FORMALIZING IMAGE PROCESSING IN HIGHER ORDER LOGIC(HOL) BY UNDERSTANDING AND USING XML-HOL-SCALA-JVM SOFTWARE FRAMEWORK TOWARDS PROCESSING OF CRYO-EM/TEM/SEM IMAGES BASED ON LEVY PROCESSES - A NOVEL SUGGESTION IN THE CONTEXT OF INSTRUMENTATION & HIGH PERFORMANCE COMPUTING ENVIRONMENT.

D.N.T.Kumar

Independent Consultant : Informatics & Nanotechnology.
email id : hmfg2014@gmail.com

Abstract :

In this research work it was proposed to understand and use Levy processes based on XML-HOL-Scala/JVM software in the context of Cryo-EM/SEM/TEM image processing.Though the discussion is very much related to Cryo-EM based topics similar logic could be applied in developing a convenient informatics framework to process other electron-microscopy images.It is one of the pioneering efforts in this direction to use Levy processes in such situations.In the near future,inorder to achieve “Smart & Intelligent” image processing platforms,we could envision incorporating advanced technology based on HOL-XML-Scala/JVM framework which will allow the seamless integration of technologies such as agents, model-based reasoning,genetic algorithms,and artificial neural nets towards robust image processing tasks on-hand.

Keywords : Cryo-EM/SEM/TEM/LEVY Processes/XML-HOL-SCALA/JVM/NanoBioSystems.

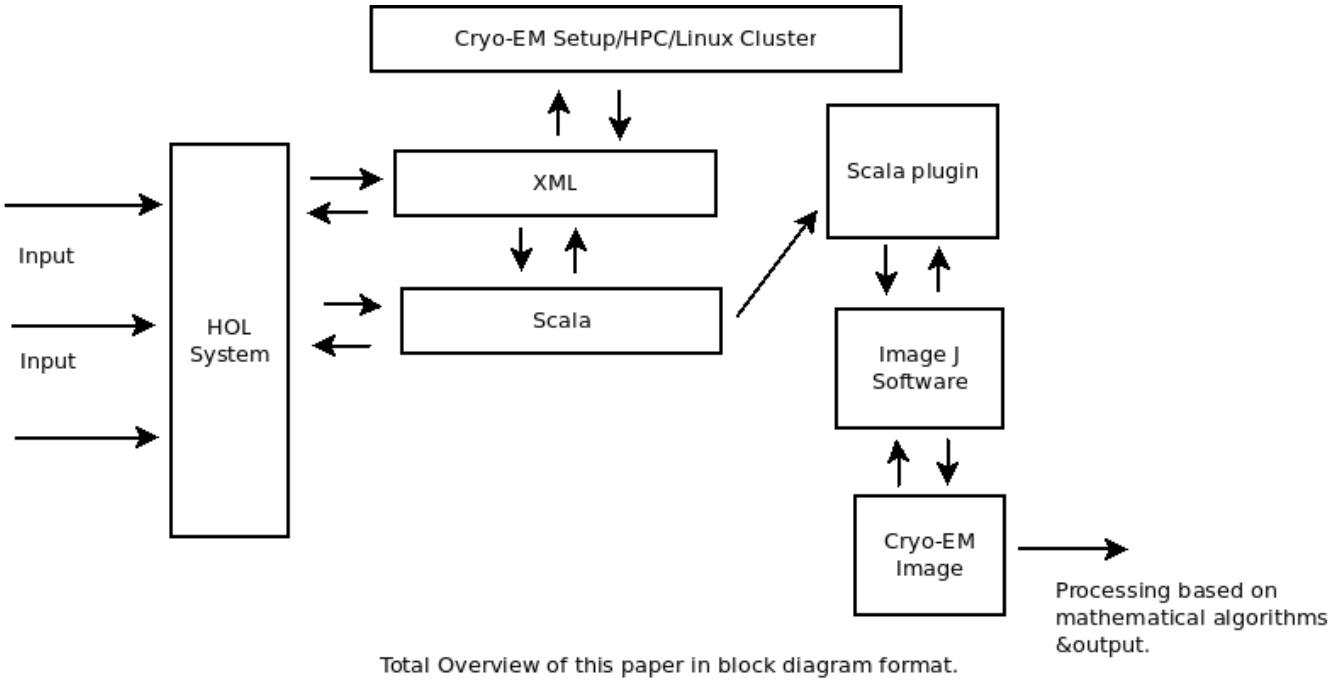
I. Introduction,Inspiration & Background :

Many disciplines are rapidly adopting XML as the universal format for information exchange. With its promising features of simple structure, published standards, platform independence, and ubiquitous support, XML's attraction can conceal the practical challenges of developing a truly usable XML Schema. Successful adoption within a technical discipline requires expert knowledge of both the application domain and XML technology, as well as being technically savvy in forging agreements with the potential adopters.In this context,we apply XML based Instrumentation and high Performance computing to process an ever evolving demanding domain like Cryo-EM Informatics.

Current platform is an independent informatics framework,designed to be generic and extensible, so that it can be applied to any instrument capable of being computer controlled. In order to design an extensible and flexible architecture,the anticipated & promising goals of the project are to:

[i] To provide as much platform independence as possible.
[ii] To create a system that is easy to develop, modify, maintain,and extend further.
[iii] To explicitly promote reuse by design and by utilizing emerging technologies that facilitate software reuse in the IoT scenario.
[iv].To greatly reduce the implementation time for facility instruments, which are reliable, robust, state-of-the-art instruments that are easily used by scientists other than the instrument's designers.
[v].To clearly define the interface between hardware and software engineers.
[vi]. To facilitate multiple iterations of the instrument description during design and implementation by means of a software architecture that is readily adaptable to such changes,by separating the implementation from description.

II. Hardware Framework Implementation :



Total Overview of this paper in block diagram format.

Figure I. TOTAL OVERVIEW OF THE IDEA.

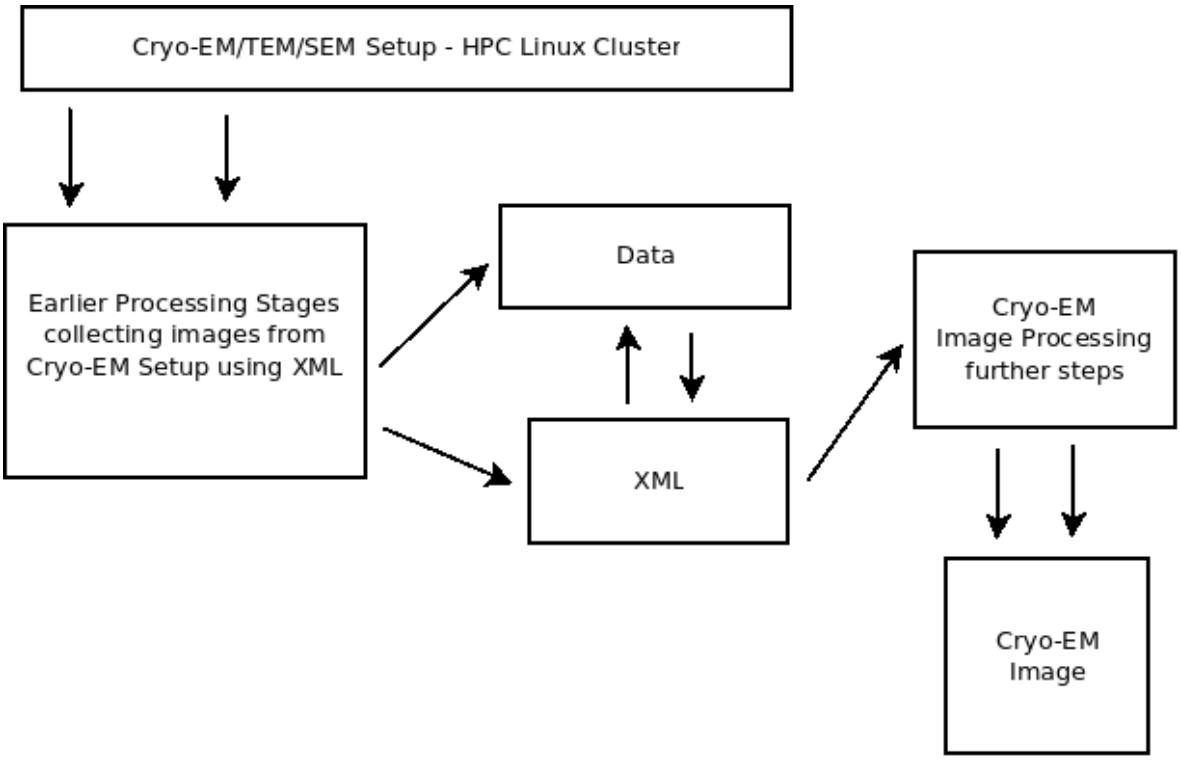


Figure II. [a] Simple Instrument Control Architecture For Image Acquisition & Processing.

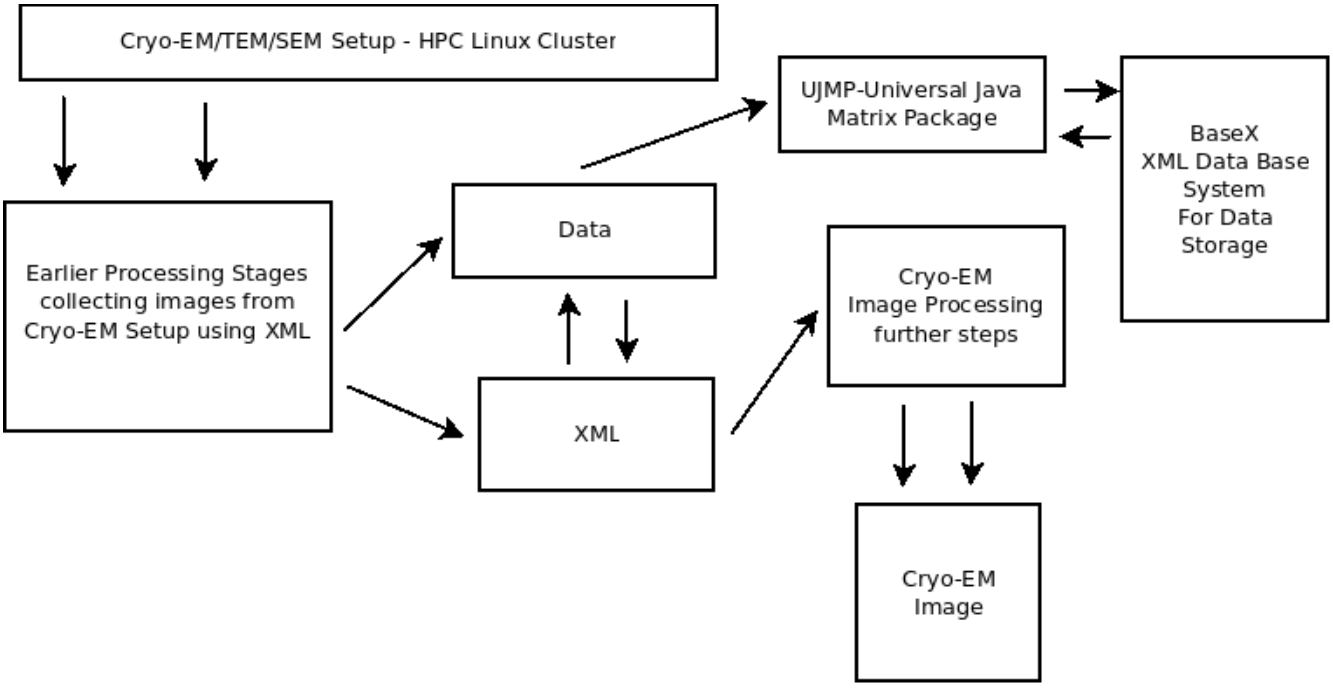


Figure II. [b] Simple Instrument Control Architecture For Image Acquisition & Processing. Gives us the options of data storage using UJMP/BaseX – XML Database.

In parallel with our efforts to improve the general Image-based Processes framework, we will be applying the framework to the development of HOL based instrument control and monitoring software.Applying the HOL-XML-Scala-JVM informatics framework to a new instrument involves taking some or all of the following steps:

1. To develop the instrument description. For HOL Instrumentation, this will involve mapping the Instrument/s from each of the instrument subsystems, including the HPC for the subsystem.

2. To develop the instrument-specific real-time pipeline algorithms.

3. To develop custom GUIs for the various instrument users.

4. To develop new visualizations.

5. To develop instrument-specific scripts.

6. To develop any instrument-specific delegates (for special purpose parsing or response handling).

III. Mathematical Algorithm Design & Software Framework Implementation :

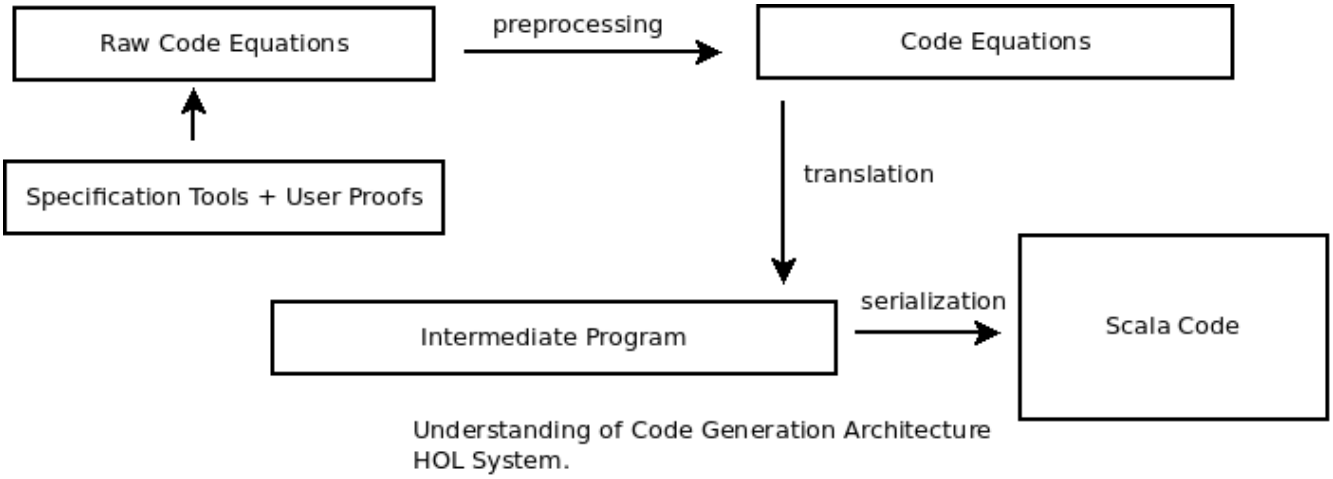


Figure III : Code Generation Architecture in the context of HOL system.
Adopted from : Code generation from Isabelle/HOL theories by Florian Haftmann with contributions from Lukas Bulwahn.

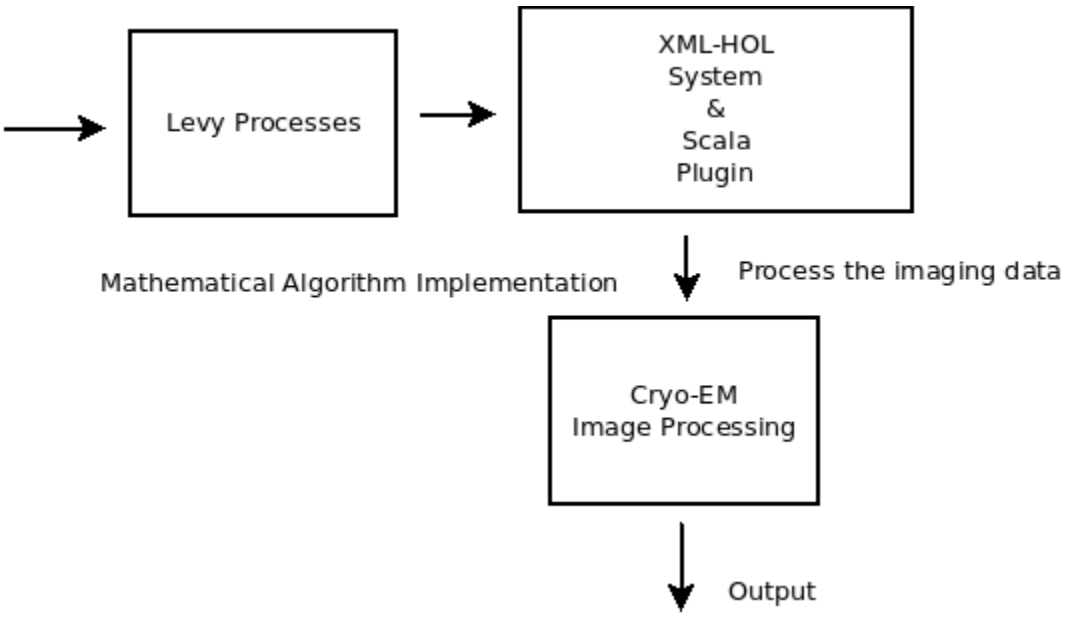


Figure IV. Mathematical Algorithm Implementation.

(Inspiration from : The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging by Ilya G Goldberg et al.
Source : Genome Biology 2005, 6:R47 (doi:10.1186/gb-2005-6-5-r47).

XML and Scientific File Formats by Robert E. McGrath - National Center for Supercomputing Applications University of Illinois,Urbana-Champaign,IL,USA.(August, 2003)).

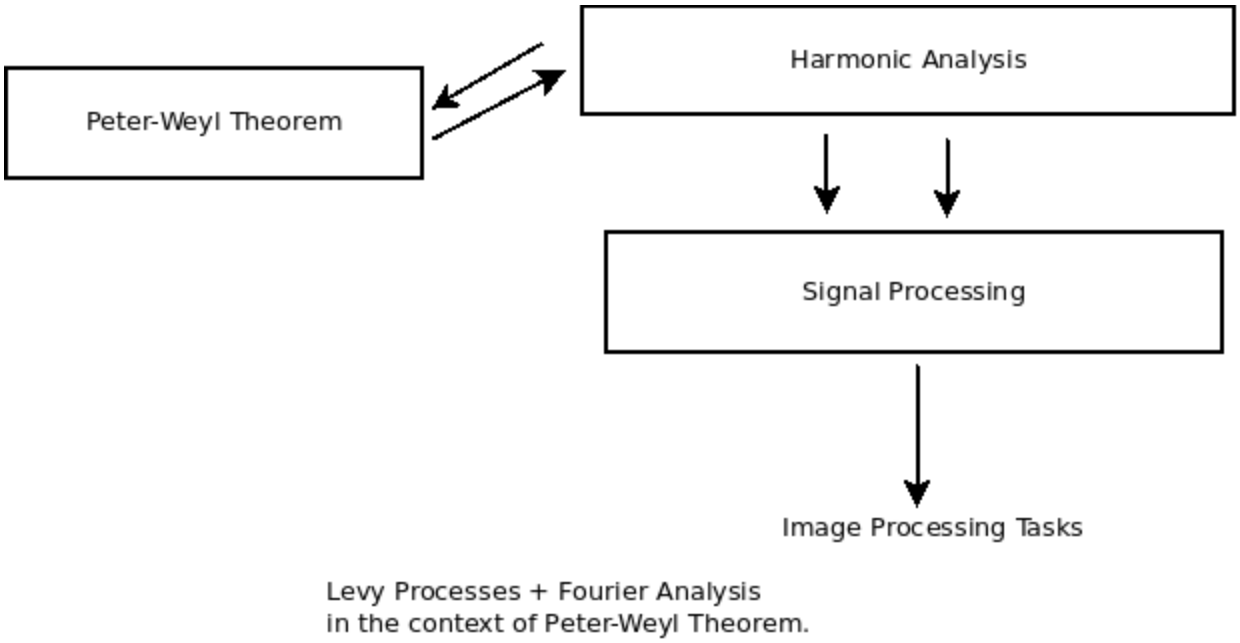


Figure V. Mathematical Algorithm Implementation.

It is possible to study the Fourier expansion of the distribution density of a Levy process in a compact Lie group based on the Peter-Weyl theorem. Levy processes and Fourier analysis on compact Lie groups [42].

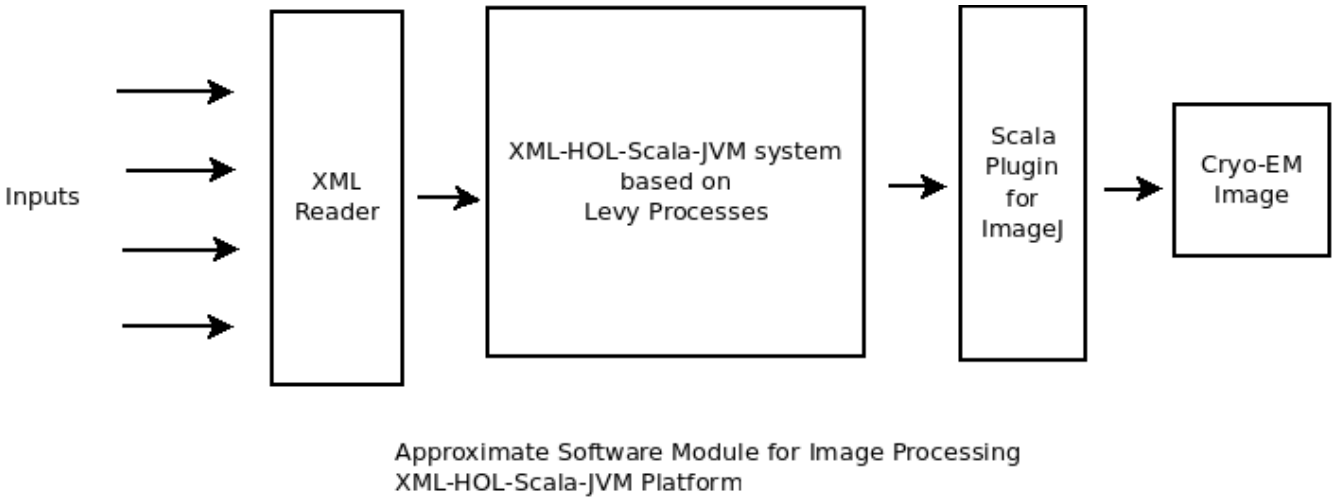


Figure VI. Software Module description in block diagram form.

Please Note : This is one of the pioneering papers on Levy Processes to process Cryo-EM/TEM/SEM images.Here only approximate methodologies are presented to highlight the importance of Levy Processes. Informatics and software architecture or frameworks is/are approximate only.Actual implementation and processing may vary. It is sincerely hoped that the readers will join me in exploring the applications of 'Levy Processes',further.

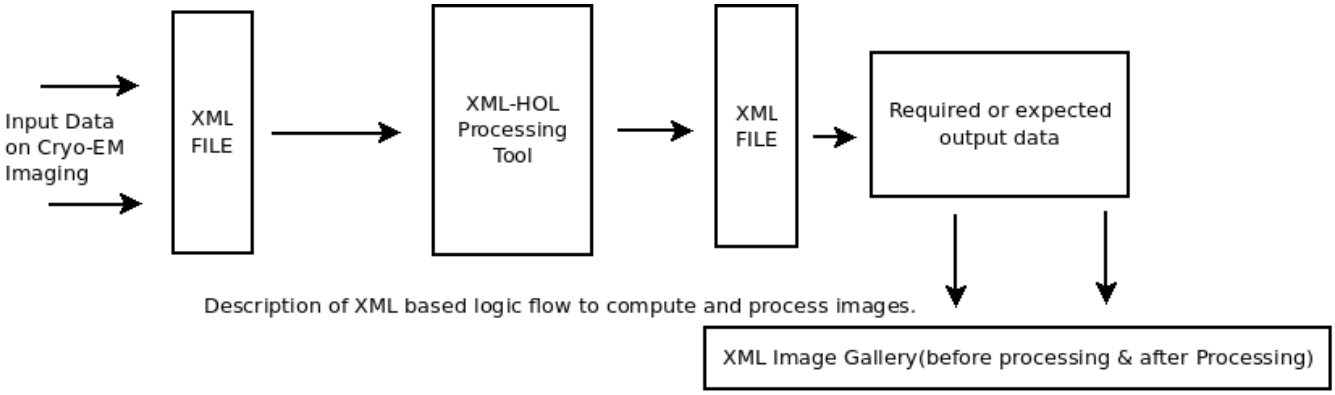


Figure VII. Description of XML based logic flow

Develop Algorithms on the following sequences :

Information Imaging Channel/Information Processing/Storage of Information/Visualization and Presentation.

Although the implementation of individual modality will be different, the “life-cycle” of these modalities is similar and can be grouped as follows:

1.Image Information Processing.
2.Information Repository.
3.Indexing and Searching.

4.Visualization and Presentation.
5.By introducing the modality concept, the integration interface and information exchange of different modalities (dimensions) can be easily identified. This provides an efficient way to adding different new modalities into the system without losing the flexibility.

Main idea of the current work is derived from :

“Java-Based Remote Viewing and Processing of Nuclear Medicine Images: Toward “the Imaging Department Without Walls” by Piotr J. Slomka et al. J NuclMed2000;41:111-118.
“Lévy Processes—From Probability to Finance and Quantum Groups” by David Applebaum; Notices of the AMS,Vol-51 Number-11;1336-1347;Nick Webber - Numerical Methods with Lévy Processes.(very useful to design an algorithm using Levy Processes)

Inspiration to use Levy Processes :

“Levy processes are the stochastic processes with stationary and independent increments and cover a large class of stochastic processes, for example, Brownian motions, Poisson processes,stable processes and subordinators (see Kyprianou, 2006). Moreover, Levy processes have been regarded as prototypes of semi-martingales and Feller-Markov processes (Applebaum, 2004b; Sato,2004). Levy processes have been successfully applied to practical applications in physics (Applebaum, 2004a),signal processing (Duncan, 2009),image processing (Pedersen et al., 2005).”

Lectures on Lévy Processes and Stochastic Calculus, Braunschweig,Lecture 2: Lévy Processes by David Applebaum,University of Sheffield,UK,July 22nd - 24th 2010.

Inspiration to use Fourier Analysis :

“Fourier Analysis is a powerful tool even when periodicity is not directly a part of the problem being solved.Discrete Fourier Transforms (DFT) are well-suited for computation by computer, especially when using Fast Fourier Transform (FFT) algorithms. Fourier Analysis can be used to remove noise from a signal or image. Interpretation of the complex Fourier Transform is not always straightforward. Convolution and Deconvolution are “simple” in Fourier transform space to restore or enhance images. There are many other image processing uses of Fourier Analysis,such as image compression [JPGs use the Discrete Cosine Transform (DCT), which is a special kind of DFT]”

Sources : Fourier Analysis and Image Processing by Earl F. Glynn,Scientific Programmer Bioinformatics;Stowers Institute for Medical Research;14 Feb 2007.

Useful to look at: Topics in Representation Theory - Fourier Analysis and the Peter Weyl Theorem. Fourier method on Lévy process, Affine process and Lévy time change process by David Pommier, March 1, 2012.(It provides a general framework to further develop Levy Processes).

IV. Analysis & Discussion :

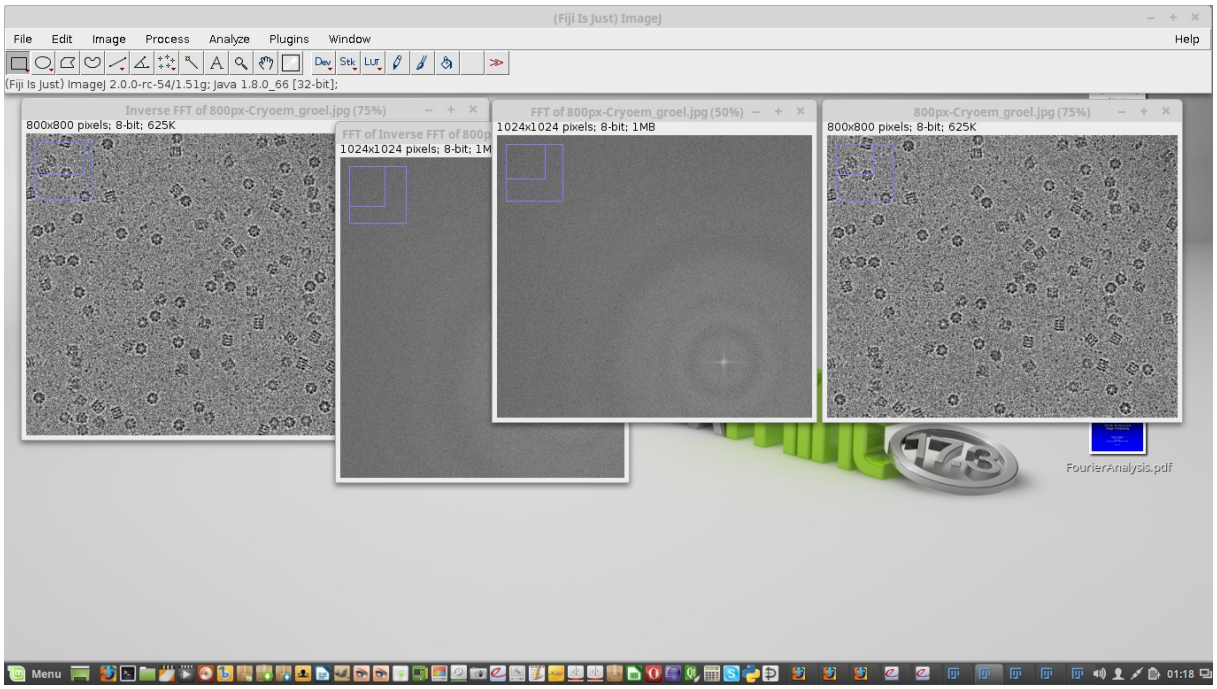


Figure VIII- Image J Software provides an excellent basis to test your informatics framework.Very easy to develop your own plug-in either in Scala or Java and then integrate with the Image J Software. Fourier Analysis on Image J Workstation.HDF5 format of images was used in testing along with others. (Testing in progress on huge datasets)

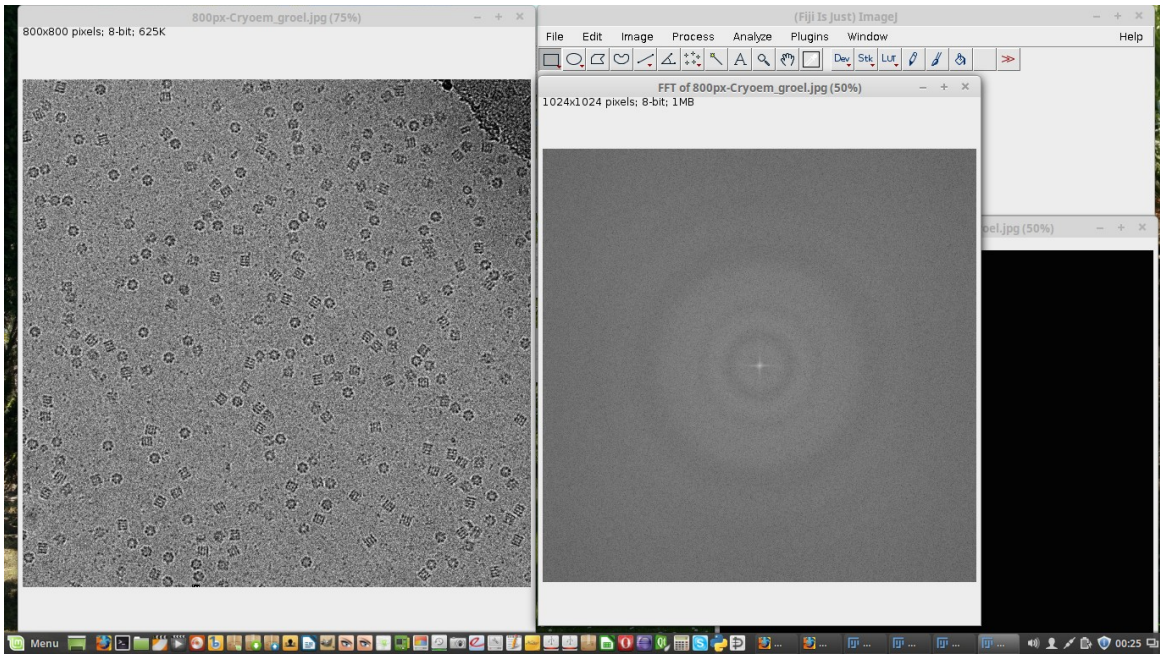


Figure IX- Cryo-EM Image depicting Fourier Analysis. On Image J workstation.

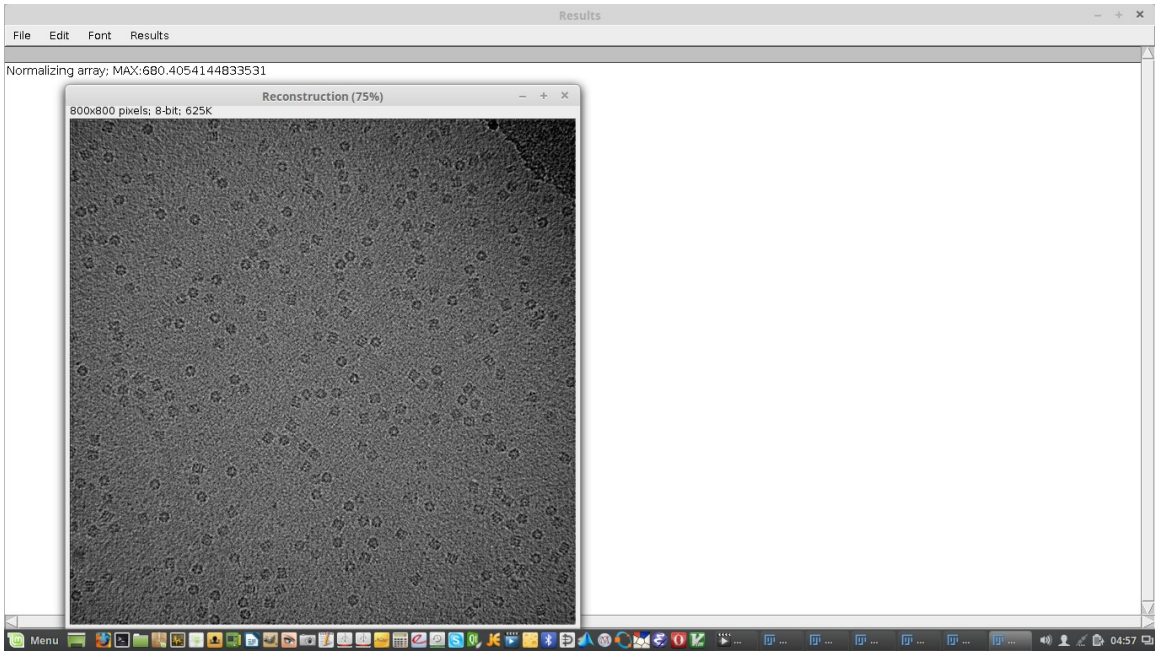
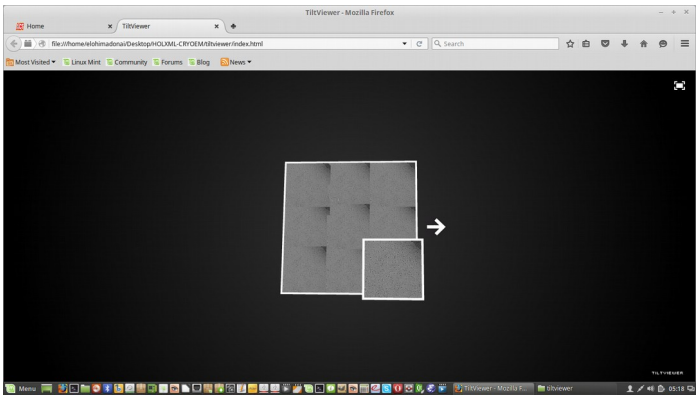
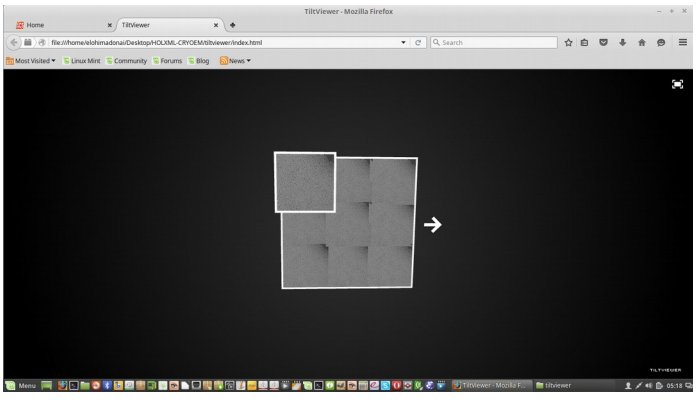


Figure X. Image J or Fiji based Image Processing Workstation. Here user defined plugin option is used to demonstrate one of the main ideas listed in this paper.Tested with both 32 bit and 64 bit Linux OS.Image J Software provides an excellent basis to test your informatics framework.Very easy to develop your own plug-in either in Scala or Java and then integrate with the Image J Software.

**** Please note :** All image processing schema are approximately shown,actual implementation may vary.Readers are advised to check the reading materials and references provided.Advantages and disadvantages have to be examined by the readers as per their requirements before usage.Thank you.

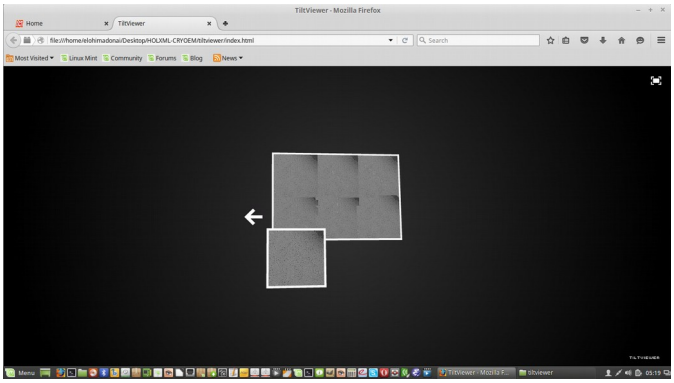


[A]

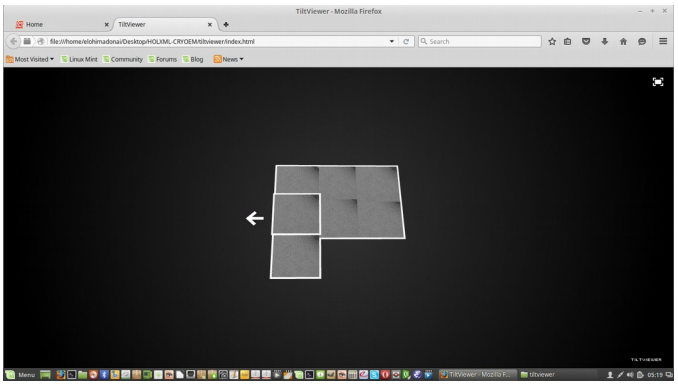


[B]

LÉVY PROCESSES AND FOURIER ANALYSIS TESTING MECHANISMS



[C]



[D]

Figure XI. XML Gallery based manipulations of Cryo-EM Images - Approximate Levy Process Demo.Here “Tiltviewer XML Gallery Software” was used in testing the mechanisms.
[A][B][C][D]

V. Conclusions With Future Perspectives :

Finally the author developed a system with multiple functions for processing and storing information extracted from image files. The system, XML-HOL-SCALA-JVM, integrates different imaging information extraction techniques with a multi-modal concept, generates an XML file for storing the extracted information, and allows users to perform editing on the XML file exported from the system. The multi-modal concept and XML format give users a flexibility to integrate new techniques and to perform further enrichment or modification on the Cryo-EM informatics data in the future. Finally, XML-HOL-SCALA-JVM system comprises an XML to Scala transformer to generate different Scala templates for a seamless multimedia presentation over the Internet of Things Environment(IoT).
So far this is one of the pioneering research papers to highlight “Levy Processes” to tackle Electron-Microscopy Image Processing Framework.

VI. Additional Information on Mathematics & Software Used:

[a].https://web.math.princeton.edu/~amits/publications.html

[b] http://spr.math.princeton.edu/ - Since ASPIRE CRYO-EM Software is in MATLAB .

[c].http://imagej.net/Fiji ; http://www.jikesrvm.org/ ; http://cmib.curie.fr/en/download/softwares/TomoJ

[d]. https://sourceforge.net/projects/scalalab/ ; CEPBA-Tools training. BCN Sept. 2010 – Barcelona Supercomputing Centre,Spain.

[e].CryoEM image of GroEL suspended in amorphous ice at 50,000X magnification.
Source : https://commons.wikimedia.org/wiki/File:Cryoem_groel.jpg
Owner of the image : https://commons.wikimedia.org/wiki/User:Vossman

[f].Isabelle - HOL System : https://www.cl.cam.ac.uk/research/hvg/Isabelle/

[g].Scala JVM Language : http://scala-ide.org/ ; http://www.scala-lang.org/ ; *http://www.scala-sbt.org/*

[h].http://www.jikesrvm.org/ ; http://www.w3schools.com/XML/xml_whatis.asp;

[i]. http://www.newportinstruments.com/cytometryml/cytometryml.htm

[j]. Schaum's Theory and Problems: Signals and Systems, Hwei P. Hsu, 1995, pp. 211-213.

[k] .https://www.math.hmc.edu/seniorthesis/archives/2006/tseacres/tseacres-2006-thesis.pdf

[l] . https://support.hdfgroup.org/HDF5/ ; https://www.simpleviewer.net/tiltviewer/

[m] . http://homepages.inf.ed.ac.uk/rbf/HIPR2/fourier.htm

[n] . Spectroscopic density-matrix Imaging and Potential Inversion: The Rotational Effect,Moshe Shapiro,Dept of Chemical Physics,The Weizmann Institute,Rehovot 76100,Israel.(Sept11,1995).

[o]. The Universal Java Matrix Package: Everything is a Matrix! Holger Arndt, Department of Computer Science, Technical University of Munich, 85747 Garching, Germany.ICML/MLOSS, Haifa,Israel.(2010-06-25); http://basex.org/products/ - XML Databases technology from University of Konstanz,Germany.

Acknowledgements :

Sincerely thank every one who made this happen in my life.No competing financial interest/s are declared in preparing this manuscript.This manuscript is meant to inspire others to develop more advanced Cryo-EM Software in this demanding area using novel methodologies.The Author strictly abides by all copyright agreements in using open source software or other such technologies used in this paper.Purpose - NON-PROFIT ACADEMIC R&D only.

References :

[1]. http://www.snaggledworks.com/em_for_dummies/cryo_em.html

[2]. https://www.ma.utexas.edu/users/hadani/publications.htm

[3]. http://www.math.tau.ac.il/~bernstei/Publication_list/Publication_list.html

[4]. https://yunhokim.wordpress.com/research/cryogenic-electron-microscopy/

[5]. http://cryoem.berkeley.edu/cryoem

[6].D.N.T. Kumar and Gagik Sh.,Understanding JikesRVM in the Context of Cryo-EM/TEM/ SEM Imaging Algorithms and Applications –A General Informatics Introduction from a Software Architecture View Point Vol. 7, Issue 1, January-April, 2016 (IJARITAC) 1-7. DOI: 10.5958/0975-8089.2016.00001.4

[7]. Nirmal Tej Kumar,An Insight into Cryo-EM Imaging Process Architecture Using GENTLE Compiler Construction System with an Informatics Design Paradigm.Vol. 7, Vol. Issue May-August,2016 (IJARITAC)80-89. DOI: 10.5958/0975-8089.2016.00008.7

[8].http://cryoem.berkeley.edu/cryoem

[9]. https://en.wikipedia.org/wiki/L%C3%A9vy_process

[10].http://www.ams.org/notices/200411/fea-applebaum.pdf

[11].http://arxiv.org/pdf/1509.05274v1.pdf

[12].http://math.cts.nthu.edu.tw/Mathematics/2012SPA-slides/1Kyprianou.pdf

[13].http://www.stat.berkeley.edu/~pitman/s205s03/levy.pdf

[14].http://www.math.uconn.edu/~gordina/levygroups.pdf

[15].http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1695527

[16].https://almostsure.wordpress.com/2010/11/23/levy-processes/

[17].http://www.stat.purdue.edu/~figueroa/Papers/LevyModelsReview3.pdf

[18].http://www.pnas.org/content/99/26/16576.full.pdf

[19].http://jmlr.org/papers/volume14/zhang13a/zhang13a.pdf

[20].http://users.minet.uni-jena.de/csb/prj/ot/

[21].https://www.researchgate.net/publication/260400026_A_Case_Study_of_Chemical_Organization_Theory_Applied_to_Virus_Dynam

[22].http://europepmc.org/articles/PMC2712341

[23]. Duke-UNC Brain Imaging and Analysis Center (BIAC, http://www.biac.duke.edu/) {Biomedical Informatics Research Network (BIRN, http://www.nbirn.net/) BIAC XML Header (BXH, http://www.biac.duke.edu/research/bxh/)}

[24].C. E. Catlett, J. Toole, ”Testbeds: From Research to Infrastructure”, in ”The Grid: Blueprint for a New Computing Infrastructure,” Ian Foster and Carl Kesselman, ed., Morgan Kaufmann, August 1998.

[25].J.G.Dennis, E.C.Van Horn, ”Programming semantics for multiprogramming computations”, Communications of ACM, March 1966.

[26].Frédéric Cao, Frédéric Guichard and Hervé Hornung, "Information capacity: a measure of potential image quality of a digital camera", Proc. SPIE 7537, 75370F (2010); doi:10.1117/12.838903

[27].E. Menegatti, M. Zoccarato, E. Pagello, and H. Ishiguro. Image-based localisation monte-carlo localisation without a map. In AI*IA 2003: Advances in Artificial Intelligence : 8th Congress of the Italian Association for Artificial Intelligence, Proceedings, page 423–435, September 2003.

[28].“Formalising Process Calculi in Higher Order Logic” by Monica Nesi, T.R. No. 411, Computer Laboratory,University of Cambridge,UK.January 1997.

[29].https://www.researchgate.net/publication/6820716 - {Web-basedDiagnosticImagingServiceUsing XMLForms}

[30].Hierchical Data Format, “Hierarchical Data Format,” http://hdf.ncsa.uiuc.edu.

[31]. Yan, Kun and McGrath, Robert E., “Experiments with JSP, XML, CORBA, and HDF,” http://hdf.ncsa.uiuc.edu/HDF5/XML/JSPExperiments/

[32]."Lessons Learned from the Creation and Deployment of a Terabyte Digital Video Library." Wactlar, H., Christel, M.,Gong, Y., Hauptmann, A., IEEE Computer, 32(2):66-73,Feb 1999.

[33].Berretti, S., Del Bimbo, A., Pala, P.: Retrieval by Shape Similarity with Perceptual Distance and Effective Indexing: IEEE Transactions on Multimedia, Vol. 2, N°4, pp. 225--239, (2000)

[34].Faloutsos, C., Barber, R., Flickner, M., Flickner, J., Niblack, W., Petkovic, D., and Equitz, W.: Efficient and Effective Querying by Image Content: Journal of Intelligent Information Systems, Vol. 3, pp 231--262, (1994)

[35].Pentland, A., Picard, R. W., Sclaroff, S.:Photobook: Tools For Content-Based Manipulation Of Image Databases: Proceedings of SPIE, Storage and Retrieval for Image and Video Databases II, Vol. 2185, pp. 34--47, April (1994)

*[36].*Smoothness of scale functions for spectrally negative Lévy processes (2006) by T. Chan, A. E. Kyprianou, M. Savov. ; http://citeseerx.ist.psu.edu/showciting?doi=10.1.1.244.6826 .

[37]. T. Ames, L. Koons, K. Sall, (1999). Instrument Remote Control [Online]. NASA Goddard Space Flight Center. Available: http://pioneer.gsfc.nasa.gov/public/irc/

[38]. T. Ames, L. Koons, (1999). IRC Presentations, Publications and Milestones [Online]. NASA Goddard Space Flight Center. Available: http://pioneer.gsfc.nasa.gov/public/irc/IRC-presentations.html

[39]. T. Ames, L. Koons, K. Sall, (1999). Instrument Markup Language [Online]. NASA Goddard Space Flight Center.Available: http://pioneer.gsfc.nasa.gov/public/iml/

[40]. J. Breed, (1999). Code 588 Website – Advanced Architectures Automation Branch [Online]. NASA Goddard Space Flight Center. Available: http://aaaprod.gsfc.nasa.gov/website/WebSite.cfm

[41].Khalil Shihab and Nida Al-Chalabi ;International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:5, No:12, 2011.(Mining and Visual Management of XML-Based Image Collections).

[42].M.Liao,LÉVY PROCESSES AND FOURIER ANALYSIS ON COMPACT LIE GROUPS,The Annals of Probability2004, Vol. 32, No. 2, 1553–1573;DOI 10.1214/009117904000000306.