CONTENTS

Remote Sensing

Since humans were travelling and discovering earth, they wanted to show their findings and educate others about where they've been. This resulted in the creation of maps, which can date back until 6200 BC where the city of C¸atal H¨uy¨uk (nowadays Turkey) is depicted. Mapping advanced in Ancient Greece and Claudius Ptolemaeus with his view of the world. His map type was commonly used in the Roman times as well in the Medieval Age. In the 16th century Gerardus Mercator developed the cylindrical projection and published his world map in 1569 (figure 1). This map and projection are still used today.

Nowadays people encounter maps every day, from emergency floor plans of buildings to the Global Positioning Systems used in vehicles for transportation and localization, also used by tourists when moving from one place to another.

The process of creating maps can be roughly divided into two parts: gathering the required data and using this data to only contain the necessary information for the map. Remote sensing is defined as the acquisition of information about an object without being in physical contact with it . Although this will only cover the gathering of the data, the selection of data for display in the map is often combined under the name remote sensing as well. This is also known as classification.

Remote sensing is the science of making inferences about objects from measurements, made at a distance, without coming into physical contact with the objects under sky. Remote sensing means sensing of the earth's surface from space by making use of the properties of electromagnetic wave emitted, reflected or diffracted by the sensed objects, for the purpose of improving natural resource management, land use and the protection of environment.

Figure 1: Mercator's 1569 map of the world (Source: Thesis Joost Kuckartz)

TYPES OF REMOTE SENSING: -

WITH RESPECT TO THE TYPE OF ENERGY RESOURCES

- **Passive sensors:** They detect natural radiation that is emitted or reflected by the object or surrounding areas. Reflected sunlight is the most common source of radiation measured by passive sensors. Examples of passive remote sensors include film photography, infrared, charge-coupled devices, and radiometers.

- **Active sensors**: They emit energy in order to scan objects and areas whereupon a sensor then detects and measures the radiation that is reflected or backscattered from the target. RADAR and LiDAR are examples of active remote sensing where the time delay between emission and return is measured, establishing the location, speed and direction of an object.

WITH RESPECT TO WAVELENGTH REGIONS

Remote Sensing is classified into three types in respect to the wavelength regions

- **Visible and Reflective Infrared Remote Sensing:** The energy source used in the visible and reflective infrared remote sensing is the sun. The sun radiates electro-magnetic energy with a peak wavelength of 0.5 µm. Remote sensing data obtained in the visible and reflective infrared regions mainly depends on the reflectance of objects on the ground surface. Therefore, information about objects can be obtained from the spectral reflectance. However laser radar is exceptional because it does not use the solar energy but the laser energy of the sensor.

- **Thermal Infrared Remote Sensing**: The source of radiant energy used in thermal infrared remote sensing is the object itself, because any object with a normal temperature will emit electro-magnetic radiation with a peak at about 10 µm.

- **Microwave remote sensing:** There are two types of microwave remote sensing, passive microwave remote sensing and active remote sensing. In passive microwave remote sensing, the microwave radiation emitted from an object is detected, while the back scattering coefficient is detected in active microwave

remote sensing. Active microwave sensors are generally divided into two distinct categories: imaging and non-imaging. The most common form of imaging active microwave sensors is RADAR. RADAR is an acronym for Radio Detection and Ranging, which essentially characterizes the function and operation of a radar sensor. The sensor transmits a microwave (radio) signal towards the target and detects the backscattered portion of the signal. The strength of the backscattered signal is measured to discriminate between different targets and the time delay between the transmitted and reflected signals determines the distance (or range) to the target.

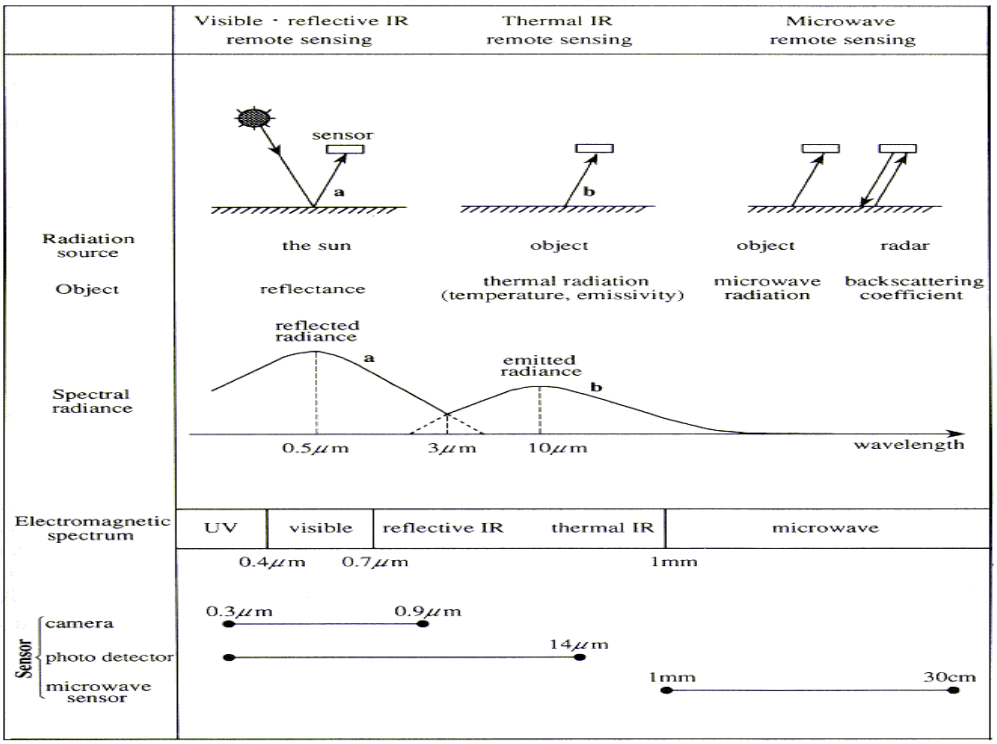


Figure 1.5.1 Three types of remote sensing with respect to wavelength regions

(Source:en.wikipedia.org/wiki/Remote_sensing)

APPLICATIONS OF REMOTE SENSING

- Geology and Mineral exploration
- Hazard assessment
- Oceanography
- Agriculture and forestry
- Land degradation
- Environmental monitoring

Gathering data

Data can be obtained using many different methods. The most common

measuring method is photography, which is possible from aircrafts or even from satellites.

The systems inside the aircraft or satellite are very similar. Both use photography techniques which will represent a measured area with a value coming from the used sensor. Combining all these measured values one can create an image of the data, and therefore execute image processing techniques.

The system measures the reflectance of the ground's electromagnetic energy using multiple sensors. The most commonly used measuring device, which can be found in many satellites, is the pushbroom scanner . The array of charge coupled devices (CCDs) are capable of measuring parts of the electromagnetic spectrum by converting the received energy in electrons, which are then counted and represented by a digital value . Multiple scanners are present in satellites, for which every scanner has its own CCD array, each sensitive for another part of the electromagnetic spectrum. The size of the measured part of the spectrum is known as the spectral resolution: the narrower the bands, the greater the spectral resolution.

Because of the height of the sensor above ground and the optic system present, one sensor measures a large area on the ground. The area D of the measured area on the ground depends on the instantaneous field of view (IFOV), the opening angle of a remote sensing detector β and the flying height H: $D = \beta H$. The IFOV is expressed in radians, the flying height and ground diameter in meters. The flying height of satellites is often 600 km or higher , so very good optical systems are built to create a small ground diameter. The area on the ground that an imaging system can distinguish is also known as the spatial resolution . The smaller the area on the ground, the higher the spatial resolution. Every measured area will be stored as a square pixel with a representative value which later can be used through image processing.

A satellite is an artificial body placed in orbit around the earth or another planet in order to collect information or for communication. There are thousands of satellites taking pictures of planets, sun and other objects. These pictures help scientists to learn about the planets and the solar system.

India has launched 75 satellites (till July 2014). The first satellite launched by Indian Space Research Organisation was Aryabhata on 19 April 1975. Some of the satellites launched by ISRO are-

Satellite	Launch Date	Description
Bhaskara-I	7 June 1979	First experimental remote sensing satellite. Carried TV and microwave cameras.
IRS-1A	17 March 1988	Earth observation satellite. operational remote sensing satellite.

Kalpana-1	12 September 2002	First meteorological satellite built by ISRO. Originally named METSAT. Named after Kalpana Chawla.
GSAT-10	29 September 2012	India's advanced communication satellite. high power satellite being inducted into the INSAT system.
GSAT-7	30 August 2013	This is an advanced multi-ba communication satellite dedicated for milita use.

Classification

Every pixel of an image contains the reflectance information of a particular area of the electromagnetic spectrum. Combining several images (also called bands), the same area measured by sensors sensitive for different areas of the spectrum, one can generate a realistic looking photograph of the area. From a photograph (which are bands red, green and blue) one can easily see where forests are located, where cities are and which locations are agricultural. With more bands combined (IR, UV, etcetera), other information can be extracted, for instance vegetation types and thermal measurements .

Classification is a method for determining where a measured area belongs to, so which class can be assigned to a measured pixel. The above mentioned examples are part of the most used application in remote sensing: landcover mapping. Some examples of classes are grass, trees, built-up and bare soil, but many more can be thought of. Every class can even be expanded to many more classes for each type of vegetation.

If a user looks at a single pixel, it's very hard to determine to which class it belongs to. Looking at a larger area people can determine the class by searching for familiar aspects like patterns and colors. Using this method, classification can only done by hand, which will be a time consuming task for large images. To use the aid of a computer in classification, two methods are available, supervised and unsupervised classification.

Supervised classification procedure involves three steps-

a)the training stage, wherein the multi-spectral parameters are extracted for various classes from the training sites identified in the image

b)the classification stage, where each pixel is assigned to a class to which it most probably belongs

c)the output stage is the stage where the data is presented in the form of maps, tables, graphs, etc. This stage is common for unsupervised classification also.

With unsupervised classification, it's the computer only defining which pixel belongs to which class. It can do this to look at the distribution of pixels among the different bands: it looks at pixel agglomerations in graphs where two bands are plotted as function of each other. Such type of graph is called a feature space .

Feature space plots are also used to determine the class separability, the greater the class separability the easier it is to determine to which class pixels belong. The measure of class separability is a value from 0 to 2.

In some programs extra information is present in classification, namely the electromagnetic spectrum distributions of several land cover types. Using these distributions together with unsupervised classification can result in even better classified images.

Mathematically every classification method used can be split up in two types: hard and soft classification. With hard classification each unit is assigned unambiguously to a single class while with soft classification multiple classes can be present in each unit and probabilities of assigning a class will be relevant . The conventional methods use hard classification based on the percentages of each class found within the pixel. Super resolution mapping mostly uses soft classification, as there is a probability with the spatial location involved.

Change Detection

Change detection is the process that helps in determining the changes associated with landuse and land cover properties with reference to geo-registered multitemporal remote sensing data. It helps in identifying change between two (or more) dates that is uncharacterised of normal variation. Change detection is useful in many applications such as landuse changes, habitat fragmentation, rate of deforestation, coastal change, urban sprawl and other cumulative changes through spatial and temporal analysis techniques such as GIS and Remote sensing along with digital image processing techniques.

GIS is the systematic introduction of numerous different disciplinary spatial and statistical data, that can be used in observing the change, constituent processes and prediction based on current practices and management plans. Remote sensing helps in acquiring multi spectral spatial and temporal data through space borne remote sensors. Image processing technique helps in analyzing the dynamic changes associated the earth resources such as land and water using remote sensing data. Successful utilization of remotely sensed data for land cover and land use change detection depends on selection of appropriate data set.

Markov Random Field

In the domain of physics and probability, a Markov random field (often abbreviated as MRF), Markov network or undirected graphical model is a set of random variables having a Markov property described by an undirected graph. A Markov random field is similar to a Bayesian network in its representation of dependencies; the differences being that Bayesian networks are directed and acyclic, whereas Markov networks are undirected and may be cyclic. Thus, a Markov network can represent certain dependencies that a Bayesian network

cannot (such as cyclic dependencies); on the other hand, it can't represent certain dependencies that a Bayesian network can (such as induced dependencies).

When the probability distribution is strictly positive, it is also referred to as a Gibbs random field, because, according to the Hammersley–Clifford theorem, it can then be represented by a Gibbs measure. The prototypical Markov random field is the Ising model; indeed, the Markov random field was introduced as the general setting for the Ising model. In the domain of artificial intelligence, a Markov random field is used to model various low- to mid-level tasks in image processing and computer vision. For example, MRFs are used for image restoration, image completion, segmentation, image registration, texture synthesis, super-resolution, stereo matching and information retrieval.

Given an undirected graph $G = (V, E)$, a set of random variables $X = (X_v)_{v \in V}$ indexed by V form a Markov random field with respect to G if they satisfy the local Markov properties:

Pairwise Markov property: Any two non-adjacent variables are conditionally independent given all other variables:

$$X_u \perp\!\!\!\perp X_v \mid X_{V \setminus \{u,v\}} \quad \text{if } \{u, v\} \notin E$$

Local Markov property: A variable is conditionally independent of all other variables given its neighbors:

$$X_v \perp\!\!\!\perp X_{V \setminus \mathrm{cl}(v)} \mid X_{\mathrm{ne}(v)}$$

where $\mathrm{ne}(v)$ is the set of neighbors of v, and $\mathrm{cl}(v) = \{v\} \cup \mathrm{ne}(v)$ is the closed neighbourhood of v.

Global Markov property: Any two subsets of variables are conditionally independent given a separating subset:

$$X_A \perp\!\!\!\perp X_B \mid X_S$$

where every path from a node in A to a node in B passes through S.

The above three Markov properties are not equivalent to each other at all. In fact, the Local Markov property is stronger than the Pairwise one, while weaker than the Global one.

As the Markov properties of an arbitrary probability distribution can be difficult to establish, a commonly used class of Markov random fields are those that can be factorized according to the cliques of the graph.

Given a set of random variables $X = (X_v)_{v \in V}$, let $P(X = x)$ be the probability of a particular field configuration x in X. That is, $P(X = x)$ is the probability of finding that the random variables X take on the particular value x. Because X is a set, the probability of x should be understood to be taken with respect to a *joint distribution* of the X_v.

If this joint density can be factorized over the cliques of G:

then X forms a Markov random field with respect to G. Here, cl(G) is the set of cliques of G. The definition is equivalent if only maximal cliques are used. The functions φ_C are sometimes referred to as *factor potentials* or *clique potentials*. Note, however, conflicting terminology is in use: the word *potential* is often applied to the logarithm of φ_C. This is because, in statistical mechanics, $\log(\varphi_C)$ has a direct interpretation as the potential energy of a configuration x_C.

Although some MRFs do not factorize (a simple example can be constructed on a cycle of 4 nodes[3]), in certain cases they can be shown to be equivalent conditions:

1)if the density is positive (by the Hammersley–Clifford theorem),

2)if the graph is chordal (by equivalence to a Bayesian network).

When such a factorization does exist, it is possible to construct a factor graph for the network.

Predictive Analysis

A Hidden Markov Model describes the joint probability of a collection of 'hidden' and observed discrete random variables. It relies on the assumption that the i^{th} hidden variable given the $(i-1)^{th}$ hidden variable is independent of previous hidden variables, and the current observation variables depend only on the current hidden state.
The Baum–Welch algorithm uses the well known EM algorithm to find the maximum likelihood estimate of the parameters of a hidden Markov model given a set of observed feature vectors.

Let X_t be a discrete hidden random variable with N possible values. We assume the $P(X_t|X_{t-1})$ is independent of time t, which leads to the definition of the time independent stochastic transition matrix

$$A = \{a_{ij}\} = P(X_t = j|X_{t-1} = i)$$

.

The initial state distribution (i.e. when t=1) is given by

$$b_j(y_t) = P(Y_t = y_t|X_t = j)$$

.

The observation variables Y_t can take one of K possible values. The probability of a certain observation at time t for state j is given by

$$b_j(y_t) = P(Y_t = y_t|X_t = j)$$

.

Taking into account all the possible values of Y_t and X_t we obtain the K by N matrix $B = \{b_j(y_i)\}$.

An observation sequence is given by $Y = (Y_1 = y_1, Y_2 = y_2, ..., Y_t = y_t)$.

Thus we can describe a hidden Markov chain by $\theta = (A, B, \pi)$. The Baum–Welch algorithm finds a local maximum for $\theta^* = \max_\theta P(Y|\theta)$. (i.e. the HMM parameters that maximise the probability of the observation.)[5]

Algorithm

Set $\theta = (A, B, \pi)$ with random initial conditions. They can also be set using prior information about the parameters if it is available.

Forward procedure

Let $\alpha_i(t) = P(Y_1 = y_1, ..., Y_t = y_t, X_t = i|\theta)$, the probability of seeing the $y_1, y_2, ..., y_t$ and being in state i at time t. This is found recursively:

1. $\alpha_i(1) = \pi_i b_i(y_1)$
2. $\alpha_j(t+1) = b_j(y_{t+1}) \sum_{i=1}^{N} \alpha_i(t) a_{ij}$

Backward procedure

Let $\beta_i(t) = P(Y_{t+1} = y_{t+1}, ..., Y_T = y_T | X_t = i, \theta)$ that is the probability of the ending partial sequence $y_{t+1}, ..., y_T$ given starting state i at time t. We calculate $\beta_i(t)$ as,

1. $\beta_i(T) = 1$
2. $\beta_i(t) = \sum_{j=1}^{N} \beta_j(t+1) a_{ij} b_j(y_{t+1})$

Super Resolution Mapping

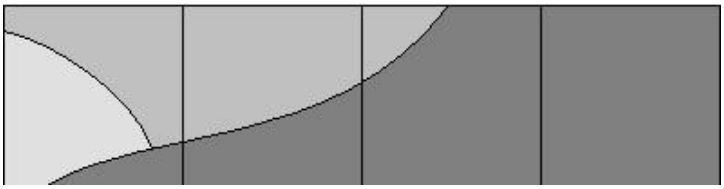
Introduction

Very detailed images, from recent satellites but more common aircraft measurements, have a high spatial resolution and therefore it's less common one pixel contains several landcover classes. Unfortunately these images are very expensive, and for measuring larger areas many images need to be bought. The images from older satellites are less expensive, but have a low spatial resolution. One pixel within these images often contains several landcover classes.

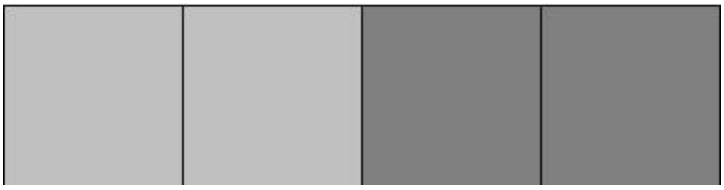
As common programs use hard classification, this pixel will get assigned to only one class . With super resolution mapping this pixel will be subdivided, where every subpixel can be assigned its own class. This will result in an image with a higher spatial resolution than the one obtained, but mathematical optimization has to be done, which will give a result with a certain efficiency.

Principle

Suppose a satellite image of the Landsat-7 satellite is available for land-cover class mapping. This particular satellite can measure 8 bands of the electromagnetic spectrum , so for every pixel there are 8 different values available. Of the 6 interesting bands for landcover mapping the CCD sensors are capable of measuring the ground with a spatial resolution of 30 meters. Within these 30 meters, a residential area could be present, but also a piece of farmland, or a piece of forest (figure 2(a)).



(a) Schematic view of an original measured area



(b) Landcover mapping using current techniques

(c) Simple subpixel mapping

Figure 2: Landcover class mapping(Source: Thesis Joost Kuckartz)

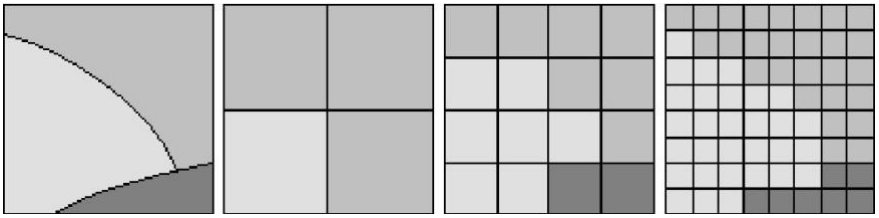
Because of hard classification every pixel will be assigned one class, so in this case an area of 30 meters will have one class. Using all the spectral band images the classification algorithm chooses the class best fitted to the pixel as a whole. This often results in a pixel showing the landcover class which is mostly

represented within the measured area. In figure 2(b) the result of this method is shown.

An improved method would be to include parts of the classes present within the single measured pixel. This can only be achieved by subpixel classification, also referred to as super resolution mapping, of which a simple possible result is shown in figure 2(c). With subpixel classification first the amount of the present classes within the pixel needs to be determined. Because every class has its own specific spectral representation, pure pixels, also called training pixels, can be used to determine which classes and the amount of classes present in a mixed pixel, and the covered percentage of that class. These pure pixels are an important measure in both conventional mapping and super resolution mapping, because they only cover one class thus the mean spectral reflectance and its variance for that particular class is known. Then if for every class a pure pixel can be found, the percentages of class availability within other pixels can be calculated.

With the available data, also coming from the earlier described classification methods, the program then looks for the pixels closest to the given spectral reflectance representation. With these pure pixels for each class, the classes and class fractions of the other pixels can be determined. This mathematical method has been researched often , but is outside the scope of this research. Because the commonly used mapping techniques don't use super resolution mapping, the original image will result in a map, showing the highest percentage of the found class type at each pixel.

For super resolution mapping the choice for the amount of subpixels needs to be made. The more subpixels are chosen, depending on the scaling factor, the closer a depiction of reality the image will be. The larger the scale factor will be, the more storage space and computation power will be required. Figure 3 shows the potential of the subpixel classification theory.



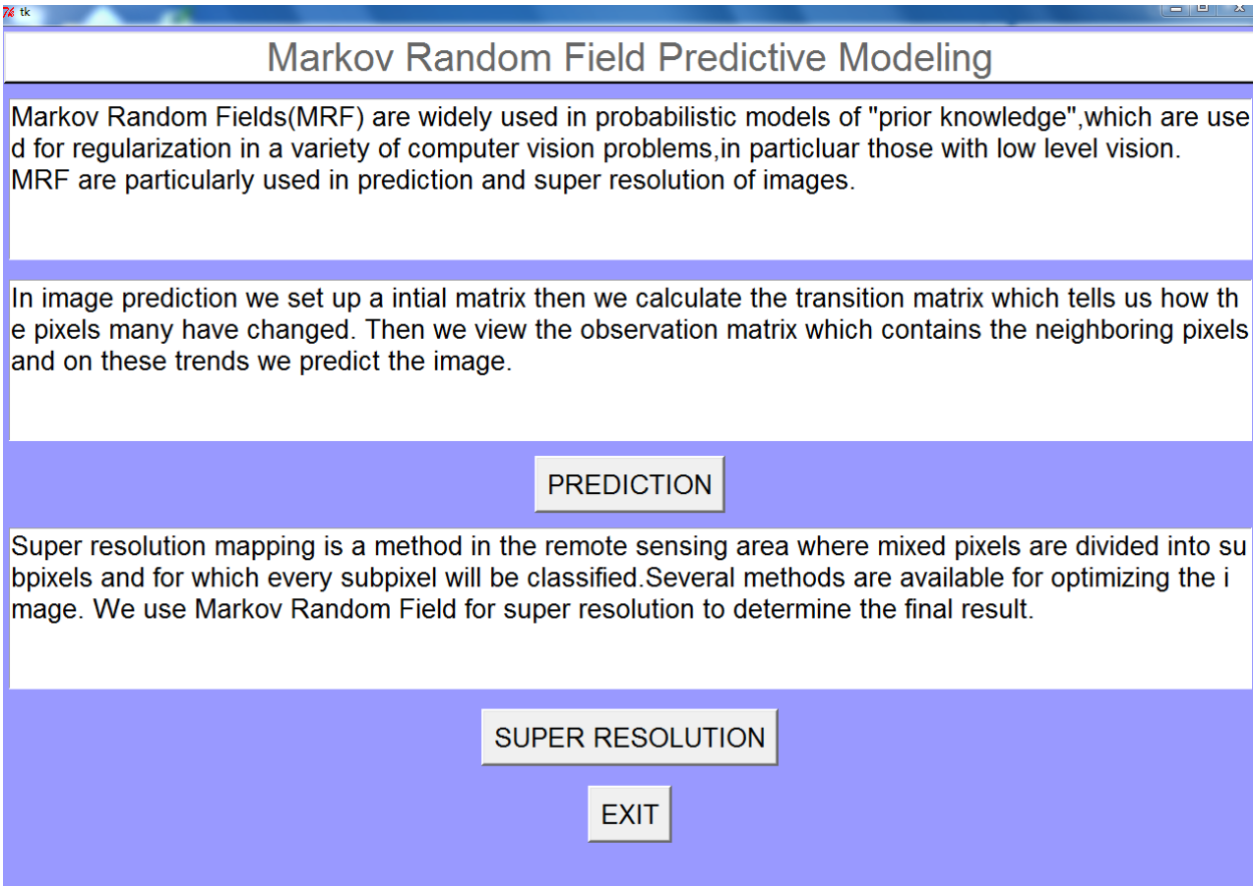
(a) Schematic view of original measured area (b) Subpixel map- -ping with scale 2 (c) Subpixel map- -ping with scale 4 (d) Subpixel map- -ping with scale 8

Figure 3: Subpixel class mapping(Source: Thesis Joost Kuckartz)

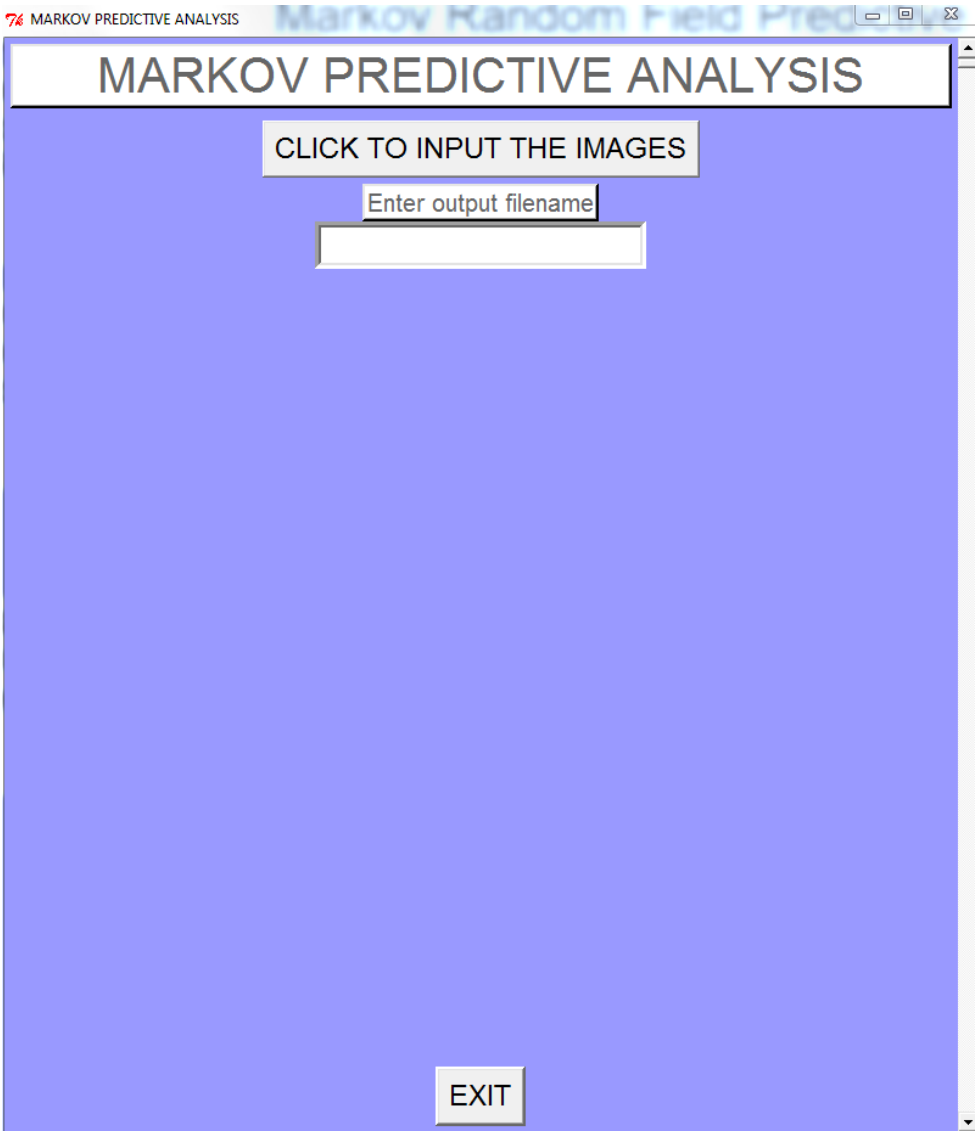
Software Features

The software was developed to test the potential of Markov Random Field in Change Prediction and Super Resolution Mapping.

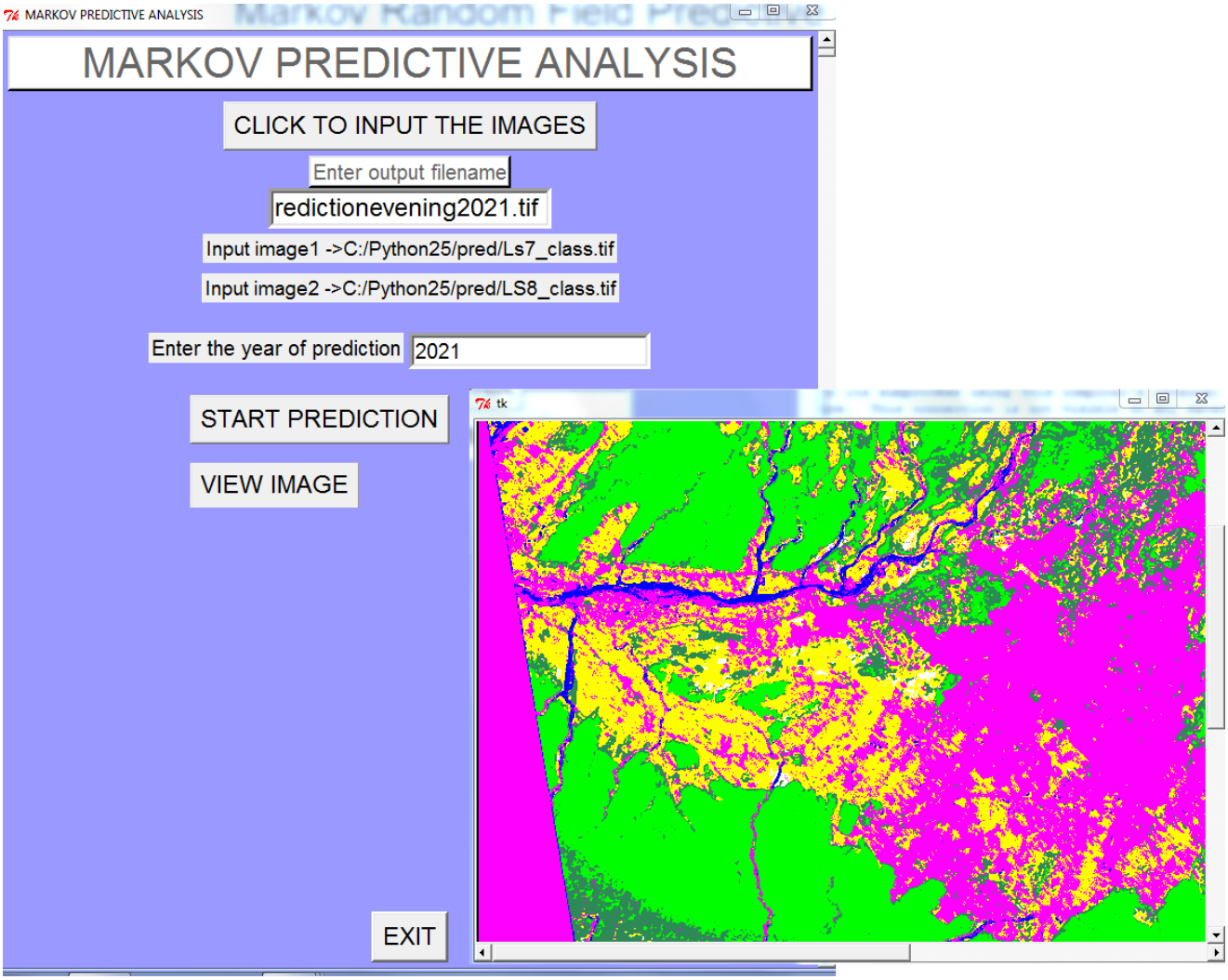
It was developed in Python programming language.MRF is used for two features i.e.Prediction and Super Resolution. Both of these were implemented in the software.The user can select either of the two option as per his choice.



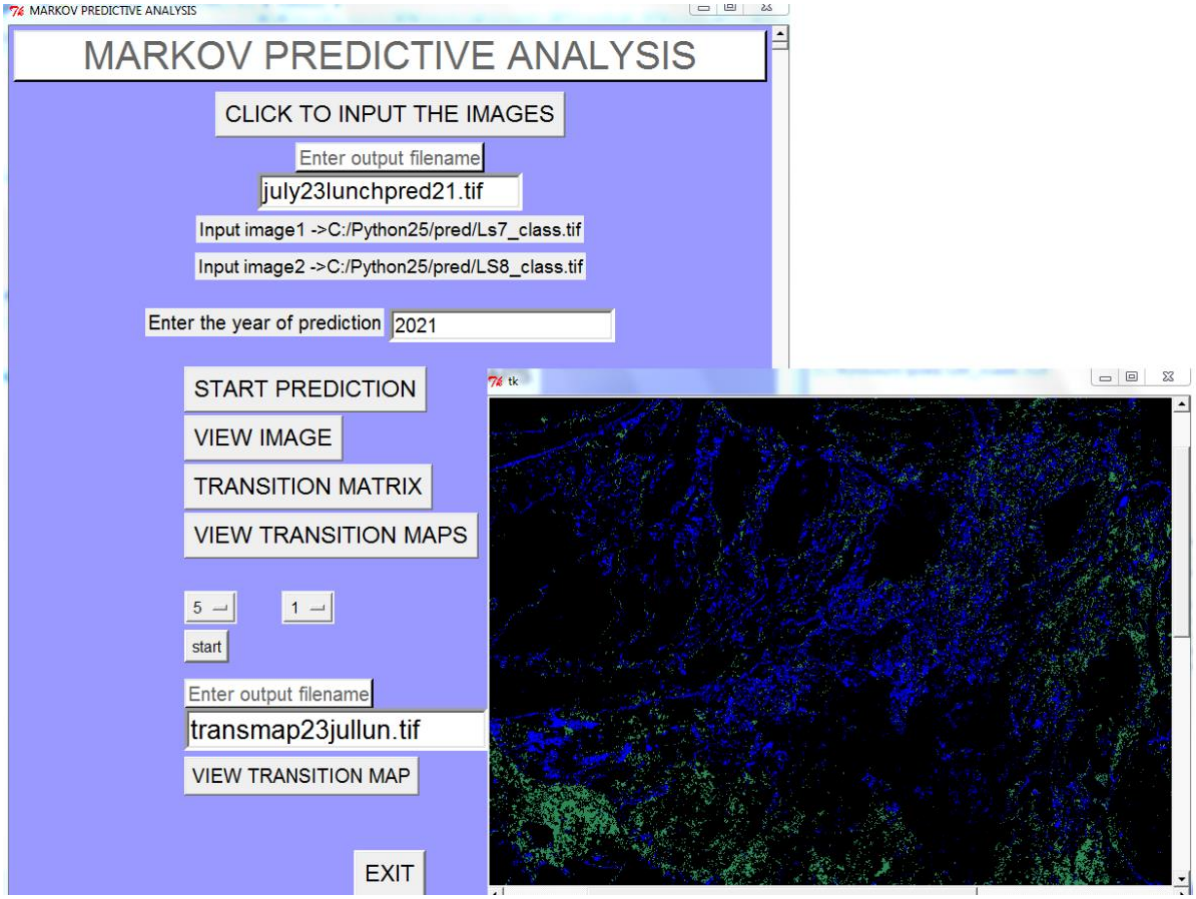
On clicking the "PREDICTION" button a new window opens providing an interface for prediction. The user has to input two images and also provide the name of the output file.

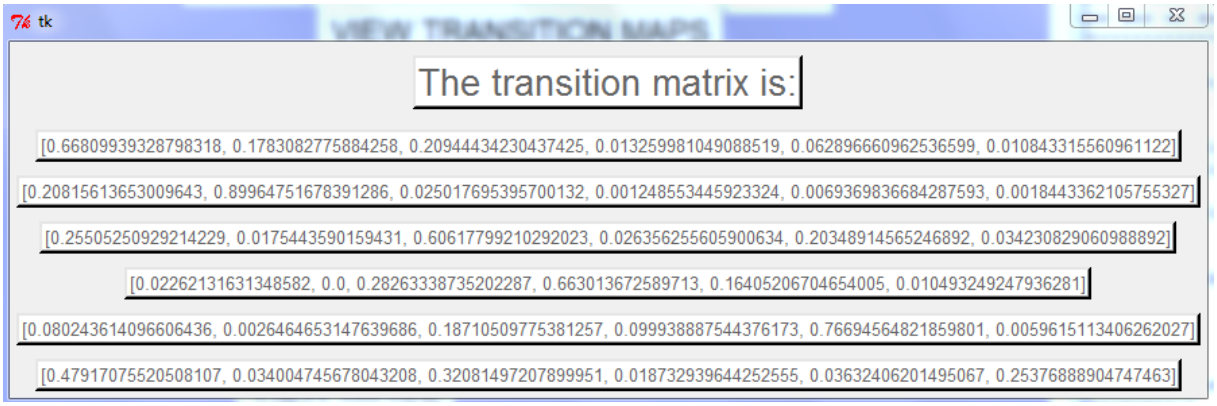

MARKOV PREDICTIVE ANALYSIS
MARKOV PREDICTIVE ANALYSIS
CLICK TO INPUT THE IMAGES
Enter output filename
EXIT

The user also has to supply the year of prediction. When the process is finished "View Image" button appears on the screen and the user can view the image by clicking on that button.

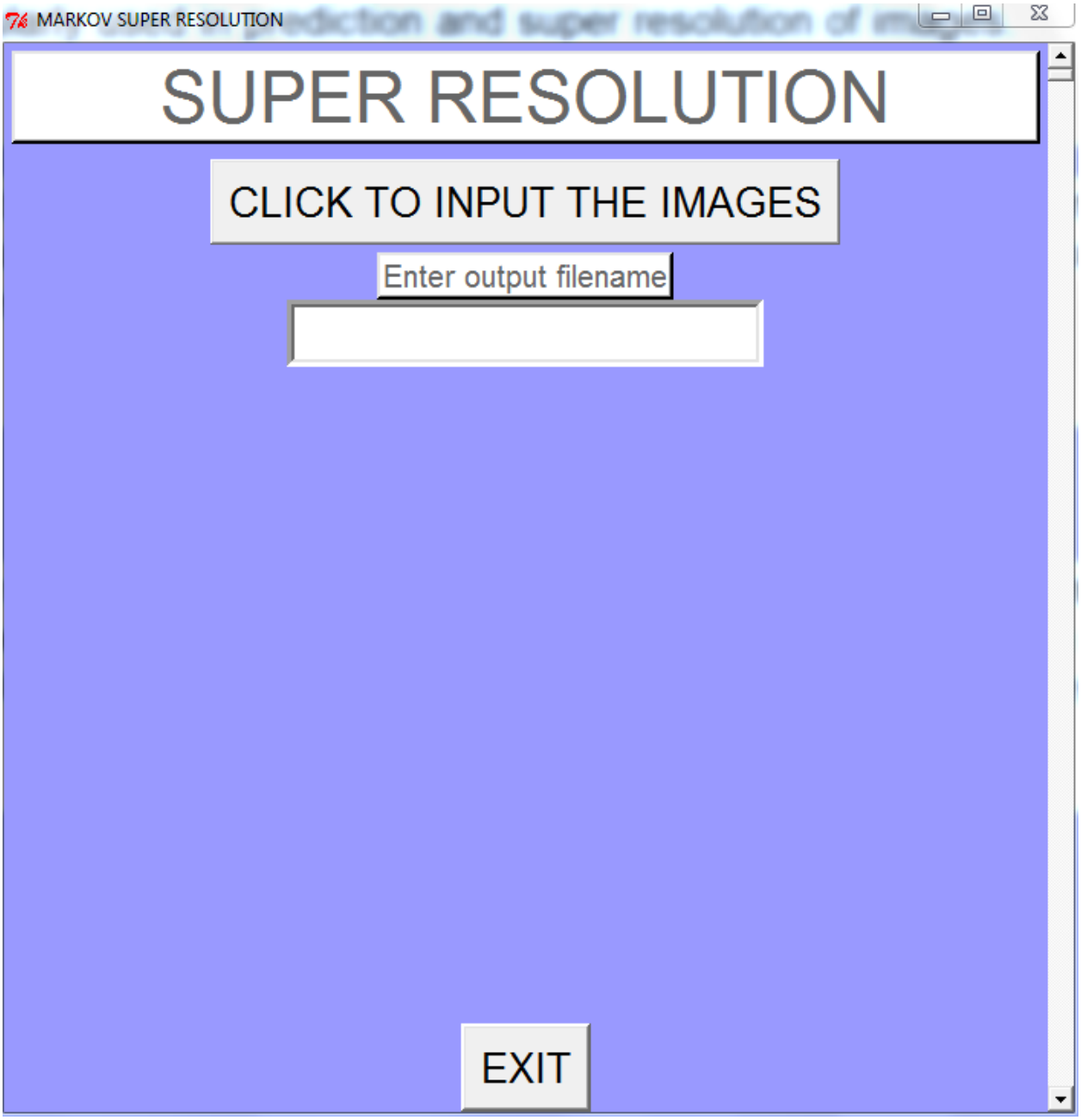


There is also a feature to view the transition matrix and transition map.To view the transition map legends have to be selected from the drop down menu and the name of output file has to be supplied.





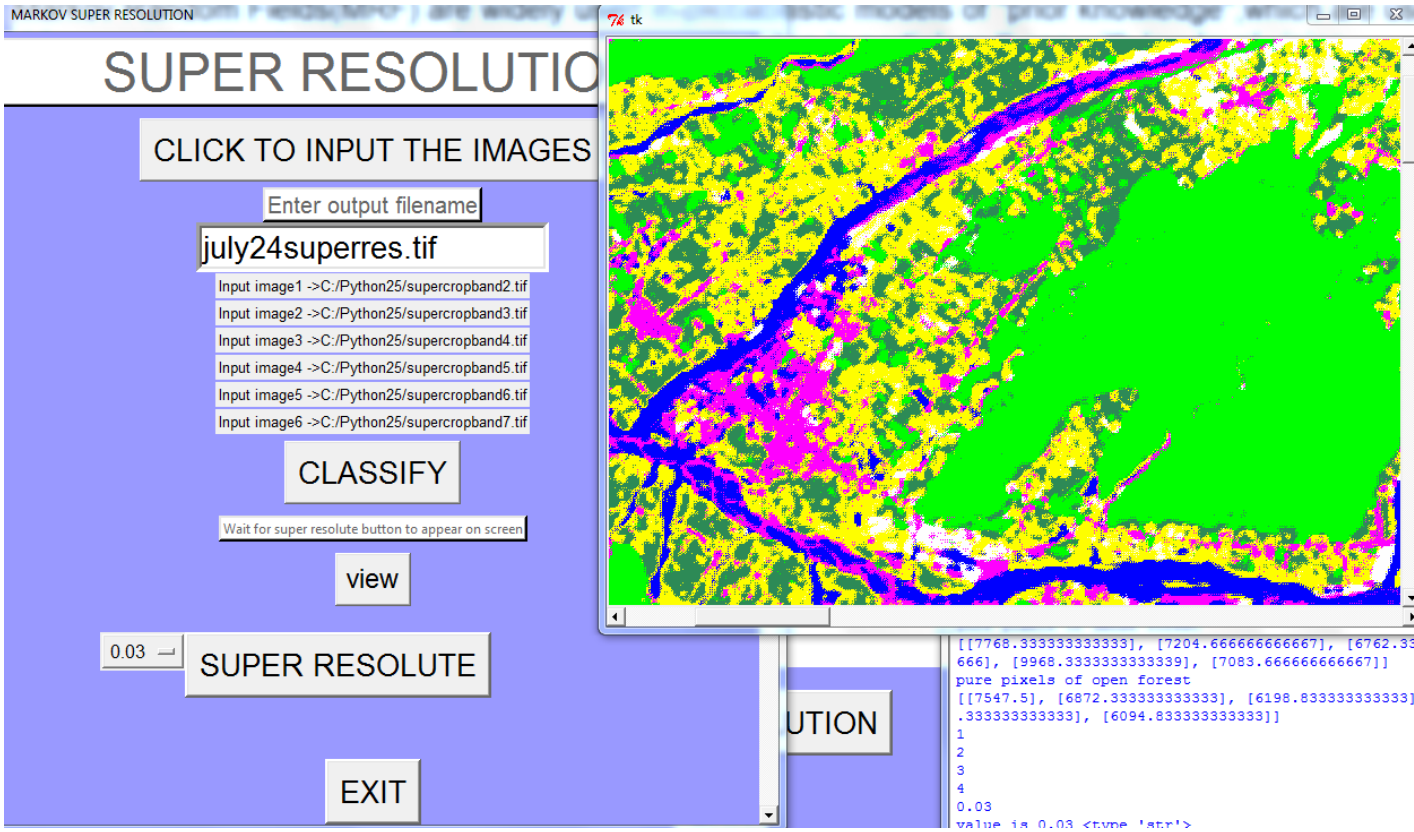
On clicking the "SUPER RESOLUTION" button from the main menu a new window opens providing an interface for super resolution. The user has to input the images and also give the name of the output file.



Once the user inputs the images(6 band image is inputted), the name of the selected files is displayed on the screen. The input images are raw data of the satellite image so they need to be unmixed and this process is achieved by clicking on the CLASSIFY button.


MARKOV SUPER RESOLUTION
SUPER RESOLUTION
CLICK TO INPUT THE IMAGES
Enter output filename
july22evensuperres.tif
Input image1 ->C:/Python25/LC81460392014024LGN00_B2.TIF
Input image2 ->C:/Python25/LC81460392014024LGN00_B3.TIF
Input image3 ->C:/Python25/LC81460392014024LGN00_B4.TIF
Input image4 ->C:/Python25/LC81460392014024LGN00_B5.TIF
Input image5 ->C:/Python25/LC81460392014024LGN00_B6.TIF
Input image6 ->C:/Python25/LC81460392014024LGN00_B7.TIF
CLASSIFY

Once the input images have been classified they can then be super resoluted. A drop down menu is provided to select the similarity between the two classes and "SUPER RESOLUTE" button is clicked. Once the super resolution is completed we can view the image by clicking on the view button. Different portions of the image can be viewed by dragging the vertical and horizontal scrollbars.



REFERENCES

[1] P.M. Atkinson. Super-resolution land cover classification using the two-point histogram. Geostatistics for Environmental Applications, pages 15–28, 2004.

[2] P.M. Atkinson, M.E.J. Cutler, and H. Lewis. Mapping sub-pixel propor-tional land cover with avhrr imagery. International Journal of Remote Sensing, 18:917–935, 1997.

[3] G.M. Foody. Hard and soft classifications by a neural network with a non-exhaustively defined set of classes. International Journal of Remote Sensing, 23:3853–3864, 2002.

[4] T. Kasetkasem, Arorab, M.K., and P.K. Varshney. Super-resolution land cover mapping using a markov random field based approach. Re-mote Sensing of Environment, 96:302–314, 2005.

[5] R.H. Kassaye. Suitability of markov random field-based method for super-resolution land cover mapping. http://www.itc.nl/library/ papers_2006/msc/gfm/hailu.pdf, 2006

[6] Python Imaging Library(PIL)

[7] Lawrence R.Rabiner. A tutorial on hidden markov models and selected application in speech recognition

[8] How to think like computer scientist-Think Python.

[9] www.wikipedia.org

[10] T.V.Ramachandra, Uttam Kumar. GRDSS for land use land cover dynamics analysis