Scientific Journal of Impact Factor(SJIF): 3.134

e-ISSN(O): 2348-4470
p-ISSN(P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 1,Issue 12, December -2014

PREDICTION OF RAINFALL USING IMAGE PROCESSING

Ankita Panchani[1], Hardik Doshi[2], Nivid Limbasiya[3]

[1]Computer Engineering, Darshan Engineering College, anki.panchani@gmail.com
[2] Asst. Prof., Computer Engineering, Darshan Engineering College, hadoshi@gmail.com
[3]Asst. Prof., Computer Engineering, V.V.P Engineering College,nlimbasiya24@gmail.com

Abstract — *The onset of monsoon is eagerly awaited in the Indian sub-continents as it has deep impact in the economic & social domain. Considering the cost factor & security issues, it is now possible to use digital cloud imagery to monitor the different parameters which affect or get affected by the monsoon in a more global scale. Digital cloud images are collected & stored as JPEG files and are used to recognize sky status, cloud status & cloud type using contourlet or pca, Cloud screening algorithm & image segmentation method respectively. After that, information about the cloud that is height, altitude, classification & appearance is provided which helps in finding the status of rainfall. Previous research data about the basic characters of cloud like shape, color, texture, edge are used to recognize the status of rainfall. Common people can get the desired detail regarding status of rainfall by just taking a photograph utilizing this technique.*

Keyword- *Cloud, Image, Rainfall, File, PCA, Sky*

I. INTRODUCTION

The main aim of the research is to use digital cloud image for estimating rainfall and detect the type of clouds using different image processing techniques. As water is inevitable part of life and rainfall is backbone for every nation's prosperity and economic scenario. Rainfall play vital role in our country's social and economic development. It is very essential to predict rainfall and estimate how much rain would occur. Rainfall would be detected from the type of cloud present and this is our aim to detect type of cloud for estimating rainfall. As, more than of 50% of globe surface is covered with cloud which play important role in hydrological cycle and energy balance of atmosphere-earth surface because of the interaction of solar radiation and terrestrial radiation. Here in this research we are going to use digital image for rainfall prediction. As digital image play essential role in many application in this fast moving Digital Universe. In most of the cloud related research requires cloud observation which includes amount and types of cloud in the sky. f's ever-changing distribution is one of the most interesting features of Earth. In our daily lives, clouds are encountered as natural as anything in atmosphere. Second thought is hardly given to their presence, as they float above us. Important role played by rainfall in hydrological cycle by providing water to the earth surface. Agriculture is totally dependent on rain and also it provides water to streams, which is important for aquatic creature. Excess of rainfall is hazardous as it causes flooding, which results in destruction of both life and property. So, rainfall have important role in many aspect of life, it is not worthy to observe where it occurred and how much it fallen but it is important to have pre-knowledge about rainfall so that proper management could be possible and prevent from destruction. Hence rainfall forecast is much necessary.

A.Introduction of Image Processing

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually Image Processing system includes treating images as two dimensional signals while applying already set signal processing methods to them.

It is among the most rapidly growing technologies, with its application in various aspects like medical, image retrieval, business etc. it is core research area in the engineering field too[1].

II. TYPES OF METHOD USED IN IMAGE PROCESSING

There are two type of methods used in image processing i.e. analog and digital image processing. In analog image processing or visual techniques to image processing can be used for the hard copies like printouts and photographs. Image analyst uses various fundamentals of interpretation while using these visual techniques. The image processing is not just confined to area that has to be studied but on knowledge of analyst. Association is another important tool in image processing through visual techniques. So analysts apply a combination of personal knowledge and collateral data to image processing [2].

Digital Image Processing techniques help in manipulation of the digital images by using computers. As raw data from imaging sensors from satellite platform contains deficiencies. To get over such flaws and to get originality of information, it has to undergo various phases of processing. The three general phases that all types of data have to undergo while using digital technique are Pre- processing, enhancement and display, information extraction.

@IJAERD-2014, All rights Reserved

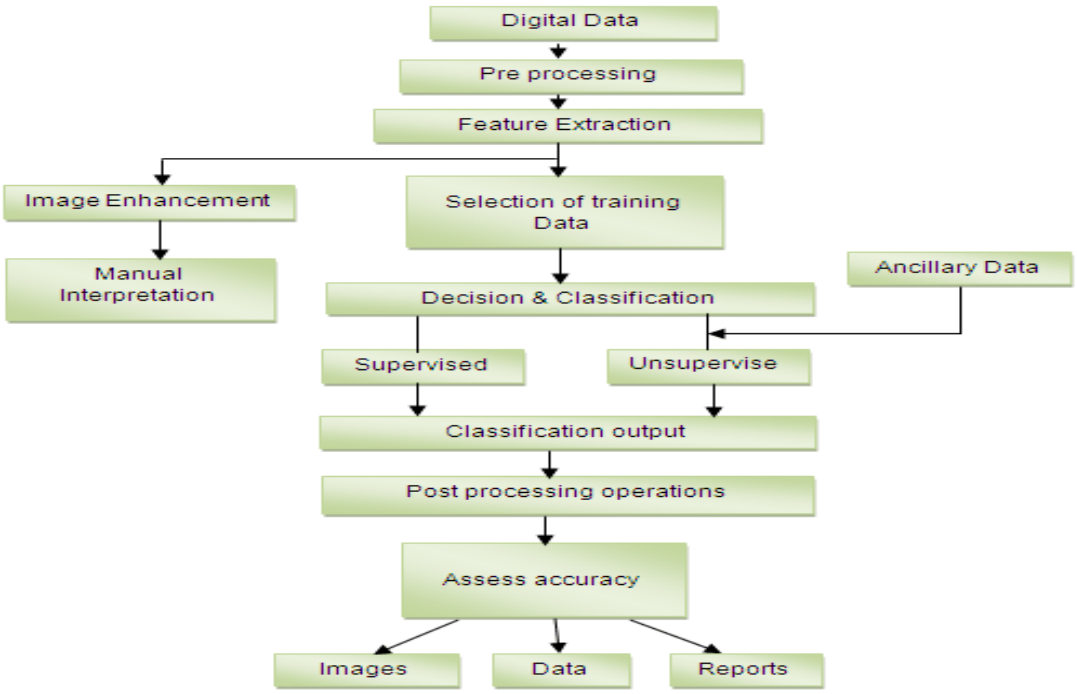

Figure1.Data flow diagram of digital image processing [2]

III. PROCESSING TECHNIQUE ON IMAGE

A.Image enhancement:

It refers to sharpening or blurring, of image features like contrast or boundaries to make graphic display more visible for display & analysis. Process of enhancement does not increase the inside information in data. That includes contrast manipulation and gray level, noise reduction, filtering, interpolation, edge sharpening and magnification.

B.Image restoration:

It is mostly combined with filtering to observe image to minimize effect of noise or degradations. Effect of image restoration is depended on accuracy and extent of the knowledge of on filter design and process of degradation. The difference between Image restoration and image enhancement is concerned with more accentuation or extraction of image features [3].

C.Image compression:

It is associated with minimizing number of bits that are required to represent an image in digital form. Use of compression are in different places like broadcast TV, military communication via aircraft, remote sensing via satellite, medical images in computer tomography, radar, facsimile transmission, teleconferencing, for educational or business documents, magnetic resonance imaging and also pictures, satellite images, digital radiology, motion, geological surveys, weather maps and so on.

IV. COMMON FILE FORMATS

Common file formats that use mostly are: JPEG that is a very efficient and much information per byte compressed 24 bit format that widely used in web and Internet.GIF that is 8-bit or 256 color and non-destructive compressed format that is mostly used for web.TIFF that is standard 24 bit publication format that compresses non-destructively.PSD that is a dedicated Photoshop format which keeps including all the layers all the information in the image.PS that is a standard vector format that has numerous sub-standards that could be difficult to transmit across platforms[4].

V. CLOUDS AND CLOUD TYPE

Water on land or in the ocean evaporates - turns from a liquid to water vapor - it rises. The water vapor cools and turns back into a liquid in the shape of tiny droplets. The result is clouds, unless it's on the ground - then we call it fog. When enough droplets get together they fall to the ground as rain or if it's very cold, they freeze and fall down as snow, sleet or hail. Clouds play an important role in the energy balance of Earth. They cool the earth by reflecting sunlight back out to space. It's a never-ending cycle but one that keeps the earth balanced. The three different types of clouds are high clouds (because they are high in the sky), medium clouds and low clouds (just over 10,000 feet high or 3 km.) [5].

@IJAERD-2014, All rights Reserved

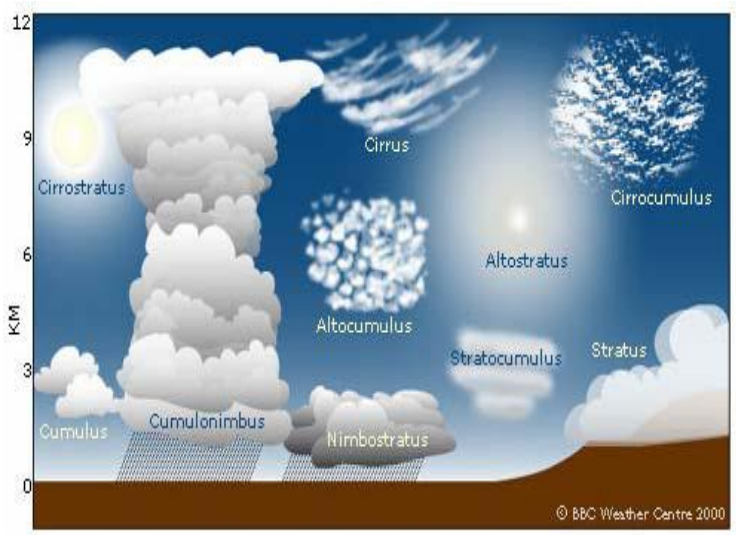

Figure 2. Different types of cloud

A.High Clouds

Cirrus (Ci): The ice-crystal cloud is a feathery white cloud that is the highest in the sky. It has a wispy looking tail that streaks across the sky and is called a fall streak. Cirrostratus (Cs): A milky sheet of ice-crystal cloud that spreads across the sky.

B.Low clouds

Cumulus (Cu): The fair weather cloud. Each cloud looks like a cauliflower. Stratocumulus (Sc): A layer of cloud that can sometimes block the sun.

C.Rain Clouds

Nimbostratus (Ns): The dark, rain carrying cloud of bad weather. It is to blame for most of the winter rains and some of the summer ones. It covers the sky and blocks the sun [6].

VI. FEATURE EXTRACTION TECHNIQUE

Feature extraction involves obtaining relevant cloud features from the cloud images. This feature may be intensity, brightness, temperature etc. This phase has application like decrementing cloud and sky or defining sky status. Finally, the system recognizes the sky status and cloud status effectively.

A.Independent component analysis

Independent Component Analysis (ICA) has emerged recently as one powerful solution to the problem of blind source separation while its possible use for face source components from their linear mixtures (the observables). ICA thus provides a more powerful data representation than PCA as its goal is that of providing an independent rather than uncorrelated image decomposition and representation.

$$X = AS$$

Where A denotes the mixing matrix, S denotes the source matrix containing statistically independent source vectors in its rows and X denotes the data matrix. In the ICA method, the only information we possess is the observations, and neither the distribution of the sources nor the mixing matrix is known. Under the assumptions that the sources are statistically independent and non-Gaussian (at most one of them can have Gaussian distribution), we find the unmixing matrix W by maximizing some measure of independence. In other words, a separation matrix, W, is estimated, which, under ideal conditions, is the inverse of the mixing matrix A.

$Y = WX$ and $W = A^{-1}$ and $Y = S$ [7]

B.Linear discreminant analysis

Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA) are two commonly used techniques for data classification and dimensionality reduction.

Class-dependent transformation: This type of approach involves maximizing the ratio of between class variance to within class variance.

Class-independent transformation: This approach involves maximizing the ratio of overall variance to within class variance.

C.Principal Component Analysis

Principal Component Analysis is a standard technique used in statistical pattern recognition and signal processing for data reduction and Feature extraction. Principal Component Analysis (PCA) is a dimensionality reduction technique based on extracting the desired number of principal components of the multi-dimensional data. The purpose of

@IJAERD-2014, All rights Reserved

PCA is to reduce the large dimensionality of the data space (observed variables) to the smaller intrinsic dimensionality of feature space (independent variables), which are needed to describe the data economically. This is the case when there is a strong correlation between observed variables. The first principal component is the linear combination of the original dimensions that has the maximum variance; the n-th principal component is the linear combination with the highest variance, subject to being orthogonal to the n -1 first principal components. The main idea of the principle component is to find the vectors that best account for the distribution of cloud images within the entire image space. These vectors define the subspace of cloud images, which we call "cloud space". PCA also known as Karhunen-Loeve (KL) transformation or eigen space projection, a frequently used statistical technique for optimal lossy compression of data under least square sense, provides an orthogonal basis vector-space to represent original data[8] .

VII. STEPS OF PCA

Acquire the initial set of cloud images (the training set).
Calculate the eigenspaces from the training set, keeping only the M images that correspond to the highest eigenvalues. These M images define the cloud space. As new clouds are experienced; the eigenspaces can be up-dated or recalculated.
Calculate the corresponding distribution in M-dimensional weight space for each known individual, by projecting his or her face images onto the "face space".

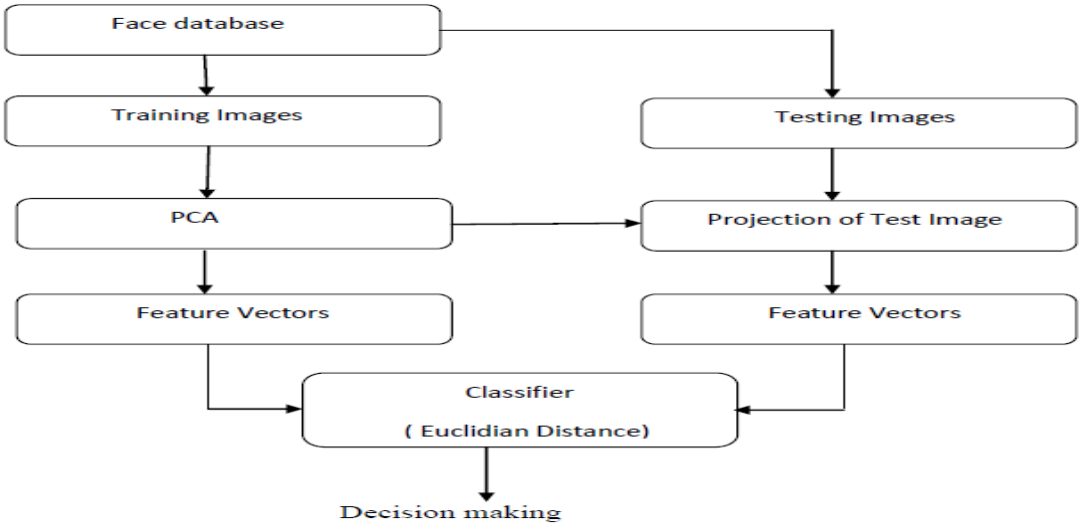

Figure 3. PCA Flowchart

VIII. MATHEMATICAL FORMULAS OF PCA

Images of faces, being similar in overall configuration, will not be randomly distributed in the huge space and thus can be distributed by a relatively low dimensional subspace [9].
Let the training set of face images be X1, X2, X3…X*n* , then the average set or mean of faces be defined as:

$$m=\sum_{i=0}^{n} Xi$$

Notice that the symbol m to indicate the mean of set X.
The average distance of each face from the mean of the data set is given by:
Q1 = X1- m; Q2 = X2- m … Qn = Xn- m
Which is the Standard deviation.
The covariance matrix is given by
C = A*A'
Where A = *[Q1 Q2 Q3……..Qn]*.
In order to reduce the dimensionality, co-variance can be calculated as
C = A'*A.
Now Eigen values and Eigen vectors are calculated for the covariance matrix. All the face images in the database are projected in to Eigen space and weight for each image is calculated. Then image vectors for each face image is obtained as

$$\text{Image Vector}=\sum_{i=1}^{n} \text{weight(i)} * \text{eigenvector(i)}$$

Sort eigen values and corresponding eigenvectors in decreasing order.
Select first d <= n Eigen vectors and generate data set in new representation.
Projected test image is compared to every projected training image and result is the training image closest to the test image [10].
To summarize the eigen faces approach to the facial expression detection involves the following steps:

@IJAERD-2014, All rights Reserved

Firstly, the train images are utilized to create a low dimensional face space. This is done by performing Principal Component Analysis (PCA) in the training image set and taking the principal components (i.e. Eigen vectors with greater Eigen values). In this process, projected versions of all the train images are also created.
Secondly, the test images also projected on face space.
Thirdly, the Euclidian distance of a projected test image from all the projected train images are calculated and the minimum value is chosen in order to find out the train image which is most similar to the test image.
Fourthly, in order to determine the intensity of a particular expression, its Euclidian distance from the mean of the projected neutral images is calculated [10].

IX. CONTOURLET

In the field of Geometrical Image Transforms, there are many 1-D transforms designed for detecting or capturing the geometry of image information, such as the Fourier and wavelet transform. However, the ability of 1-D transform processing of the intrinsic geometrical structures, such as smoothness of curves, is limited in one direction, then more powerful representations are required in higher dimensions. The Contourlet transform which was proposed by Do and Vetterli in 2002, is a new two-dimensional transform method for image representations. The Contourlet transform has properties of multiresolution, localization, directionality, critical sampling and anisotropy. Its basic functions are multiscale and multidimensional. The contours of original images, which are the dominant features in natural images, can be captured effectively with a few coefficients by using Contourlet transform. The Contourlet transform is inspired by the human visual system and Curvelet transform which can capture the smoothness of the contour of images with different elongated shapes and in variety of directions. However, it is difficult to sampling on a rectangular grid for Curvelet transform since Curvelet transform was developed in continuous domain and directions other than horizontal and vertical are very different on rectangular grid [11].

X. PROPOSED SYSTEM

In proposed system the, input image is the digital cloud image. The input image is given to system in order to discriminate the status of sky from the input cloud image. In existing system wavelets was taken into account to differentiate between the sky and cloud. Now the features are extracted from the image to represent the sky status. After that the feature extracted images of sky status is processed with cloud screening algorithm and we progress with the feature extracted image to represent the status of cloud

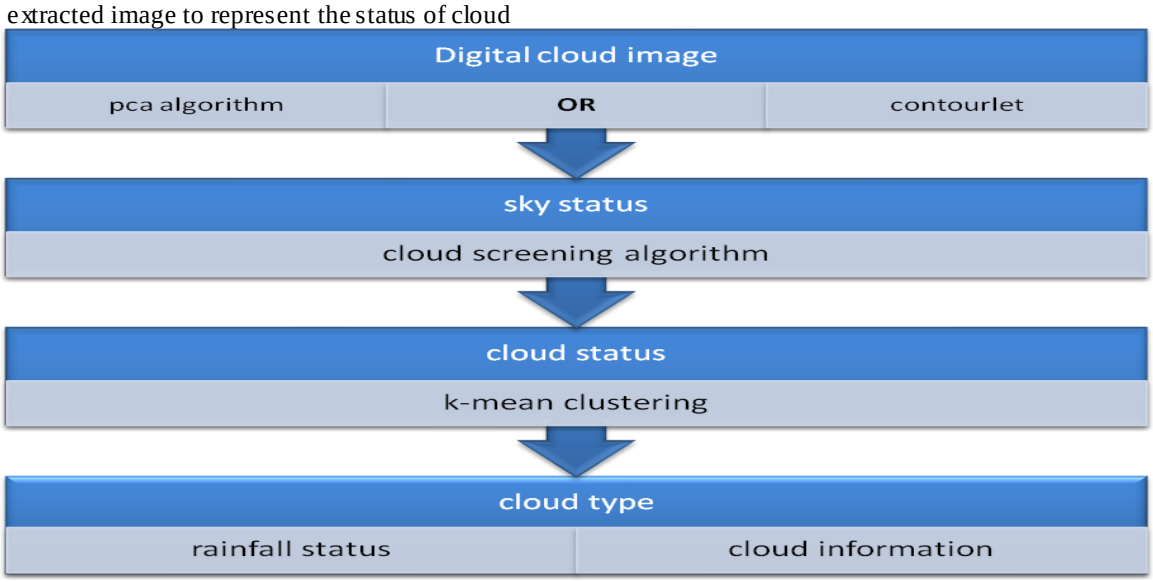

Figure 4.Proposed system architecture

A.Input Data

The input data are the digital cloud images. File system are used to store the images. The dimension of image would be normal or say 400x300. The images would be loaded either from the site which in regular interval upload images or taken with minimum 10 megapixel camera with good clarity. Images taken in different atmosphere help in study better result and derive fruitful approximation of system's efficiency.

B.Sky Status

The second phase is discriminate cloud and sky, i.e. sky status. In previous year research sky status is found using the wavelets and law's texture description. But we use contourlet or pca algorithm for finding the sky status. As pca algorithm and contourlet theory provides that feature extraction by these techniques provides better result.. For example: blood cells detection, molecule identification ,brain tumor, blood cells detection, molecule identification and bone crack detection, if contoulet or pca algorithm is used it produces a better result. Our application is cloud so the cloud is the separation point.

@IJAERD-2014, All rights Reserved

C.Cloud Status

The progressively we obtain the status of cloud. The cloud varies according to the thickness and brightness. Threshold and high density is detected. The density with the middle level is considering the sky. In mild time or moody time sky is considered as cloud. To find the mask the histogram equalization is used the value with the highest weight will be considered as sky and the others will be considered as cloud. Also the cloud mask algorithm is used here. The formula to find the cloud status is given as follows.

Cloud Status = Total No. of Segment – Sky Status

The cloud mask algorithm consists of certain tests. Single pixel threshold tests are used first. Dynamic histogram analysis is used to get threshold. Thick High Clouds (Group 1): Thick high clouds are detected with threshold tests that rely on brightness temperature in water vapor bands and infrared. Thin Clouds (Group 2): Thin clouds tests rely on brightness temperature difference tests. Low Clouds (Group 3): Low clouds are best detected using solar reflectance test and brightness temperature difference. Spatial uniformity test is also used over land surface. High Thin Clouds (Group 4): This test is similar to Group1, but it is spectrally turned to detect the presence of thin cirrus. Spatial uniformity test and Brightness temperature difference test are applied. Temporal uniformity test is also processed.

D.Cloud Type

The major task here is to find the type of cloud as per the cloud status each and every cloud will be having its own shape and density and the values are matched accordingly. The type of cloud is identified by using clustering. We use K-means clustering to combine the pixels in order to differentiate the clouds. The thickness of the clouds will be in the base part. The color, Shape and Texture are the concepts used in order to find the type of cloud. The formula to find the cloud type is shown as follows:

$$H(n) = \sum_{n=0}^{255} C[i,j]$$

Cloud id = Highest Density of Cloud Status

K-Means algorithm is an unsupervised clustering algorithm that classifies the input data points into multiple classes based on their inherent distance from each other. The algorithm assumes that the data features form a vector space and tries to find natural clustering in them. The points are clustered around centroids μi∀=1….k which are obtained by minimizing the objective

$$V=\sum_{i=1}^{k}\sum_{xj\in si}(xj-\mu i)\text{^}2$$

Where there are k clusters Si, i = 1, 2,…., k and μi is the centroid or mean point of all the points xj ϵ Si.

As a part of this project, an iterative version of the algorithm was implemented. The algorithm takes a 2 dimensional image as input. Various steps in the algorithm are as follows:

Compute the intensity distribution (also called the histogram) of the intensities.

The Centroids are initializing with the k random intensities.

Repeat the below steps until the cluster a label of the image does not change anymore.

Cluster the points based on distance of their intensities from the centroid intensities.

Compute the new centroid for each of the clusters.

E.Rainfall status and cloud information

The rainfall is estimate from the type of cloud. Cumulonimbus and nimbostratus are the rainfall clouds. So we take the color and shape and also the width and find the rainfall status the temperature is also taken into account.Cloud information gives the theoretical proof of cloud that is altitude, height, appearance and classification are given.

XI. CONCLUSION

The rainfall would be estimate accurately by determining type of cloud using the methods like pca or contourlet,cloud screening algorithm and k-mean clustering. Theoretical study says that it must provide better performance compared to previous technique like wavelet, k-mediods and other. Considering the cost factors and security issues, the digital cloud images were used to predict rainfall rather than satellite images. The feature extracted from image and using contourlet or pca sky status is found. By applying cloud screening algorithm the cloud status is obtained and finally type of cloud is detected using k-mean clustering technique. Thus, rainfall is predicted by analyzing color, density and intensity of cloud images. The result will be predicted by analyzing the color and density of the cloud images. Hence, using this analysis rainfall can be estimated in my future work using digital cloud image.

ACKNOWLEDGEMENT

I heartily thankful to Prof. Hitesh Bheda for his valuable guidance towards research direction and continuous support for performing various practical approaches.

@IJAERD-2014, All rights Reserved

REFERENCES

[1] K.Kaviarasu, P.Sujith and Mr. G. Ayaappan, "Prediction of rainfall" IEEE International Conference on Computational Intelligence and Computing Research, 2010.
[2] K. Richards and G.D. Sullivan," Estimation of Cloud Cover using Colour and Texture" Intelligent Systems Group, University of Reading, RG6 2AY,2006.
[3] Tony F. Chan and H.M. Zhou,"
[4] Kuo-Lin Hsu, X. Gao, and Soroosh Sorooshian,"Rainfall Estimation Using Cloud Texture Classification Mapping", Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721.
[5] Luis Gómez-Chova, Gustavo Camps-Valls, Javier Calpe-Maravilla, Luis Guanter, and José Moreno, "Cloud-Screening Algorithm for ENVISAT/MERIS Multispectral Images", IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 45, NO. 12, DECEMBER 2007.
[6] Minh N. Do and Martin Vetterli," The Contourlet Transform: An Efficient Directional Multiresolution Image Representation", IEEE TRANSACTIONS ON IMAGE PROCESSING.
[7] Sanjay D. Sawaitul, Prof. K. P. Wagh, Dr. P. N. Chatur," Classification and Prediction of Future Weather by using Back Propagation Algorithm-An Approach", International Journal of Emerging Technology and Advanced Engineering ,ISSN 2250-2459, Volume 2, Issue 1, January 2012
[8] Gabriele Poli, Giulia Adembri, Monica Gherardelli and Maurizio Tommasini, "DYNAMIC THRESHOLD CLOUD DETECTION ALGORITHM IMPROVEMENT FOR AVHRR AND SEVIRI IMAGES" , IGARSS, 2010, pp. 4146-4149.
[9] Ms.Chinki Chandhok, Mrs.Soni Chaturvedi and Dr.A.A Khurshid," An Approach to Image Segmentation using K-means Clustering Algorithm", International Journal of Information Technology (IJIT), Volume
[10] – 1, Issue – 1, August 2012 ISSN 2279 – 008X, 11-17.
[11] Malay K. Kundu, Priyank Bagrecha, "Color Image Retrieval Using M-Band Wavelet Transform Based Color-Texture Feature",*ICAPR*, 2009, Advances in Pattern Recognition, International Conference on, Advances in Pattern Recognition, International Conference on 2009, pp. 117-120, doi:10.1109/ICAPR.2009.34

@IJAERD-2014, All rights Reserved