieval," Fourth International Conference Networked Computing and Advanced Information Management, pp. 690-694, 2008.

[13] J.-L. Shih, L.-H. Chen, "Colour image retrieval based on primitives of colour moments," IEEE Proceedings- Vision, Image and Signal Processing, vol.149. no. 6, pp.370-376, Dec. 2002.

[14] J. Huang, S. R. Kumar, M. Mitra, W. Zhu, R. Zabih, "Image Indexing Using Color Correlograms," Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97), pp.762-768.

[15] Hoang Ng-Duc, Thuong Le-Tien, Tuan Do-Hong, Cao Bui-Thu, Ty Ng-Xuan, "Image Retrieval Using Contourlet Based Interest Points," 10th International Conference on Information Science, Signal Processing and their Applications, pp. 93-96, 2010.

[16] Y. Tsai, "Salient Points Reduscion for Content-Based Image Retrieval," International Journal of Electrical and Computer Engineering, pp. 11-14, 2010.

[17] Y. Li, L. Guan, "An Effective Shape Descriptor for the Retrieval of Natural Image Collections," Canadian Conference Electrical and Computer Engineering, pp. 19601963, May 2006.

[18] D. Dimov, "Fast Shape Based Image Retrieval," International Conference Computer Systems and Technologies- CompSysTech'2003, pp. 296-302, 2003.

[19] Y. Xie, M. OhEigeartaigh, "Shape Descriptor Based on Structural Curvature Histogram for Image Retrieval," World Congress Computer Science and Information Engineering, pp.411-415, 2009.

[20] Г. Гочев, Компютърно зрение и невронни мрежи, Технически Университет, София, 1998.

[21] Zhu, C., Song, F., Wang, Y. et al. Breast cancer histopathology image classification through assembling multiple compact CNNs. BMC Medical Informatics and Decision Making 19, 198 (2019). https://doi.org/10.1186/s12911-019-0913-x.

[22] S. Kaymak, A. Helwan, D. Uzun, "Breast cancer image classification using artificial neural networks," 9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW 2017, 22-23 August 2017, Budapest, Hungary.

[23] Cherrat Em, Alaoui R, Bouzahir H. "Convolutional neural networks approach for multimodal biometric identification system using the fusion of fingerprint, finger-vein and face images," PeerJ Comput. Sci. 6:e248 DOI 10.7717/peerj-cs.248, 2020.

[24] V. Gancheva, I. Georgiev, "Software Architecture for Adaptive In Silico Knowledge Discovery and Decision Making Based on Big Genomic Data Analytics," Proceedings of the 45th International Conference on Application of Mathematics in Engineering and Economics (AMEE'19), AIP Conf. Proc. 2172, 090009-1–090009-8; https://doi.org/10.1063/1.513358, 2019.

[25] K. Abouelmehdi, A. Beni-Hessane, H. Khaloufi, "Big Healthcare Data: Preserving Security And Privacy," Journal Of Big Data, (2018) 5:1, Pp. 1 – 18

[26] I. Ivanov, "Counteraction and protection against modern cyberattacks, phishing and ransomware," International conference "Modern security and modern technologies", New Bulgarian University, Sofia, Bulgaria, 2019, pp. 93 – 98, ISBN 978-619-7383-13-3.

[27] S. Mbonihankuye, A. Nkunzimana, A. Ndagijimana, "Healthcare Data Security Technology: HIPAA Compliance," Wireless Communications and Mobile Computing, 2019:1-7, DOI:10.1155/2019/1927495

[28] Ivanov, I., "Analysis of vulnerabilities in web applications," Proceeding of Science Conference "Current Security Issues," Veliko Tarnovo, vol. 6, 2020, pp. стр. 233 – 236. ISSN 2367-7465.

[29] C. W. Shaffrey, N. G. Kingsbury, and I. H. Jermyn, "Unsupervised image segmentation via Markov trees and complex wavelets", In Proceedings of IEEE International Conference on Image Processing, 2002, pp. 801-804.

[30] J. Magarey, and N. Kingsbury, "Motion estimation using a complex-valued wavelet transform", IEEE Trans. on Signal Processing, special issue on wavelets and filter banks, vol. 46, No 4, April 1998, pp 1069-84.

[31] P. R. Hill, D. R. Bull, and C. N. Canagarajah, "Rotationally invariant texture features using the dual-tree complex wavelet transform", in Proceedings of IEEE International Conference on Image Processing, 2002, pp. 901-904.

[32] B. Liao, and F. Peng, "Rotation-invariant texture features extraction using dual-tree complex wavelet transform", International Conference on Information, Networking and Automation, 2010, pp. V1-361 - V1-364.

[33] S. Vetova, and I. Ivanov, "Image features extraction using the dual-tree complex wavelet transform", 2nd, International Conference on Mathematical, Computational and Statistical Sciences, 2014, pp. 277-282.

[34] 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.

[35] N. Kingsbury A. Zymnis, "3D DT-MRI Data Visualisation using the Dual Tree Complex Wavelet Transform," in Proceedings EURASIP Biosignal Conference, Brno Czech Rep., June 2004, paper 31.

[36] N. Kingsbury, A. Zymnis, Alonso Pena, "DT-MRI Data Visualisation Using The Dual Tree Complex Wavelet Transform," 3rd IEEE International Symposium on Biomedical Imaging: Macro to Nano, vol. 111, pp. 328-331, Los Alamitos 2004.

[37] H. Chen, N Kingsbury, "Efficient Registration of Nonrigid 3-D Bodies," in IEEE Transactions On Image Processing, vol. 21, no. 1, pp. 262-272, January 2012.

[38] T. Hong, N. Kingsbury, "Estimation of the Fundamental Matrix Based on Complex Wavelets," International Conference on Networking and Information Technology (ICNIT), DOI: 10.1109/ICNIT.2010.5508498, pp. 350-354, July 2010.

[39] N. Kingsbury, "A Dual-Tree Complex Wavelet Transform with Improved Orthogonality And Symmetry Properties," in Proceedings International Conference on Image Processing, pp. 375-378, 2002.

[40] D. Cui, Y. Liu, J. Zuo, Bo Xu, "A Modified Image Retrieval Algorithm Based on DTCWT," in Journal of Computational Information Systems, vol.7, no.3, pp. 896-903, 2011.

[41] R. Chauhan, Dr. R. Dwivedi, Dr. R. Bhagat, "Comparative Analysis of Discrete Wavelet Transform and Complex Wavelet Transform For Image Fusion and De-Noising," in International Journal of Engineering Science Invention, vol. 2, issue 3, March 2013 pp.18-27.