ANCIENT TAMIL SCRIPT AND RECOGNITION AND TRANSLATION USING LabVIEW

M.Mahalakshmi, MalathiSharavanan

***Abstract*— A simple method for recognition and translation of ancient Tamil inscriptions is done and a detailed study of Ancient Tamil characters and current century characters are presented here the ancient Tamil stone inscriptions are to be recognized and translated using LabVIEW. The images of the ancient script are segmented using segmentation technique the segmentation are done using image segmentation techniques such as PSO, DPSO, FPSO which are the optimization techniques evolved in ancient script segmentation process. For image enhancement we employ contourlet transform where the images are enhanced and noise in the images are removed by fuzzy median filters. The proposed methodology has been simulated under Matlab and LabVIEW.**

***Index Terms*—Segmentation Algorithm (PSO, DPSO), Contourlet transform.**

I. INTRODUCTION

The earliest inscriptional evidences for ancient Tamil occur in Tamil-Brahmi. The conventional consensus for the upper bound dating for Tamil-Brahmi is 3rd Century BCE. Though there have been recent contending evidences for pushing the dates further back as far as 5th Century BCE, they do not have mainstream acceptance.

Tamil-Brahmi had been in vogue for several centuries since it adaptation until it morphed itself into the more cursive Vattaezhuthu around 5th Century CE.The present Tamil script is not a descendant of Tamil-Brahmi. The precursor to the present Tamil script originated around 7th century CE as a derivative of Pallava Grantha with hybrid elements of Vattaezhuthu. Vattaezhuthu itself was in vogue in Tamil Nadu until 11th century CE, until it was completely usurped by the then more prevalent Pallava-derived Tamil script.

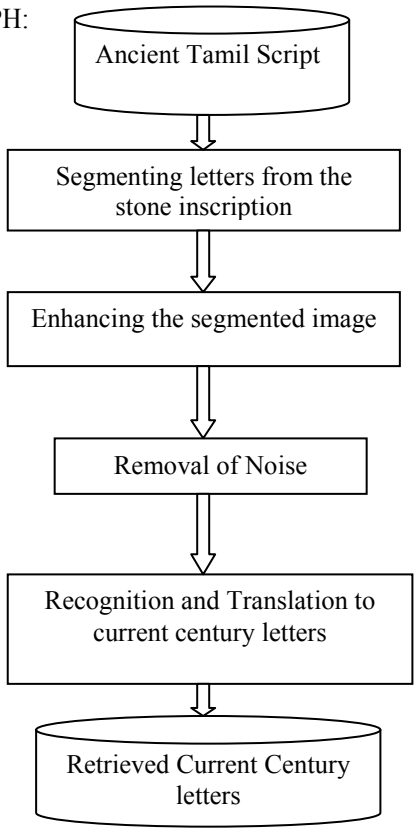
Epigraphs are translated forms of ancient inscriptions that are found on stones, palm leaves etc.There is evidence that epigraphs contain the lifestyle of kings, kingdoms and also medicines. Nowadays, allopathic medicines have led to a lot of side effects, if not immediately, at a later stage .So, people have drifted to Ayurvedic, Siddha and Homeopathy.These medicines originated from ancient medicine practices that were inscribed in temples or palm leaves or were transferred from ancestors. Tamil language has since then been written in wide range of continuum of scripts.

M.Mahalakshmi, Prathyusha Institute of Technology and Management,Chennai. (email: maha.murugasen@yahoo.co.in)

MalathiSharavanan, Prathyusha Institute of Technology and Management, Chennai. (email: mals.sharu@yahoo.com)

LabVIEW plays a vital role in creating epigraphs. The pr - ocess of recognition of the Tamil letters is done by the segmentation algorithm (Particle Swarm Optimization), DPSO (Darwinian PSO), FO-DPSO (Fractional-Order DPSO) and exhaustive methods based on the image histogram shape with the filtering techniques and the image enhancement is executed using countourlet transform.. The segmented ancient script fonts are recognized from a corpus of Grantha scripts and Brahmi scripts and they are translated to Tamil that is presently used.

FLOW GRAPH:



In the [1] existing method Hidden Markovian Model for contourlet transform and the segmentation of a particular image is not done instead they perform the segmentation for the handwritten letters, using adaptive context of multi-dimension fusion method, from the appropriate coarse scale has merged into the segmentation result, and then gets the final image segmentation image here the contourlet is used for the high scale multi decomposition segmentation .

Analysing [8] Contourlet Transform already proposed nine zone segmentation algorithm is implemented but the cursive letter identification is not done through the contourlet transform or curvelet transform ,they give the accuracy but not towards the forward implementation of Matlab techniques.

978-1-4673-4866-9/13/$31.00 ©2013 IEEE

Contourlet transform was proposed as a directional multiresolution image representation that can efficiently capture and represent singularities along smooth object boundaries in natural images. [2] Deals with the main objective of contourlet transform and they say about the efficiency of an object but do not give main idea of implementation of Matlab in contourlet transform and also it has a drawback of localization in frequency domain.

Most probably we switch over to this contourlet transform is to outperform the denoising of a proposed image. [3] Very few ancient scripts (hand written, images) have been recognized, this recognition was carried out by using Optical Character Recognition (OCR) Systems. Hand Written Documents cannot be easily read by OCR software but OCR does not have 100% accuracy. Additional Document Clearance Error is also required still even the image is scanned by OCR has a difficulty in identification of the image letters as O and 0 etc.

Here we go for the Ancient Tamil Recognition and Translation to create a easy mode of study of the epigraphs and ancient Tamil scripts which trends improve the medicinal field.

II. SEGMENTATION

Segmentation plays a fundamental role in subsequent higher level operations in image processing such as recognition object based image Segmentation and feature extraction is very difficult task, where successful results require more complex techniques and good quality data, recently attention was focused on gray level image segmentation with the advance of computational power and instrumentation has evolved the research on the ancient Tamil sculptures works.

In this proposal the ancient image from the sculptures are retrieved and recognized with the help of the segmentation process. For implementing the process of segmentation we go for some algorithms such as (Particle Swarm Optimization), DPSO (Darwinian PSO).

A. Particle Swarm Optimization Segmentation Algorithm:

PSO is a robust stochastic optimization technique based on the movement and intelligence of swarms.PSO applies the concept of social interaction to problem solving. It was developed in 1995 by James Kennedy (social-psychologist) and Russell Eberhart (electrical engineer).It uses a number of agents (particles) that constitute a swarm moving around in the search space looking for the best solution.The first step in the segmentation stage is to extract the wavelet packet based feature set of the new image and then classify each pixel in this image as a desired object pixel (true) or undesired object pixel (false) according to the prior textural model. This classification processes is carried out by using the linear fisher discriminate algorithm.

Finally, this stage is completed by applying the PSO algorithm to get the level set function that truly segments the image as we will clarify in the next sections. Each particle in the PSO population consists of the set of parameters that control the shape of the segmenting curve the training stage to define the level set function that implicitly represents the segmenting curve as in the following:

1) Extraction of control points (features)
2) Matching features
3) Estimation of the optimal transformation.

The dimensional PSO process of each particle may then move the particle to another dimension where it will remember its positional status and keep recognition within the positional The swarm, on the other hand, keeps track of the best particles in all dimensions, each of which respectively indicates the best (global) position so far achieved and can thus be used in the regular velocity update equation for that dimension.

Similarly the dimensional PSO process of each particle uses its personal best dimension in which the personal best fitness score has so far been achieved. Each particle is treated as a point in an N dimensional space which adjusts its "flying" according to its own flying experience as well as the flying experience of other particles.

The best particle in all best dimensions represents the optimum solution (and the optimum dimension).After we configure the PSO algorithm and adjust the curve parameters accordingly to the desired letter, we carry out the segmentation process according to the following sequence:

1. Select the curve parameters randomly from the range specified and create the Corresponding level set functions.
2. Segment the image by using the curves derived from the generated level set functions.

We herein extend the concept of the algorithms to image segmentation. In other words proposing a new thresholding based segmentation method which is robust, makes up our main goal.

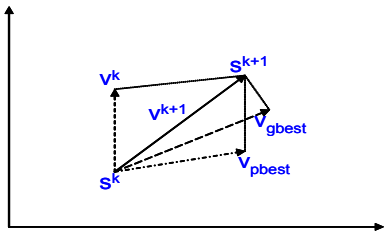


Fig 2.1 PSO swarm movement graph

s^k : current searching point. s^{k+1}: modified searching point. v^k: current velocity. v^{k+1}: modified velocity. v_{pbest} : velocity based on pbest. v_{gbest}: velocity based on gbest

The first is a multidimensional extension, the so called MD PSO, which presents a substantial improvement over PSO via inter-dimensional recognition. Accordingly, each particle keeps track of its last position, velocity and personal best positionThe modification of the particle's position can be mathematically modeled according to the following equation:

$$\mathbf{V_i^{k+1} = w V_i^k + c_1\ rand_1\ (\ldots)\ x\ (pbest_i - s_i^k) + c_2\ rand_2\ (\ldots)\ x\ (gbest - s_i^k)} \quad \mathbf{(1)}$$

where, v_i^k: velocity of agent i at iteration k,
w: weighting function,
c_j : weighting factor,
rand : uniformly distributed random number between 0 and 1,
s_i^k : current position of agent i at iteration k,
$pbest_i$: pbest of agent i,
gbest: gbest of the group.
The following weighting function is usually utilized by using (1)

$$w = wMax-[(wMax-wMin) \times iter]/maxIter \quad (2)$$

Where wMax= initial weight,
wMin = final weight,
maxIter = maximum iteration number,
iter = current iteration number.

$$s_i^{k+1} = s_i^k + V_i^{k+1} \quad (3)$$

PSO is the only algorithm that does not implement the survival of fittest.Finally the swarm keeps track of the global best dimension, best, among all the personal best dimensions. Further, in order to show the advantages of the new methods been commonly used in the literature to determine the (n -1) Optimal n-level threshold on given images.

B. Darwinian PSO Segmentation Algorithm:

Having a local optimum problem with the PSO is such that it may work well on one problem but may fail on another. In order to overcome this problem many authors have suggested other adjustment to the parameters of the PSO algorithm combining fuzzy logic (FAPSO) where the inertia weight w is dynamically adjusted using fuzzy ''IF–THEN'' rules or Gaussian approaches (GPSO) where the inertia constant w is no longer needed and the acceleration constants q1,q2 and q3 are replaced by random numbers with Gaussian distributions.

More recently, Pires et al. used fractional calculus to control the convergence rate of the PSO. Alternatively many authors have considered incorporating selection, mutation and crossover, as well as the DE, into the PSO algorithm. The main goal is to increase the diversity of the population by either preventing the particles to move too close to each other and collide or to self-adapt parameters such as the constriction factor, acceleration constants, or inertia weight. The fusion between GA and the PSO originated the GA-PSO which combines the advantages of swarm intelligence and a natural selection mechanism, such as GA, in order to increase the number of highly evaluated agents, while decreasing the number of lowly evaluated agents at each iteration step. Similar to this last one, the EPSO is an evolutionary approach that incorporates a selection procedure to the original PSO algorithm, as well as self-adapting properties for its parameters.

The FODPSO presented in is an extension of the DPSO in which fractional calculus is used to control the convergence rate of the algorithm. The characteristics revealed by fractional calculus make this Mathematical tool well suited to describe phenomena such as irreversibility and chaos because of its inherent memory property.

The PSO, DPSO and FODPSO methods are parameterized algorithms, therefore, one needs to be able to choose the parameter values that would result in faster convergence. The cognitive, social and inertial weights were chosen taking into account several works focusing on convergence analysis of the traditional PSO.Thus the PSO and DPSO algorithms are extended for the segmentation process.

III.IMAGE ENHANCEMENT

In the stone inscriptions if the background has similar grey level with corresponding letters the details of this object is hard to detect. So if the wide letters in the kalvettu image which represent as strong edges are easy to find, the thin curves and dots are hard to detect due to their similar grey levels to the background.

The proposed strategy softens the strongest edges and amplifies the faint edges. We try to reduce the ratio of strong features to faint features so that the slim the curved letters and dots become visible. Since the Contourlet transform is well-adapted to represent images containing edges, it is a good candidate for enhancement in ancient Tamil letter identification as well as edge enhancement in those l images. Contourlet coefficients can be modified via a nonlinear a dynamic range compression.

The image enhancement is a very strongest approach which is employed here is to outfocus the ancient script images and enhance them in the field of the recognition and translation for the proposed work.

A. Contourlet Transform:

Here presented a new multi-scale method for ancient Tamil script/consonant enhancement based on the Contourlet transform. The Contourlet transform has better performance representing edges than wavelets for its anisotropy and directionality, and is therefore well-suited for multi-scale edge enhancement. We modify the Contourlet coefficients in corresponding sub bands via a nonlinear function and take the noise into account for more precise reconstruction and better visualization.

We compare this approach with enhancement based on the Wavelet transform, Histogram Equalization. The application of this method on images from the drive database states that the proposed approach outperforms other enhancement methods on low contrast and dynamic range images**.**

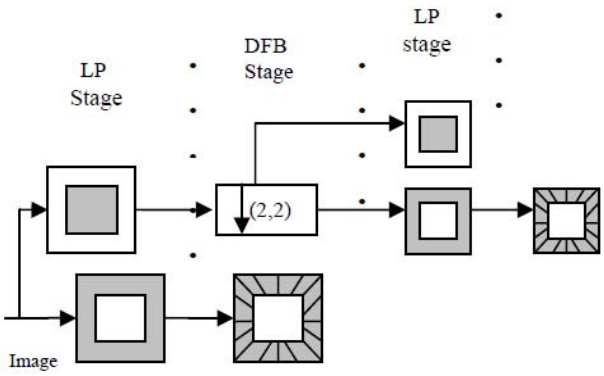


Fig 3.1 Contourlet Transform

A double filter bank structure of the contourlet the Laplacian Pyramid is used to capture the point discontinuities, and then followed by Directional Filter Bank (DFB), which is used to link these point discontinuities into linear structures.The contourlets have elongated supports at various scales, Directions and aspect ratios.

These properties of CT, i.e., directionality and anisotropy made it a powerful tool for image enhancement The contourlet transform is an extension of the wavelet transform which uses multiscale and directional filter banks. In practice wavelets are not effective in representing the images with smooth contours in different directions. Image representation in the contourlet transform domain addresses these problems by having the following properties such as:

• **Multiresolution**: The representation should allow images to be successively approximated, from coarse to fine resolutions.
•**Localization**: The basic elements in the representation should be localized in both the spatial and the frequency domain
•**Directionality**: The representation should contain basis elements oriented at a variety of directions, much more than the few directions that are offered by separable wavelets.
•**Anisotropy**: To capture smooth contours in images, the representation should contain basis elements using a variety of elongated shapes with different aspect ratios.
Among these directionality and anisotropy made contourlet transform an effective tool for content based image retrieval It is constructed by combining the Laplacian pyramid with the directional filter bank (DFB) as proposed in. Laplacian Pyramid (LP) is used to capture the point discontinuities and Directional Filter Bank (DFB) is used to link these points discontinuities into linear structures.

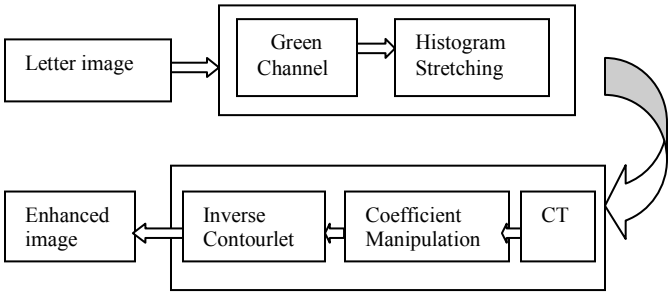


Fig 3.2 Image Enhancement Module

This transform is suitable for applications involving edge detection with high curve content. This approach involves taking the contourlet transform of test gray scale images. The ancient Tamil letter is taken as input to the histogram module where the green channel and histogram matching is performed. By using the contourlet transform the multiscale decomposition is done along with the directional filter bank properties, segmented is made or the letter is enhanced using the PSO and DPSO techniques and the coefficient manipulation is automated for the prehistogram images.

At first the image from the stone inscription is retrieved from the stone inscription they are passed into the green channel where the RGB colour module is extended the output of the green channel is fed into the Histogram Sketching where this does the sketching of those retrieved images. The very next step is to follow the contourlet transform done using the combination of the multi directional filter bank and Laplacian Pyramid which executes the letter image with directionality and high aspect with anisotropy then the process of coefficient manipulation is processed where the comparison of the threshold of each image is followed ,finally the process of image contourlet is done to have the segmented image and the enhanced image following from the module of image enhancement using contourlet transform.

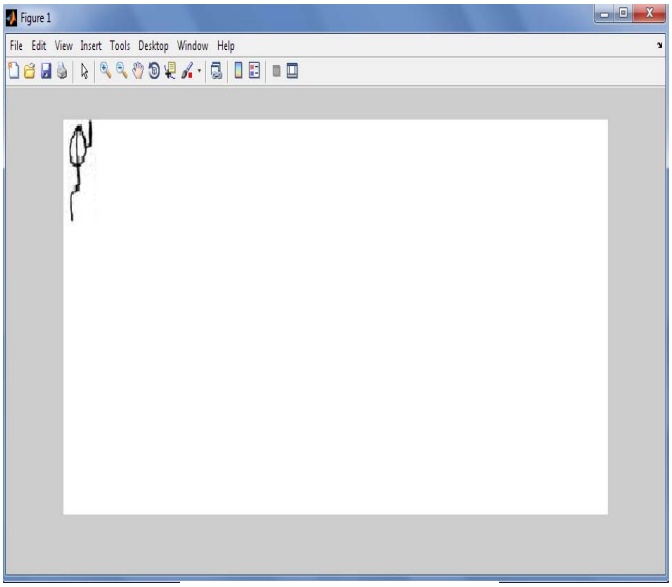
B.LP Decomposition

To obtain multiscale decomposition Laplacian Pyramid is used. LP decomposition at each level generates a down sampled low pass version of the original image and difference between the original and the prediction, that results in a band pass image.

In LP decomposition of an image is the main part for recognition in this proposal. Along with the recognition this also follows up with the noise removal from the particular character.LP combines with the median filter and thus the output for recognized and translated ancient century letter is retrieved. Normally we go for the fuzzy median filter for the removal of the noise which has the combination of the Laplacian Pyramid and Guassian Pyramid which in turn gets the noise free output.

VI. SIMULATION RESULTS AND DISCUSSION

Simulation is done using Matlab and LabVIEW results are presented here. Figure 4 shows. The simulations in Matlab environment is executed for the segmentation and contourlet transform similarly the output of the recognized and translated character is explored in LabVIEW model.

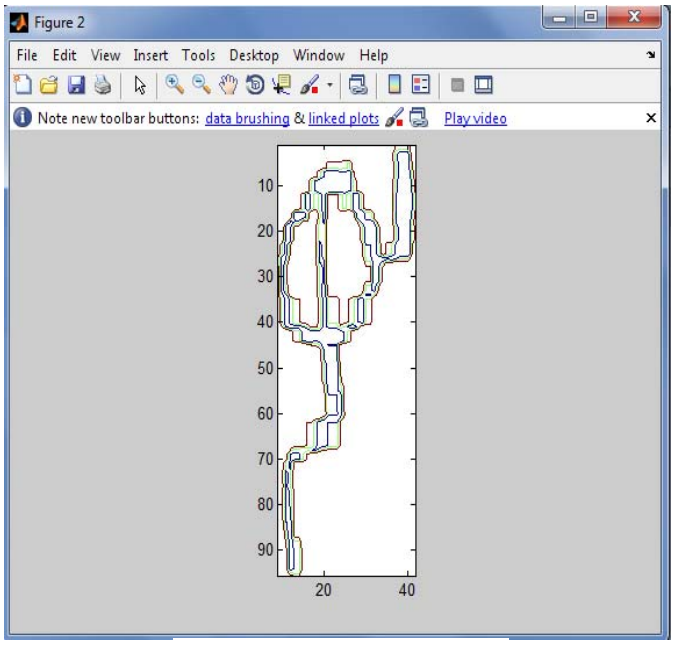


Segmented Input Character

Fig 4.1 Input Character of the segmented image from stone inscription.

The segmented output is executed using the PSO, DPSO technique is shown in the above Fig 4.1, where the input from

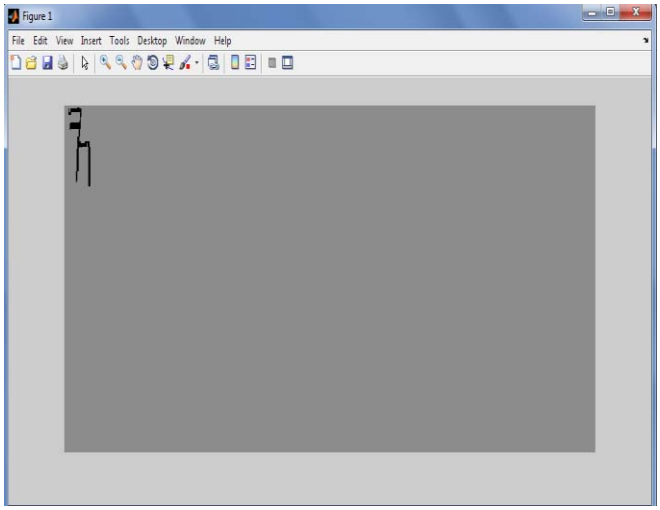
the stone inscription is fed into the Matlab and the algorithmic process is done to have the output.



Enhanced Input Character 1

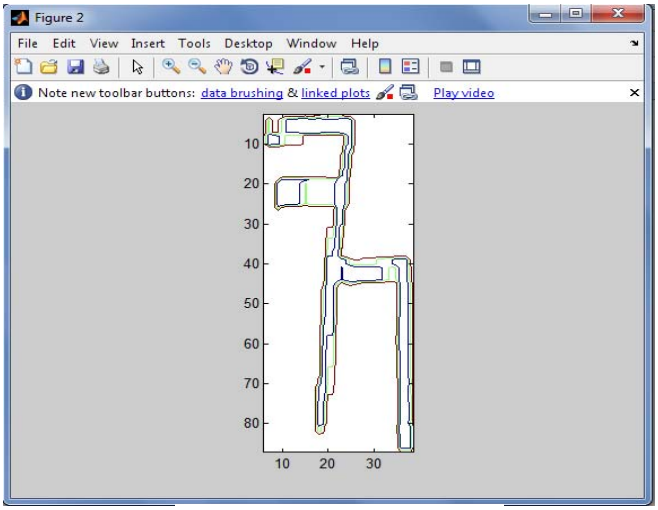
Fig 4.2 Input Character 1 executed with the countourlet transform

The process of contourelet transform is executed and the curves are easily identified and the output is showed in the above Fig4.2.When we go for the wavelet transform the curves and dots in the stone inscriptions are not segmented as well as enhanced so we switch over to the contourlet transform which gives the enhanced image output with RGB identification and makes the process easy i.e. Letters are easy to identify



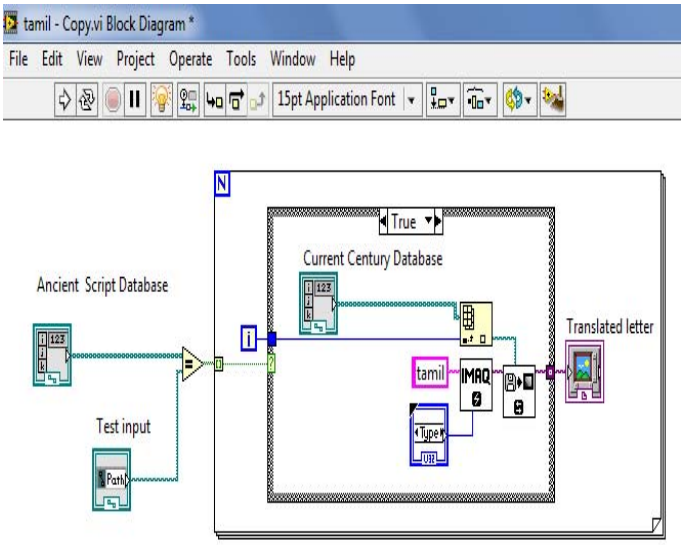
Segmented Matlab module

Fig 4.3 Input Character 2 The second letter segmented module is shown in the above fig as to illustrate that this is been executed in the Matlab.



Enhanced Input Character 1

Fig 4.4 Input Character 2 executed with the countourlet transform



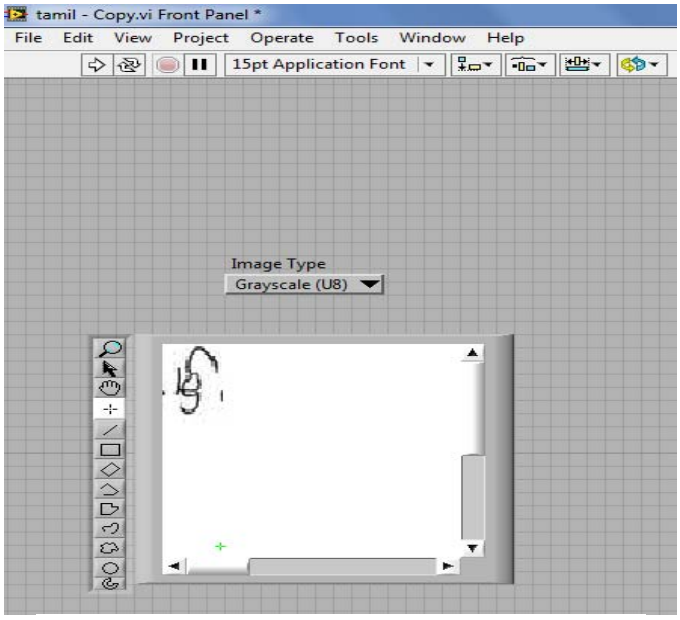
Block Representation in LabVIEW

Fig 4.5 Block Diagram of the LabVIEW

The block diagram gives the database creating model for the century identification in Fig 4.5 .Thus followed by the recognized and translated output in the vi front panel module.

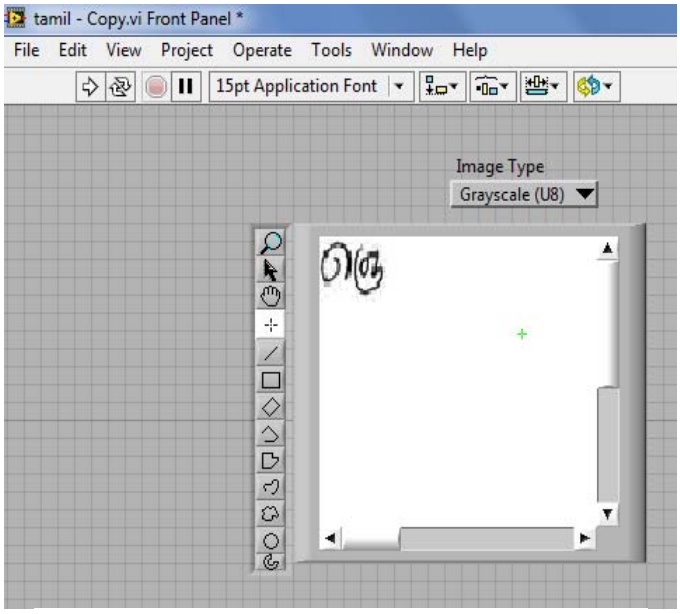
The block representation of the LabVIEW model is presented in the Fig 4.5 this shows the database management and how the century identification is done and the output us retrieved.

The database are compared in the LabVIEW and the letters are recognized and compared with the current century reference database which is already stored.Thus by comparing the letter the equivalent letters are displayed in the front panel of the LabVIEW.Here we go for the unique specification of the NI Vision Module for displaying the output.



Recogoinized and Translated Vi Front Panel Module:

Fig 4.6 Output Character 1 run under the simulated condition in the LabVIEW front pannel



Recogoinized and Translated VI Front Panel Module:

Fig 4.7output character 2 this simulation is done under the LabVIEW environment

The m-files of the segmented Brahmi scripts are given as input to the VI module and thus the current century letter equivalent of the segmented Brahmi scripts are recognized and translated using LabVIEW where the segmentation is done using the matlab algorithm and the image enhancement is also proposed by Matlab as shown in the Fig 4.1 and Fig 4.3 finally the mfile is called in Labview and translated equivalent day today Tamil character and displayed in the forntpannel of the LabVIEW module as shown in the Fig 4.6 and Fig 4.7

V. CONCLUSION

This work outlined the implementation of recognition and translation of ancient Tamil scripts to current century consonants with the process of segmentation and contourlet transform. The scope of this work is segment the ancient inscription and translates them into the current century letter by the implementation of the segmentation algorithms and image enhancement module. The process of segmentation is done in MATLAB where as the M-file is called in LABVIEW to have the recognition and translation work done.

REFERENCES

[1] Harbin "The Image Segmentation Algorithm Based on Contourlet Transformation Proceedings of the 2010 Vol 14, 19 IEEE International Conference on Information and Automation.

[2] J.Mitola III and G.Q.Yue Lu and Minh N. Do "A New Contourlet Transform With Sharp Frequency Localization" University of Illinois at Urbana-Champaign, Urbana, IL 61801.

[3]S.RajaKumar,Dr.V.SubbiahBharathi SathiyabamaUniversiy Chennai, India"Eighth Century Tamil Consonants Recognition FromStone Inscriptions" Chennai.

[4] Sharmila S. Abirami S.Murukappan Bhaskaran R.,India. (IJIDCS) "International Journal on Internet and Distributed Computing Systems". Vol: 2 No: 1, 2012 109.

[5] "Base Line Correction for Handwritten Word Recognition" Shinji Tsuruokat Nobuyuki Vatanabet Nariyasu Minamidet Fumit aka Kimurat Yasuji MiyalieS Shridhars Department of Electrical and Computer Engineering,University of Michigan-Dearborn, Dearborn, Michigan 48128-1491, USA.

[6] Enhancing retinal image by the Contourlet transform Peng Feng a,*, Yingjun Pana, BiaoWei a, Wei Jin b, Deling Mia a Key Laboratory of Opto-electronics Technology and System, Ministry of Education, ChongQing University, ChongQing 400044, PR China Faculty of Information Science and Technology, Ning China.

[7] Yue Lu and Minh N. Do "A NEW CONTOURLET TRANSFORM WITH SHARP FREQUENCY LOCALIZATION H." IEEE Trans.Image Proc - Vol. 14, No. 12, December 2009.

[8] G. Zhu, I. F. Akyildiz, and G Gheith A. Abandah, Fuad T. Jamour Computer Science Engineering Department The University of Jordan "Recognizing Handwritten ArabicScript through Efficient Skeleton-Based Grapheme Segmentation Algorithm" IEEE Pross 2009 3,5

[9] http://en.wikipedia.org/wiki/brami_script.