



A File Based GIS anchored Information Retrieval Scheme (FBGISIRS) through Vectorization of Raster map

Anirban Chakraborty

Assistant Professor, Department of Computer Science
Barrackpore Rastraguru Surendranath College
Barrackpore, Kolkata-700 120, W.B., India
theanirban@rediffmail.com

Dr. J.K. Mandal*

Professor, Department of Computer Science & Engineering
University of Kalyani
Kalyani, Nadia, W.B., India
jkm.cse@gmail.com

Arun Kumar Chakraborti

Student, Department of Computer Science
Barrackpore Rastraguru Surendranath College
Barrackpore, Kolkata-120, W.B., India
great.arun786@gmail.com

Abstract: This paper presents a technique to vectorize raster maps/images using classical files system. The technique calculates coordinates of necessary object constructing points from the raster map and stored same into the files. Based on needs of users, mouse clicking/moving tools can be used along with the assignments of different colours to various objects. Heterogeneous data can be associated to different objects. A mapping technique has been generated to retrieve the attributes from the files for visual display/interpretation. Provision has also been made to edit digitized objects for correctness or future upgradation. During editing new objects may be inserted or deleted and colour of an object may be reassigned along with alteration of Associated data.

Keywords: Raster map, vector map, digitization, GIS, information retrieval, file based vectorization.

I. INTRODUCTION

Digitization is the process by which a vector map can be generated from the input raster map. There is no computer understandable object into a raster map because it is a collection of pixels which have some intensity value only. So it is very difficult to generate any GIS feature on this type of maps. The first task in GIS is to digitize the raster map to obtain a vector map. The output of digitization is the vector map which is a collection of polygons, lines and points [2, 3, 4]. Along with the boundary point's coordinates the polygons may have the boundary color and filling color features. Similarly the line and point objects may have the color feature along with the object constructing point's coordinates into the file. It is very easy to generate GIS feature on a vector map because the computer can understand the objects as polygons or lines or points [1].

Section 2 of this paper deals with the scheme of the implementation techniques. The implemented results are given in section 3. Analysis and comparisons are drawn in section 4. Conclusions are drawn in section 5.

II. SCHEME

The process of digitization/vector map generation technique is presented in this section. The total scheme may be achieved through four sub tasks which are as follows:

1. Calculation of coordinates of object constructing points
2. Color assignment & Data Association
3. Storing information
4. Editing

All of these are discussed in section A to section D respectively.

A. Calculation of Coordinates of Object Constructing Points

The digitization technique calculates the coordinates of the object constructing points and stores the coordinates of each object into the files. Now the most important task is to retrieve the co-ordinates of the object constructing points. Obviously, most of the users like to make use of the mouse for this purpose and it is easy to handle also. Here also requires the heavy concentration of the user, because if the mouse fetches some wrong points while gathering object constructing points, then obviously it will cause improper digitization. For example, if mouse-moving method is used and if the user becomes unable to move the mouse minutely through the curvy edges of a given raster map, then the digitization will also not be proper. There is no need to calculate the coordinates of all object-constructing points. For polygon and line digitization only some essential points' coordinates is required to produce the vectorized object. The absolute shape of an object can be constructed by help of these essential points. But more number of points/more closely spaced points generate better vectorised object. Since there are three types of objects (polygon, line and point) exist in a vector map this process can be divided into three sub-tasks. These are:

1. Polygon construction
 2. Line construction
 3. Point object construction
- All of these are discussed in section a through section c .

a. Polygon Construction

The polygons play the most important roles in the maps, e.g. the map of a place is composed of various objects, which in other sense can be viewed as to be composed of various "polygons". So it is necessary to gather the co-ordinates of

polygon generating points very accurately. The polygon is drawn as a combination of many line-segments. If a line is drawn between two co-ordinate points (x_1, y_1) and (x_2, y_2) then its length is $l = \sqrt{[(x_2 - x_1)^2 + (y_2 - y_1)^2]}$. However, when these line segments are used to draw a polygon, the curvy shape of the polygon will not be understood fully, if the points are not closely spaced, because then the line-segments will have some reasonable amount of length. But, if we gather some closely spaced points for drawing a polygon, then as x_1, x_2 and y_1, y_2 are closely spaced so $(x_2 - x_1)$ and $(y_2 - y_1)$ both will have very small values and so length $(= \sqrt{[(x_2 - x_1)^2 + (y_2 - y_1)^2]})$ will also be very small and a set of this like very small-length lines will be able to construct a irregular-polygon accurately.

Two techniques can be applied for the purpose, mouse-clicking method and mouse-moving method. The mouse-clicking method is easier than the other technique. When it is required to digitize a polygon whose boundary is not very much curvy in shape, this technique can be used. Here just the mouse-pointer is placed on the boundary of the polygon, a left-click is made to inform about starting of the digitization of a polygon and then clicks are made on the boundary line. Finally, a right-click is made to inform about the ending of the digitization of that polygon.

The chief advantage of this technique is that, it is easier and claims lesser user attention. Also it is less time consuming. But, if this technique is being adapted to digitize a polygon, whose boundary is very much curvy one, the result obtained is not very much accurate one. For example, let us consider the Political map of India (Fig. 1) and the state Jharkhand (pointed by red-line) has to be digitized.



Figure. 1: Political Map of India (Polygon pointed with red is being digitized)

If the method of mouse-clicking is used, then the resulting shape will be as follows (Fig. 2):



Figure. 2: Result of digitization using mouse-clicking method

To overcome the above mentioned problem, the mouse-moving method can be used. Here also, at first the mouse-pointer is placed on the boundary line of the polygon and a left-click is made to denote starting of digitization. Next the most important part is to move the mouse-pointer over the boundary line and it has to do very minutely because a slight deviation of the mouse-pointer causes fetching of some other points, which are not actually constructing the polygon. So, it causes to re-generation of an improper polygon. Finally, a right-click has to be made to denote the ending of the digitization of the polygon.

Obviously this method quite is tedious, time-consuming and requires a very minute user-attention and concentration compared to the above mentioned mouse-clicking method.

But, the major advantage is that if the user drags the mouse-pointer exactly through the boundary, then the exact shape of the polygon could be re-generated after digitization. As an example and comparison to mouse-clicking method, if the state Jharkhand (shown by red-color, in the above political map of India; Fig. 1) is digitized using mouse-moving method then the shape shown in Fig. 3 is obtained :-



Figure. 3: Result of digitization using mouse-moving method

Obviously, the obtained digitized shape using mouse-moving method is far much better than that obtained by mouse-clicking method. Thus, it is upon the user to choose the correct technique, based upon the current scenario.

The polygon boundary calculation technique is shown in figure 4.

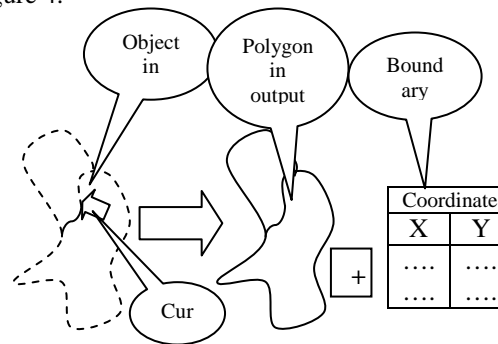


Figure. 4: Polygon boundary calculation technique

b. Line Construction

To digitize a line, it is not necessary to gather the co-ordinates a huge number of line constructing points. Only some points of the line are sufficient to maintain the absolute shape of the line. Therefore, it is recommended to apply the less tedious and less time-consuming mouse-clicking method here, rather than to use mouse-moving method, which will collect a huge number of points (co-ordinates), in expense of time and maximum user concentration. Moreover, to store a very large number of point co-ordinates, the size of the file will also be very large.

c. Point Object Construction

The coordinate of the points can be calculated by the same method that is clicking of mouse.

B. Color Assignment & Data Association

Color has to assign to each objects of the map (e.g. while digitizing the map of a country, each state can be colored differently). Polygons have two types of color which are boundary color and fill color. Lines and points also have color attribute. To represent different type of objects separately different colors may used. In this technique a color chooser button is placed in the digitization page. By this color chooser, colors are assigned to each objects of the map. After assigning color, various data, associated with a particular region (such as, Rainfall, Population etc. for a state in the map) can also be assigned with the object. These associated values, along with the boundary color and fill color are stored in different files, to retrieve latter along with the digitized object. Object's color attributes and association of values of a vector map are shown in figure 5.

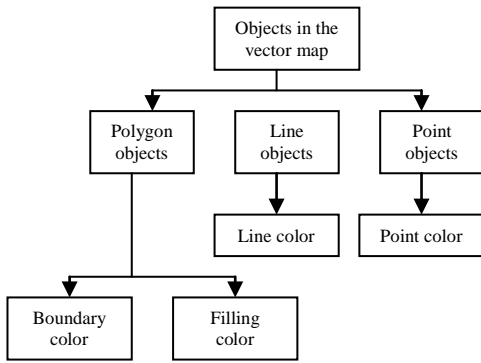


Figure. 5: Object's color attributes of a vector map

C. Storing Information

The saving of the digitized object is actually done following algorithm discussed below:

a. Algorithm:

Step 1: The user has to give a profile name, which, if is same with some existing profile is not accepted. The saved work can latter be opened for further modification (adding or deletion)

Step 2: For every profile one new folder will be created and for every digitized polygon (e.g. states of a country) the X & Y co-ordinate points will be saved in a text file.

Step 3: The values of the assigned colors or the data associated with a particular region (polygon) will also be saved into files in the same folder.

This digitization tool has an extra facility which is temporary storage of partially completed digitization. If the total number of objects is large, the digitization process may be time consuming. For this reason this technique is integrated. Here the digitized objects (maps) are stored as flat-files. The user use different Profile names while saving any digitization work. However, it is very easy to deal with an incomplete digitization work here. For example, in the created profile, say "Political Map of India", if the user digitizes only five states at the beginning and then quit, the software saves the work automatically. Latter, if the user again open the same profile and digitize some more states, the updation on editing will also be done automatically.

D. Editing

To edit (i.e. for adding new polygons, deleting some portions, changing color etc.) the user have to select the desired profile and then editing can be done. If the user wants to delete an object then he/she has to select that object. To select a polygon object, the mouse-pointer has to take inside the polygon object. The polygon selection scheme is shown in figure 6.

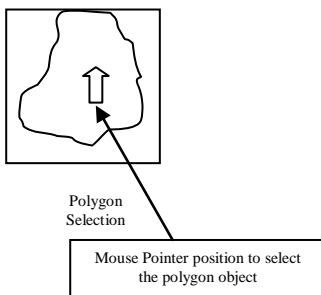


Figure. 6: Polygon Selection

In case of polygon selection the technique search the polygon contains the point pointed by the mouse-pointer and then the polygon will be selected. For this purpose, the "Ray-Casting Algorithm" has been used with some proper modification.

III. IMPLEMENTATION & RESULTS

The implemented results, pages and its operations are presented and discussed in this section. The implementation has been done using Net Beans (Java), which is based on file systems without using databases. It is also the advantage of this tool that, the user needs not to resize the image at all. Moreover, the File-Open window helps the user to choose any Map from any location of the computer, in a most user-friendly way. Organization of the modules is shown in Fig. 7.

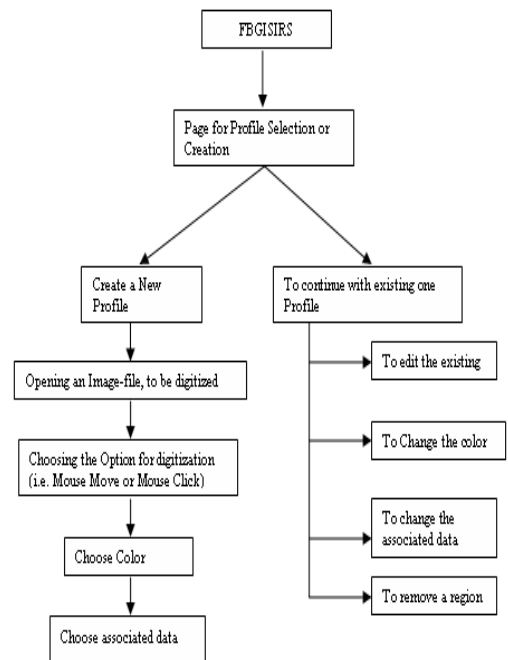


Figure. 7: Structure of the modules of the FBGISIRS

When the software runs, it opens with a profile selection and creation form as shown in Fig. 8. There are basically two choices available to the users, one is to select an existing profile from the drop down list and another one is to create a new profile keeping in mind that each profile is made to digitize only one map. To digitize a separate map a separate profile has to be created. If we want to create a new profile then by clicking on "Create New Profile" the software will ask for the name of profile. The name of the profile should be given for the map for which we want to perform digitization operation to make it self explanatory about its content. On doing so the profile will be created and it can be viewed from the drop down list. On selecting the desired profile from the drop down list, the software will automatically search for any map in the profile. If it finds existing one then it will directly move the next form but if it cannot find one then it will allow us to select a map. This map selection for a particular profile is done only once making it user friendly and easier to work on, since the user don't require the source map every time he/she wants to perform digitization. It is not necessary for the users to complete the digitization for a map at one go, it can be edited any time by simply selecting the profile where that map resides and the rest is all same.



Figure. 8: Main Page



Figure. 9: Page to assign the name to a profile

After entering the name and clicking in OK (as depicted in the screen shown in Fig. 9) the new profile will be saved and a proper message will be displayed, as given in Fig.10.



Figure. 10: Page to display successful creation of a profile

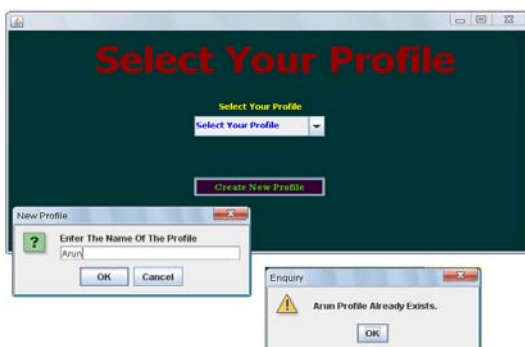


Figure. 11: Page showing error

However while creating a new profile, if the name assigned to the profile already exists then a proper message will be displayed thereby making the user aware of the existing profiles (Fig. 11).

So, while digitizing, various suitable names will have to be used, for example, while working on Indian map, we may create a profile named “Political Map of India” to store its various states, another profile named “Indian Rivers” to store its rivers’ locations and so on.

When the user select a profile from the drop down list, the software will search for any image file; if it finds one then a new form will open as shown in Fig 13. But if it can't find any then a window will open shown in Fig 12, allowing us to select an image file against which we can perform the digitization work.

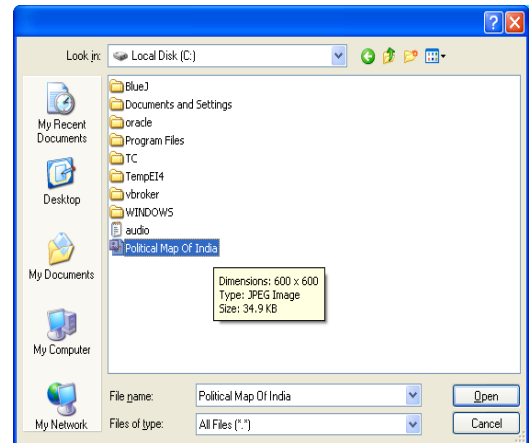


Figure. 12: File-Open Page

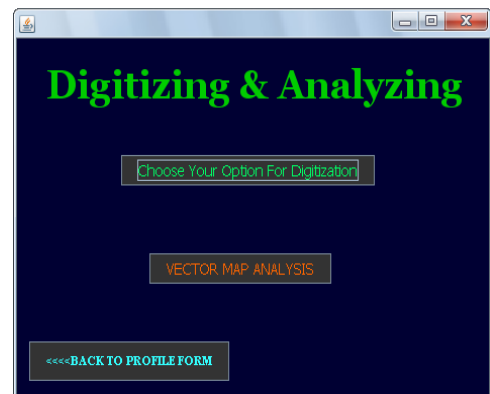


Figure. 13: Digitization option

In this page, the user has to select the option for digitization by clicking into “Choose Your Option for Digitization”. As mentioned earlier, on depending upon the current need, the user is now free to digitize the map using either of mouse-clicking or mouse-move method, keeping in mind the relative merits and demerits of the two, as discussed in section a.

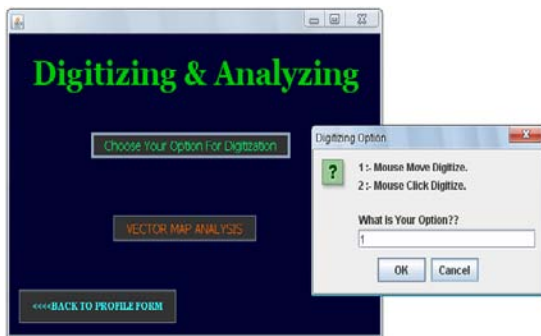


Figure. 14: Selecting the option for Digitization

The next screen which will be appearing is shown in Fig. 14. Here the user has to enter 1 as option in the text box for digitization by mouse-move. For other option, option 2 is to be selected.

Let, the political map of India (JPEG Image) be chosen for digitization, shown in Fig. 15, below.



Figure. 15: Map chosen for digitization

In mouse clicking method of digitization, first a right click and then left click on the boundary of the region which we want to digitize, is required. The right click has to be done before the start of digitization for any new region of the map and then left click again on the boundary. But if the user does not want to continue any more or there are no more regions left to digitize then after completing the last region digitization by left clicking simply the window is closed.

In mouse move method, first a right click is done and then just the mouse pointer is moved over the boundary of the region for that region. It has to be noted that the pointer of the mouse must be on the target (i.e. on the region which we want to digitize) because as soon as the right click is done all the coordinates thereafter will be considered as the boundary coordinates of the target region until a left click is made to stop the digitization for that region. But if the user does not want to continue any more or there are no more regions left to digitize then after completing the digitization of the last region simply the window is being closed.

As an illustration, here we have digitized some region of the political map of India shown in Fig 15. The resultant vector representation is shown in Fig 16.

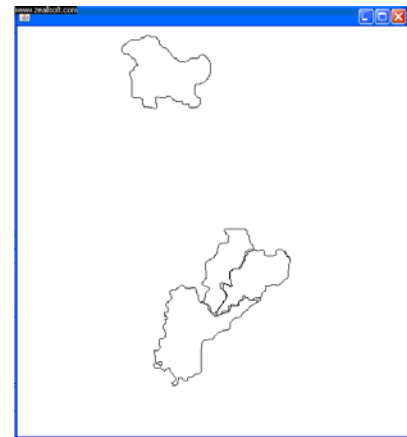


Figure. 16: After digitization

Colors are applied to the digitized regions; placing mouse-pointer inside the region, followed by a right click will open a window for choosing colors, as shown in Fig 17.



Figure. 17: Window to choose a color

After choosing the desired color and pressing “OK”, causes the region to be colored with the chosen color.

For association of the spatial values, with various regions again the same steps as above, are to be repeated. First, the mouse-pointer has to be placed inside the region and then a right-click has to be made. A window for entering the spatial values will be opened, as shown in Fig. 18.

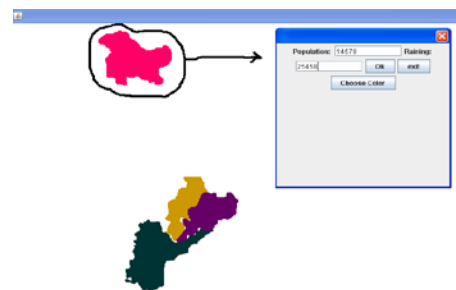


Figure. 18: Window for adding information

In Fig. 18, for the state Jammu and Kashmir (the circled region), the values for raining and population have been entered. After entering the values in the text boxes, the “OK” button is clicked and to quit from the window “EXIT” is pressed.

To display the spatial information associated with a region, the mouse-pointer is moved inside the region and a left-click is made and the information is displayed in a window as shown in Fig. 19.

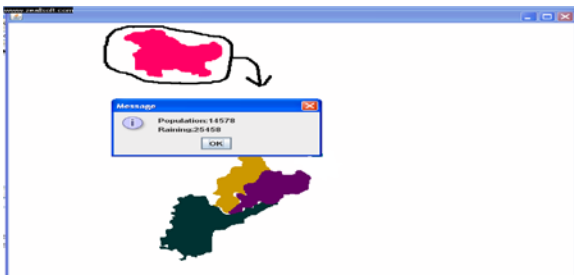


Figure. 19: Displaying the spatial data associated with a region

In Fig. 19, the mouse-pointer is placed inside the state Jammu & Kashmir and a left-click has been made, to display the associated spatial data.

To view an existing digitized map, the user first has to choose the profile (i.e. the existing digitized map) and screen in Fig. 20 is opened.



Figure. 20: To view an existing digitized map

Now if the button “VECTOR MAP ANALYSIS” is clicked, the corresponding digitized map can be viewed.

It is also easy to edit an existing map. For example if the user, while digitizing the political map of India quit after digitizing only 4 states and later if wants to digitize other states, then this step will have to be gone through.

Here from the starting page first the profile (i.e. the existing digitized map) on which editing has to done have to be chosen. Then simply the original map will get opened and editing can be done in the same way, as of digitizing a new map.

This method enables the user to digitize a portion of the map by mouse-clicking and the other portions by mouse-moving, if it is needed by the current scenario.

For deleting, we have to select a desired profile and then clicking on "Vector Map Analysis" we can view the vector map. The digitized map is shown in Fig. 21 and let the region marked with black colored circle has to be deleted.

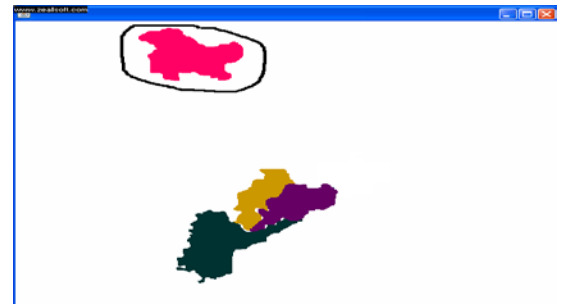


Figure. 21: Deleting a region

We have to place the mouse pointer anywhere inside the target region and then have to press “Delete” key on the keyboard, the region will be deleted along with all its associated values.

IV. ANALYSIS & COMPARISON

Various digitization techniques are available through which vector maps can be produced from the raster maps. One of such technique has been devised by Mandal, Das and Moitra [4]. In this technique the vectorization is performed by the mouse drag method. There is no temporary storage allocation processes, editing is also not possible and the space (table space) required is also more than the proposed technique. Another technique was there devised by Mandal, Halder, Mal named as GISIRS [1] , in which both mouse moving and mouse clicking methods were not there and moreover it was based on database system.

In TNTmips [5], another well-known and industrially accepted Remote Sensing and GIS software of MicroImages, the mouse click method is used to calculate the coordinates of object constructing points. But none of them has the capability to handle both of mouse move and mouse click methods. Moreover in TNTmips, no color can be assigned with the objects at the time of digitization process. The Comparisons of the proposed scheme with these three techniques are given in the table1.

Table 1: Comparison of the proposed scheme with existing techniques

Features	Proposed Technique	GISIRS [1]	DOVTG [4]	TNTmips [5]
Coordinates calculation	Based on both Mouse Move and Mouse Click	Based on mouse click	Based on mouse move	Based on mouse click
Color assignment	Color attributes may be assigned to the objects. For polygons there are boundary and fill color attributes.	Color attributes may be assigned to the objects. For polygons there are boundary and fill color attributes.	Colors are assigned. But no boundary color for the polygons	Colors can not be assigned to the objects

Databases	1. No DBMS is required only based on file systems, hence very much cost-effective. 2. For each Profile (i.e. for each digitized map one separate folder is assigned, in which different files are holding the co-ordinates, color values, attribute values and so on.)	1.Tables does not contains all the boundary point's coordinates, size of the coordinate labels are not large. 2. Color table exists	1. Since coordinate tables contains all the boundary point's coordinates the size of the coordinate labels are large. 2. Color table exists but no boundary color for polygons.	1. Tables does not contains all the boundary point's coordinates so the size of the coordinate tables are not large 2. Color table does not exists
Edit facility	Available	Available	Not available	Available

V. CONCLUSION

Proposed digitization scheme can be used to digitize and assigned color values to each object. In existing digitization scheme if large number of objects are to be digitized then the required time will be long. In this case partially digitization is also possible by temporarily storing the partially completed digitization records. After the digitization, the user can edit it by generating the vector map on the actual raster map.

Since only some boundary points which are essential to construct the absolute shape of an object are stored, this technique minimizes the size of the file, hence space required.

VI. ACKNOWLEDGEMENT

The authors express a deep sense of gratitude to the Department of Computer Science, Barrackpore Rastraguru Surendranath College, Kolkata-700 120, India and Department of Computer Science and Engineering, University of Kalyani for providing necessary infrastructural support for the work.

VII. REFERENCES

[1] Halder, P., Mandal, J.K. and Mal, S. : A GIS Anchored Information Retrieval Scheme (GISIRS) based on

Vectorization of Raster Map. International Journal of Engineering Research & Industrial Applications [ISSN-0974-1518, vol-2, No. III, pp 265-280,2009] .

[2] Halder, P., Mandal, J. K.: A GIS Based Pruning Management System of Tea Garden using Data Base Management System (GISBPMS). International Journal on Computer Science and Applications ,Chilki, Maharastra, India [ISSN- 0974-1003, vol-2, No. I, pp 09-13, 2008]

[3] Mandal, J.K., Das, S. Nath, A. and Sarkar, P.: A Pest Management System of Tea Garden Using GIS Based Database Management System. Proceedings of First International Conference on Information Technology INTL- INFOTECH 2007(ISSN-0973-6824), Haldia, Haldia Institute of Technology (2007) 261-264

[4] Mandal, J.K. Das, S. and Moitra, A.: A GIS Based Database Oriented Vectorization of Tea Garden Image for Decision Making. ACVIT- 2007, Aurangabad, Maharastra, India (2007) pp 839-846

[5] Hongye, B. : Digitization of ancient Maps Based on GIS Technology: The yu ju tu Maps. World Library & Information Congress : 75th IFLA General Conference & council, 23-27 August,2009 ,Milan, Italy.

[6] User manual of TNTmips, the MicroImages Remote Sensing and GIS software.