LINEAR ALGEBRA -II

Submitted By: -

Uzair Rehman	2020509
Aiman Jamil	2020060
Sirat Abdul Sattar	2020462

Abstract

A multispectral image refers to an image that captures and represents information across multiple discrete spectral bands or wavelength ranges. Unlike traditional images that capture only the visible spectrum (red, green, and blue), multispectral images acquire data in additional spectral bands beyond what the human eye can perceive. These additional bands can cover a wide range of wavelengths, including ultraviolet (UV), near-infrared (NIR), and thermal infrared (IR), among others.

A multispectral image may be composed of several bands. For example, the Landsat 5 satellite produces 7-band images, with the wavelength of the bands being between 450 and 1250 nm.

Principal component analysis, or PCA, is one of the most used dimensionalities reduction techniques that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set. In this complex engineering project, we have applied the techniques of PCA to reduce the dimensionality of multispectral images and performed error analysis on the images transformed due to PCA.

Introduction

Multispectral imaging uses more than 3 spectral filters to capture images. By capturing and analyzing data across multiple bands simultaneously, multispectral imaging allows for the extraction of valuable information that is not visible in conventional images. Multispectral imaging has various applications all around the world, some of which are listed below:

- Land mine detection
- Weather forecasting
- Space-based imaging
- Ballistic missile detection
- Documents and artworks
- Military target tracking

PCA is a statistical technique for reducing the dimensionality of a dataset. This is accomplished by linearly transforming the data into a new coordinate system where (most of) the variation in the data can be described with fewer dimensions than the initial data. Principal component analysis has applications in many fields such as population genetics, microbiome studies, and atmospheric science.

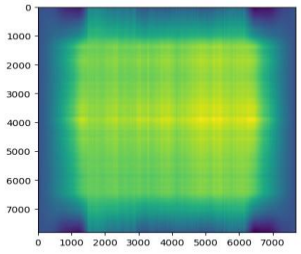
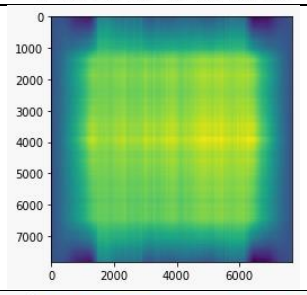
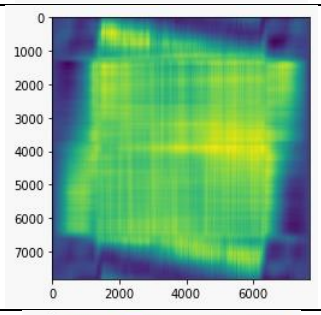
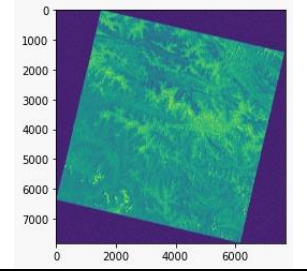
The Landsat program keeps a repository of the earth's images taken, we have downloaded two datasets one containing the multispectral image and the other containing the images of singular bands of the region of Chitral from https://earthexplorer.usgs.gov.

Methodology

The strategies that we have used to perform basic operations, apply PCA, and perform an error analysis are discussed below.

Basic operations performed on multispectral images.

These were the operations we performed on our dataset of images:

Percentage of eigenvectors	Image after transformation	Mean Squared error
10%		0.198
20%		0.198
75%		0.015
95%		0.0017

Visualization of Bands

In our dataset, we had multiple raster graphic images representing different bands. Using the Geospatial Data Abstraction Library (GDAL) of Python, we were able to read the raster images and convert them into arrays, which allowed us to access and manipulate the pixel values of the images programmatically. Once the data was converted, we utilized the Matplotlib Python library for data visualization.

Cropping the size

Using the Python Imaging Library, we cropped the image in 4 dimensions that were left, top, right, and bottom.

Concatenation of bands

During this phase, we scaled the images that represented different bands of our multispectral image. After scaling the images, the numpy library was used to concatenate the bands together into one image. After this, the concatenated image was visualized using matplot.

Methodology for Applying PCA To Multi-Dimensional Image

Principal component analysis is a popular statistical technique used for the dimensionality reduction of a dataset. The main concept behind the approach is to transform the data into a new coordinate system such that the maximum amount of variation in the data is preserved with fewer dimensions than the original dataset.

We have used the popular Python framework named 'sklearn' to apply the technique of Principal component analysis on our multi-dimensional image. Following are the steps taken to apply PCA using sklearn:

- Importing the required libraries.
  - matplotlib = For plotting the multi-dimensional and PCA applied to the image.
  - Sklearn.decomposition.PCA = To apply PCA on a multi-dimensional image.
  - numpy = To work with multi-dimensional arrays and matrices.
- Reading the multidimensional image using Matplotlib's imread function.
- Preprocessing Step:
  - Changing the data type of the image array to uint8 for easy handling of an array.
  - Normalizing the array
  - Converts the image array to 2-D to feed the image to sklearn's fit function.
- Creating PCA function:
  - Using sklearn.decomposition.PCA(n_components).fit to find n_principal_components
  - Calling the transform() function to apply the founded components on the image
  - Calling .inverse transform() function to find projections of principal components on the image

Error analysis on the use of PCA

In our analysis, we performed an error analysis to assess the quality of the transformed images obtained after applying PCA with different numbers of principal components. The transformed matrices resulting from PCA with various numbers of components are stored within the "Matrices" list. The entries in this list correspond to the transformed matrices obtained when applying PCA with 10, 30, 50, and 95 eigenvectors, respectively.

Since the dimensions of the original image and the transformed images are the same, you utilized mean squared error (MSE) to quantify the magnitude of the error between the original image and each of the transformed images. MSE is a common metric used to measure the dissimilarity or error between two images.

Result

According to the experiment, increasing the percentage of the PCA reduces the magnitude of the Mean squared error between the original image and the transformed image. The suitable number of principal components is 95% because it gives less MSE as compared to other numbers of principal components.

Conclusion

Multispectral remote sensing involves the acquisition of visible, near-infrared, and short-wave infrared images in several broad wavelength bands. PCA helps in interpreting the data. Principle Component analysis simplifies the complexity in highdimensional data while retaining trends and patterns. Patterns have applications in various fields like Medical Field, Satellite imagery, and Agriculture, etc.

References

[1] Wikipedia , "Earth observation satellite," 21 Jan 2017. [Online]. Available: https://en.wikipedia.org/wiki/Earth_observation _satellite