



# Dimensionality Reduction on Cloud Images Based on Various Climate Zones

L. Gowri, K. R. Manjula

**Abstract:** In recent decades, cloud image classification has become a research hotspot in the field of weather forecasting. Initially, cloud images that fall on various climate zones are categorized based on their regions. Dimensionality reduction is performed in the cloud images by applying Principal Components Analysis (PCA) to enhance the classification accuracy of cloud images. The proposed system uses the training set to learn the features of cloud images and classifies the test case images into low, medium and high. The experimental results are obtained by implementing the INSAT weather image data set using MATLAB tool. The proposed methodology can be used in various applications like Rainfall Prediction, Oceanography and Cyclone Forecasting.

**Keywords:** INSAT weather satellite images, PCA, principle features,

## I. INTRODUCTION

Cloud image classification is a research hotspot in the field of weather forecasting applications in recent decades. Pattern recognition techniques are used to analyze various applications like weather prediction and forecasting. The research on cloud images is necessary to consider common techniques for finding patterns. PCA is one of the pattern recognition techniques which is mostly used for face recognition [1, 2], image compression and also commonly applied in remote sensing images. The remote sensing is primarily a multi-disciplinary science which includes the various disciplines such as photography, optics and spectroscopy. The classification techniques are used in variety of approaches such as neural networks, fuzzy logic, Artificial Intelligence and Genetic Algorithm [3]. Generally image classification techniques can be categorized into three types: supervised classification, unsupervised classification and hybrid of supervised/unsupervised classification. Basically, supervised classification is often used for quantitative analysis of satellite images. The supervised classification is divided into four steps (4Cs): Choose training portions, Create feature selection, subdivide the Corresponding training fields and Classify the image. Unsupervised classification is used in clustering the values based on analytical procedure.

Hybrid supervised/Unsupervised methodology is mainly focused on either possibility. This paper aims to differentiate clear sky and clouds in a given image and if clouds are present then identify types of clouds such as low level, mid level and high level clouds. This paper highlights first the related work and then the methodology and the PCA algorithm applied elaborately and finally the results and experiments.

## II. RELATED WORK

High dimensionality of remote sensing data, feature extraction and feature selection techniques have been broadly discussed by many authors with different perspectives. The related works of this paper focus on PCA based cloud classification satellite images. The conventional PCA is one of the most popular choices in the field of face recognition, image compression and image fusion. PCA is used to transform image data into orthogonal set which gives most relevant principal features from the images. Many researchers have addressed cloud classification in different viewpoints. S.C.Ou and Tankano [4] calculated only cirrus cloud parameters through various attributes of images like mean, gray scale value and so on. The cloud images were taken from Advanced Very High Resolution Radiometer (AVHRR) satellite. In training phase, sample sizes of images were used by  $16 \times 16$  pixels. Bankert, R. L. [5] used Probabilistic Neural Network (PNN) technique for classification of cloud images with the size of 16 pixel  $\times$  16 pixel sample region on Advanced Very High Resolution Radiometer (AVHRR) satellite. William E. Shenk et al. [6] four registered channels of the Nimbus-3 Medium Resolution Infrared Radiometer (MRIR) have been used for discrimination of individual cloud types and cloud types combination over tropical oceans. Adriana Romero et al. [7] used unsupervised sparse features for greedy layer-wise pre-training algorithm. Single layer convolution network used to extract only neighboring pixels. Jorge Costa Pereira et al. [8] specified spectral activity in terms of unequivocal information on the number of active spectral components in aquatic system. Bryan A. Baum [9] presented a Fuzzy Logic Classification (FLC) methodology is used to analyze cloud fraction with the help of  $32 \times 32$  pixel sample size. This method is applied in single-layered and multilayered type clouds. Cloud-top height was determined using ATSR data by Turner, P.J. [10]. The cloud top height is calculated more accurately from satellite images for weather forecasting and analyzing.

**Revised Manuscript Received on 30 July 2019.**

\* Correspondence Author

**L.Gowri**, School of Computing SASTRA Deemed University, Thanjavur, Tamil Nadu, India.

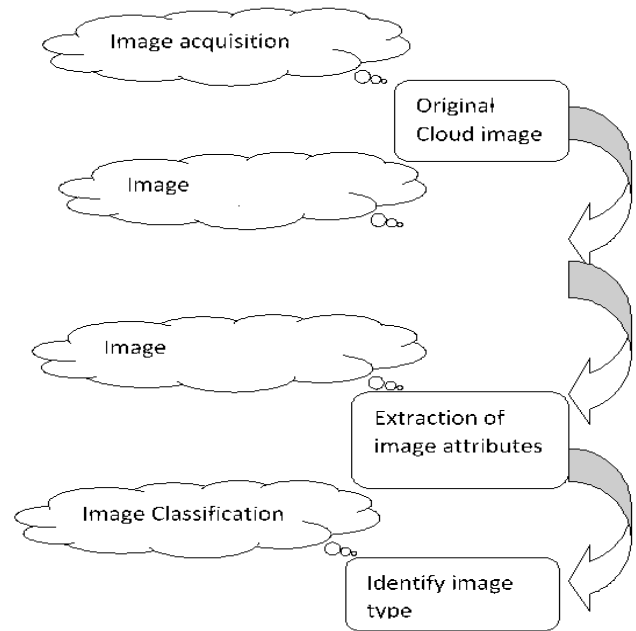
**Dr.K.R.Manjula**, School of Computing SASTRA Deemed University, Thanjavur, Tamil Nadu, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

In this paper we proposed PCA for various climate zones of satellite images. PCA can handle a huge amount of data due to data reduction and data complexity. And also PCA is used to avoid additional calculation steps of images. Cloud detection consists of identifying cloud portions in complex scenery. In recent decade textual, spectral and spatial signatures have been used to categorize cloud types. Imagery classification studies are used in discrimination of cloud types in Tropical Region, Mid Latitude Region and Polar Regions

### III. METHODOLOGY

Cloud classification is the main research area in weather prediction. This paper classifies cloud images in terms of low, mid and high level clouds using Principal Component Analysis (PCA). This method is mainly used for feature extraction from cloud images present in INSAT weather satellite. The main objective is detecting different cloud types from visible and infrared satellite images. Principal feature transformations are mainly used for feature extraction as well as image enhancement. Before using pattern recognition, the least principal features are neglected. These neglected features are allowed to avoid the insignificant part of image information and additional computation times. Once feature extraction process is over then feature selection will take place based on priority of principle characteristic significance. The transformation functions are determined during the training phases of PCA to derive principal features using Eigen values and Eigenvectors. The Eigenvectors are useful to define related set of spatial characteristics of an image. The predominant Eigen values and Eigenvectors are used to detect first level of image discrimination. Once features are identified, then training phase is used to train number of samples and label class names as low, mid and high level. In testing phase, new image will be classified based on the features. The special features are used to discriminate cloud types with respect to pixel density. Figure-1, give an idea about the steps of image discrimination. The following steps are used to describe the steps of image processing.



**Figure-1 Steps in cloud type classification**

1. **Image Acquisition** – This step assist to collect new image. These images can be acquired through different satellite resources. Before processing, the images convert into gray scale image with requires size.
2. **Image Segmentation**- This is used to differentiate foreground and background layers of the image and also to identify cloud fraction in the given image
3. **Image Representation** - This step identifies the cloud attributes through principal component
4. **Image Classification**- Images are classified with their particular types based on matching attributes of the image.

#### A. Principal Features of cloud images

PCA is statistical methods which is mainly used for reduce dimensionality of image as well as remove redundancy of image. The Eigenvectors are representing characterization of spatial information of satellite images. Principal features were extracted with the help of following steps:

#### Algorithm: Principal Feature Extraction

**Input:** satellite image with gray scale value

**Output:** Detection of cloud types

**Step 1: Construction of Row vectors:**

Take the image as  $m \times n$  matrix size and convert into a first row  $1 \times mn$  size matrix, Take next image as  $m \times n$  matrix size and convert into a second row  $2 \times mn$  size matrix and so on.

**Step 2: Calculate mean vector**

Calculate mean cloud vector  $\mu$  of  $n$  rows

$$\mu = \frac{1}{n} \sum Y_j$$

$j = 0, 1, 2, 3, \dots$

**Step 3: Constructing training space**

Create  $\phi$  by subtracting mean ( $\mu$ ) vector from each cloud image

**Step 4: Calculate covariance matrix**

Covariance matrix  $Cov\_mat$  is considered by multiplying matrix  $\phi$  with its transpose matrix

Cloud types	Tropical Region	Mid Latitude Region	Polar Region
Low	0 to 2000 m	0 to 2000 m	0 to 2000 m
Middle	2000 to 8000 m	2000 to 7000 m	2000 to 4000 m
High	6000 to 18000 m	5000 to 13000 m	3000 to 8000 m

$$Cov\_mat = \phi \cdot \phi^T$$

**Step 5: listed dominate values of image from image space**

Selecting highest Eigenvectors from the image space

**Step 6: Calculate projection of image**

A new image is identify by deriving projection on the image space

$$\text{Projection} = \text{image space} * (\text{new image values} - \text{mean value})$$

**Step 7: Discriminate cloud types**

Fix threshold values and identification occurs at particular cloud types





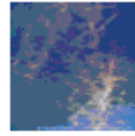

The above algorithm is used in detection of cloud images. First, construct image data into  $1 \times mn$  format. Next calculate mean value of each direction of images. All the  $x$  values are subtracted from the mean value of  $x$  and the entire  $y$  values are subtracted from the mean value of  $y$ . Calculate covariance matrix of the image through which the Eigen values and Eigen vectors are calculated. Finally choose feature vector which gives more dominant similarity of the images from the image space.

**Table 1 : Types of regions and its cloud types**

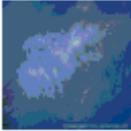

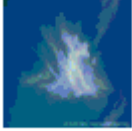



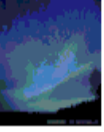

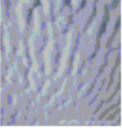

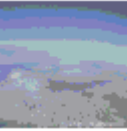







Table 1 is used to represent differentiate regions based on their altitude. Low level clouds are made of water droplet whose heights are below two kilometer in altitude. This is mainly used to predict rainy status. Some example for low level clouds includes cumulus, stratus, nimbostratus and fog. The mid level clouds altitude in between two and six km. Some example for mid level clouds includes altocumulus and altostratus. High level clouds altitude in above six km. The particle of these clouds can be either ice or water droplet.





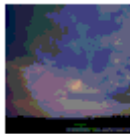

**IV. RESULTS AND EXPERIMENTS**

In this paper, cloud detection of INSAT satellite images and weatherscape images are used with respect to various zones. The training samples are used to create image space. In order to test new images through this image spaces the following output could be derived from MATLAB 13.a, Intel core i5 processor at the speed of 2.50GHZ

Original Images	Cloud Type	Processing Time	Eigen Values and Vectors	Processed Images
	High-Level Cirrus	5.602 s	EVal – 1 Evec- 514.3729	
	High-Level Cirrus	0.919 s	EVal – 1 Evec- 530.0393	
	High-Level Cirrus	0.313 s	EVal – 1 Evec- 614.1452	

## Dimensionality Reduction on Cloud Images Based on Various Climate Zones

	High-Level Cirrus	0.301 s	EVal – 1 Evec- 492.8527	
	High-Level Cirrus	0.286 s	EVal – 1 Evec- 1049.4	
	Mid-Level Alto cumulus	0.431 s	EVal – 1 Evec-3782.7	
	Mid-Level Alto cumulus	0.423 s	EVal - 1 Evec-1226.7	
	Mid-Level Alto cumulus	0.338 s	EVal -1 Evec -309.8860	
	Mid-Level Alto cumulus	0.386 s	EVal -1 Evec -423.4747	
	Mid-Level Alto cumulus	0.298 s	EVal-1 Evec-1006.4	
	Low-Level Strato cumulus	0.287 s	EVal – 1 Evec- 3297.0	
	Low-Level Strato cumulus	0.371 s	EVal – 1 Evec- 883.2221	

	Low-Level Stratocumulus	0.301 s	EVal – 1 Evec- 772.8233	
	Low-Level Stratocumulus	0.287 s	EVal – 1 Evec- 2002.5	
	Low-Level Stratocumulus	0.307s	EVal – 1 Evec- 973.5092	

**V. CONCLUSION**

This paper vividly expressed the PCA based cloud image classification for satellite imagery. The experiment results specify that the PCA method is capable of discriminating cloudy and clear sky images. PCA algorithm handles a large amount of data and provides more accurate cloud classification. Eigen vectors are used to train cloud images and classify accordingly. Future research will concentrate on the effective way of detecting cloud images though different sub types of each main type clouds. The availability of spatial information is useful for significant image classification effectively. The success of cloud image classification depends on many features. Future work will focus on additional features like texture, shape, density, cloud fraction and different levels of layered cloud images.

**REFERENCES**

1. Estornell, Javier, et al. "Principal component analysis applied to remote sensing." *Modelling in Science Education and Learning* 6 (2013): 83-89.
2. Bajwa, Imran Sarwar, and S. Irfan Hyder. "PCA based image classification of single-layered cloud types." *Proceedings of the IEEE Symposium on Emerging Technologies*, 2005.
3. Pereira, Jorge Costa, et al. "Unsupervised component analysis: PCA, POA and ICA data exploring-connecting the dots." *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 165 (2016): 69-84.
4. Bryan A. Baum, Vasanth Tovinkere & Jay Titlow, Ronald M. Welch, 1997. Automated Cloud Classification of Global AVHRR Data Using a Fuzzy Logic Approach. *Journal of Applied Meteorology*. pp 1519-1539.
5. Bankert, R. L., 1994: Cloud classification of AVHRR imagery in maritime regions using a probabilistic neural network. *Journal of Applied Meteorology*. 33, 909–918. Boulder, CO 80307.
6. Shenk, William E., Robert J. Holub, and Robert A. Neff. "A multispectral cloud type identification method developed for tropical ocean areas with Nimbus-3 MRIR measurements." *Monthly Weather Review* 104.3 (1976): 284-291.
7. S. C. Ou, K. N. Liou, W. M. Gooch, and Y. Takano, [1993] "Remote sensing of cirrus cloud parameters using advanced very-high resolution radiometer 3.7- and 10.9-m channels," *Appl. Opt.* Vol. 32, pp. 2171–2180.
8. Romero, Adriana, Carlo Gatta, and Gustau Camps-Valls. "Unsupervised deep feature extraction for remote sensing image classification." *IEEE Transactions on Geoscience and Remote Sensing* 54.3 (2016): 1349-1362.

9. Prata, A.J., Turner, P.J., [1997] "Cloud-top height determination using ATSR data" *Remote Sensing Env.*, 59, 1, pp. 1-13.
10. Macías-Macías, Miguel, et al. "A comparison of
11. PCA and GA selected features for cloud field Classification." *Ibero-American Conference on Artificial Intelligence*. Springer Berlin Heidelberg 2002.