ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

A Study of Linear Algebra for Computer Vision

Anshul Garg [1], Kiran Kaundal [2]

Assistant Professor, Department of Computer Application, Chandigarh Group of Colleges, Landran, Mohali, Punjab, India [1,2]

ABSTRACT: Human body use their brain and eyes to see and sense the world, computer vision aim is the same. It is mainly concerned with analysis and understanding of information from an image or multiple images which involve development of algorithmic and theoretical basis. Linear algebra plays an important role to achieve the aim of computer vision. In this paper, the main focus is on how linear algebra helpful in computer vision. This paper is representing the concept of some basic linear algebra like Matrix, Least square, single value decomposition, homogeneous coordinates, eigenfaces and how they are helpful in computer vision.

KEYWORDS: Computer Vision, Least Square, Homogeneous Coordinates, Eigenvectors, Eigenfaces, Transformations.

I. INTRODUCTION

Computer Vision is a field of study in which basically we learn how to reconstruct, understand and interpret the 3D image into the 2D image in terms of structure's properties in the scene. Its main goal is to create a model, replica and most important exceed the human vision with the help of computer hardware and software by using mathematics, science etc. Computer Vision is quite difficult as vision is an inverse problem and image data can be noisy and the data can be uncertain [13]. The hierarchy of computer vision is

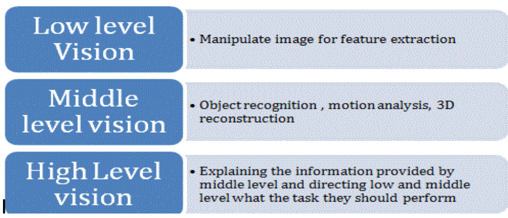

Fig-1 the Hierarchy of Computer Vision is shown in

Computer vision is divided into three levels. In the lowest level we can manipulate the images is we want to extract the features, Motion analysis and 3D reconstructions can be done in the middle level, and High level vision how low and middle level should perform their task. There are many applications of computer vision like Image formation, Image processing, Feature detection and matching, image rendering, 3D constructions and much more. The common approaches for solving the computer vision's problem are linear models and statistical model. In statistical models, the probability is generally used to determine what can occur in an image on the basis of given parameters of the image whereas in linear models we take the help of linear algebra in which we take multiple images at different but close to each other from different angles [13]. Properties those are identified in images like points, lines, and planes are stored as vectors in camera matrix. Then each of these vectors can be transformed through projective geometry such that the relative coordinate system of every image is aligned and the one can compare the two images. In the next section, the discussion is on various mathematical notations that are helpful to understand the computer vision.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

II. RELATED WORK

The research on computer vision usually deals with relatively domain- independent considerations. The results are very beneficial in a broad range of context. As the understanding of computer vision is growing, its applications are increasing appreciably. Computer vision technology can be used in many different field like in food quality evaluation, image processing, camera modeling, face detection, forensic etc. To understand computer vision, linear algebra plays an important role. Many other linear algebra terms like SVD, Least Square ,Convolution ,Eigen vector ,homogeneous coordinates ,plays a great role to understand computer vision problem .By using these algebraic terms a number of computer vision's algorithms can be derive to get more appropriate results.

III. LINEAR ALGEBRA

Linear algebra is the branch of mathematics that deals with vector spaces and linear mapping between spaces. In this section, the focus is on the various mathematical notations that can be very useful in computer vision like Homogeneous coordinates Least Squares, Singular Value Decomposition, Eigen faces, Matrix, and Geometric Introduction etc. which are used as main tools in solving the problem of computer vision.

A. Homogeneous Coordinates:-

In computer vision, we use the concept of projection of linear system; it is easier to express points in homogeneous coordinates. In Homogenous coordinates, we add an extra dimension to the standard Euclidean coordinate system means if we have (x, y) points in x y-plane, it's representation in homogeneous coordinates will be (x, y, w) in the same plane [7] .The concept of geometric interpretation of homogeneous coordinates (x, y, w) can be explained by the figure-1.

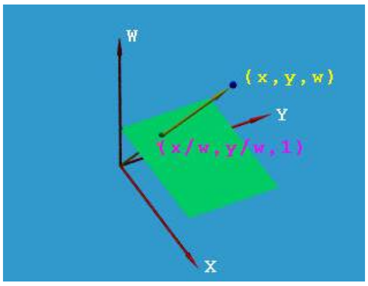

Fig-2[2] Geometric interpretation of homogeneous coordinates

In the above figure, the line joining this point and the coordinate origin intersects the plane w = 1 at a point (x/w, y/w, 1). This transformation treats a 2D homogeneous point as a point in 3D space and projects this 3D point to the plane w=1. Therefore, when homogeneous point moves on a curve described by homogeneous polynomial f (x, y, w) =0, its related point moves in 3D space [7]. Another main purpose of using these coordinates is to capture the concept of infinity when using planes and lines to understand curved surfaces or the intersection of two parallel lines, the concept of infinity becomes very useful [8]. The point (3, 2) in Cartesian coordinates becomes (3, 2, 1) in homogenous coordinates and as this point moves out towards infinity we can illustrate it in homogeneous coordinates as (3, 2, 0).

B. *Matrix :-*

Matrix is an important tool that can be used in computer vision. Images in computers can be represented in the form of Matrix .We can explain the concept through the following figure-3

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

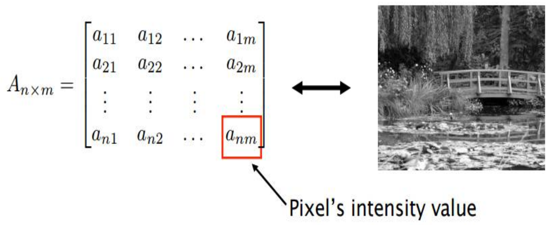

Fig-3 Matrix Representation of Image

In the above figure ,each individual coordinate will be referred as one pixel and the each element of the matrix being a whole number ranging from 0 (for black) to 225 (for white) to form the image
Camera Matrix is the most important matrix that we use in computer vision. The camera projection matrix is the way to represent the 3-D transformation of real world vectors into projected camera interpreted vectors [13]. For this we must know about projection transformation. In this section we are discussing about the basic transformation and their matrix representation that we can use to transform the world vector to image vector. These transformations usually act on vectors through multiplication with the matrix representation of transformation.

- *Projective*

  Projective transformation under composition form a group called projective linear group denoted by PL (3) [7]. The block representation is

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A & t \\ v & b \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H_p \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Where A is non singular matrix, t is translation in 2D vector and v= $(v_1 + v_2)^t$. Translation, Rotation, Scaling, Affine transformations are the special type of Projective transformation.

- *Translation:-*

  When an object is displaced by a given direction and distance from its original position. If the displacement is given by vector

$$\vec{v} = t_x \hat{i} + t_y \hat{j}$$

Then new coordinates can be calculated as

$$x' = x + t_x$$

$$y' = y + t_y$$

The Matrix representation in Homogeneous Coordinate System is

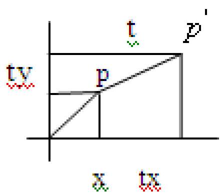

Fig-4 Translation

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

$$P = (x, y) \rightarrow (x, y, 1)$$
$$t = (t_x, t_y) \rightarrow (t_x, t_y, 1)$$

$$P' \rightarrow \begin{bmatrix} x+t_x \\ y+t_y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} I & t \\ 0 & 1 \end{bmatrix} P = TP$$

- *Rotation:-*

In the 2D rotation, the object is rotated by angle θ about the origin. The convention is the direction of the rotation is counter clockwise if θ is a positive angle and clockwise if θ is a negative angle.

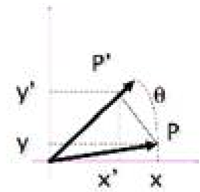

Fig-5 Rotation

$$x' = \cos\theta x - \sin\theta y$$
$$y' = \cos\theta y + \sin\theta x$$

The Matrix representation in Homogeneous Coordinate System is

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- *Scaling:-*

Scaling is the process of expanding or compressing the dimensions of an object.
Sx=Positive Scaling (Change in the length w.r.t. x direction)
Sy=Positive Scaling (Change in the length w.r.t. y direction)
In the homogeneous coordinate system, the 2D scaling can be represented as shown below:-

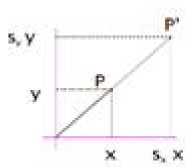

Fig-6 Scaling

$$x' = S_x x$$
$$y' = S_y y$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

Matrix Representation is $$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

- *Affine Transformation*

  The affine transformation is a transformation which maintains straight lines. It means that sets of parallel lines will remain parallel after transformation, not necessarily to preserve angles or lengths.

  The matrix representation of affine transformation is

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} A & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H_a \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Where $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

A is non singular matrix, t is translation vector, 0 is null vector.

- *Similarity:-*

  One another transformation i.e. Similarity is also a subset of the affine transformation. Angles between lines, the ratio of two lengths, and the ratio of areas are invariant under this transformation [11]. A Similarity transformation is in the form

$$H_S = \begin{bmatrix} s\cos(\theta) & -s\sin\theta & t_x \\ s\sin(\theta) & s\cos(\theta) & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

The matrix representation is

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} sR & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H_s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Where R is the nonsingular matrix, s is the isotropic scaling, t is translation 2D vector, 0 is the null 2-D vector

- *Euclidean transformation*

  Euclidean transformations are a subset of similarity transformations. Euclidean transformations preserve distance and hence are also known as isometrics.

Isometrics: $$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H_e \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Where

$$H_E = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & t_x \\ \sin(\theta) & \cos(\theta) & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

Here also, R is the nonsingular matrix, t is translation 2D vector, 0 is the null 2-D vector

D. Least Square:-

Least Square is a very important tool to solve the problem of computer vision. This is the method to minimize the ||Px-b|| for m × n matrix [13], P and the vector b. There is three types of least square that we can use in computer vision.

- *Robust least squares*

This method is used to minimize only the subset of the residual not solving the ||Px-b|| for the values of x=[x1, x2………, xn].The problem is solved only for a subset i<n of residual which means computing the residuals for each possible subset of i values from x and then select the solution which can be used as minimal residual [6].

- *Non-linear least squares*

This method is used to solve the problems where the functions are not linear in the unknown parameters. Iteratively re-linearizing about the approximation of unknown parameters can be used to solve this type of problems.

- *Total Least square:-*

Total least square assumes that there can be an error in the measurement of b as well as in the values of the data matrix, P.

In Computer Vision, we generally focus only on minimizing || Px || with constraint ||x||=1 not vector b. This minimum value can be solved by using single value decomposition.

E. Singular Value Decomposition

Single value decomposition is the matrix method which is very important in computer vision.

Definition: Any real matrix P of order m x n can be defined as the product of three matrix

P=U D V^T .Where U and V are orthogonal unit vector i.e. $UU^T = I$, $VV^T = I$.

Single value decomposition can be used in solving the homogeneous linear equations PX=O for vector X, This equation can be used when problems like the elements of camera projection matrix and homographic transform arise [9]. This is also useful while calculating the inverse of a singular matrix.SVD is also very useful in attempting to compare two objects [6].

F. Eigen Faces

Eigen faces is the name provided to a set of eigenvectors when they are used in human face recognition in computer vision. [5] In this, the variation in the collection of the face image is captured and further this information is used to encode and compare the images of individual faces in terms of the part based or featured based like nose, eyes, ears, lips etc. Eigen faces are the eigenvectors of the covariance matrix of face images matrix [2].SVD can be used as principal component analysis to produce the data in the form of eigenvectors of the covariance matrix.[3] Eigen faces are used to represent face images in an efficient manner such that they reduce the space and computation complexity.

Here we are discussing how Eigen faces are used in face image detection.

According to **M. Turk and A. Pentland** , By selecting the subset of eigenvector $\hat{U} = (u_1, u_2, ..., u_m)$ linked with m largest Eigen values, the Eigen faces span an M - dimensional subspace of actual image space which results in face space. To detect the face, one has to calculate the distance from or within the face space [2].

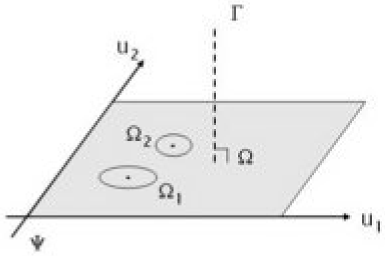

Fig-7[2] Visualization of 2D face space, with the axes representing two Eigenfaces.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

Let the set of face image be $\Gamma_1, \Gamma_2 \ldots\ldots\ldots\ldots, \Gamma_m$, Then the average face of set i.e. Ψ = Average face of face space i.e. $\psi = \frac{1}{M}\sum_{1}^{M}\Gamma_i$. The face difference from the average face will be $\Phi_i = \Gamma_i - \psi$. In the fig-7, u_1 and u_2 axes are eigenfaces of face space and Ω is the presentation of new face in the set. Face detection can be measured as finding out the image patches near to the face space. The projection distance δ (distance between the face image and its projection into face space can be calculated as $\delta = \left\|\left(I - \hat{U}\hat{U}^T\right)(\Gamma - \psi)\right\|$. Where **I** is the identity matrix, Γ is a N^2x1 vector, corresponding to an NxN face image I. In face recognition, the new face i.e. Γ can be calculated by $\Omega = \hat{U}^T(\Gamma - \psi)$

IV. APPLICATION OF LINEAR ALGEBRA IN COMPTER VISION

Nowadays, computer vision is used in solving many problems which include medical imaging, face and object detection and identification, camera modeling, structure from motion and much more [2]. The techniques provided by linear algebra plays an essential role in solving problems in each of these fields. The key benefit of using linear algebra in computer vision is that they provide more truthful and close solution to the real world problems. The fundamental of computer vision is geometry and one common way to work with geometry is through algebra [12]. In geometry, we generally deal with angles, projections, intersections of points, planes and lines etc. which can be calculated only with the help of linear algebra. Some most important applications of Linear algebra in computer vision are

A. *Camera Modeling*

Camera modeling is important parts of computer vision. A camera model is a mathematical projection between a 3D object space and a 2D image. These models can be used for calculating the geometric information from the images [10]. On the basis of parameters, there are various types of camera model but most generally only pinhole camera model is used in academic research and many applications. The pinhole model, assumes that camera rays pass via a single point, the optical center and linear relationship will occur between image point position and associated camera ray's direction [10].

B. *Computer Gaming:-*

Linear algebra is widely used in computer gaming. In computer gaming, plotting shapes like rotating them, moving them around, forces, collisions etc. In 2D as well as in 3D and 3D effects programmed with shadings use linear algebra and hardware to calculate a large amount of calculations efficiently using parallel processing [9].

C. *Image Processing:-*

In image processing, linear algebra is very helpful. An image can be manipulated and valuable information can be extracted using linear algebra computations [9]. Some examples are

- Image compression and image encryption
- Basic image processing functions like scaling, rotation, translation and adding two images.
- Object detection/recognition use Eigen values and Eigen vector.

D. *Convolution:-*

Convolution also named as Kernel is very popular and simple to use application of linear algebra in computer vision. In it, a smaller matrix is used to manipulate only the part of image matrix through mathematical convolution [13]. It has many applications in sharpening, noise reduction, sharpening, embossing/edge detection, dynamic range correction, and template matching.

E. *Pose estimation:*

The Pose is the position and orientation of an object. The task of determining the pose of an object from a 2D image is called pose estimation. It is very good for mixing the reality and virtual reality. Linear algebra plays an important role in pose estimation as in it the concept of perspective and orthogonal projection is used.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

F. Structure-from-motion

Structure from motion is the process to estimate the 3-D structure of a scene from a set of 2-D images. This is used in many applications, such as 3-D scanning and augmented reality. Structure from motion can be calculated in many different ways. Many factors are considered when an approach is decided to solve the problem such as the number and type of cameras used, and whether the images are ordered. The camera motion and 3D structure can only be recovered up to scale if the images are taken with a single calibrated camera [10].

V. CONCLUSION

Nowadays, Computer vision is getting faster and more precise in visualizing the world around us with the help of linear algebra. In this paper, basic linear algebraic terms are discussed which are very helpful to understand the concept of computer vision. There are numerous fields like image processing, camera modeling, forensics, geosciences, robotics, security and surveillance and much more where computer vision can be used, some of them are discussed in this paper which is very beneficial for further studies for students and researchers.

REFERENCES

[1] Slavković ,M., and Jevtić D.," Face Recognition Using Eigenface Approach", SERBIAN JOURNAL OF ELECTRICAL ENGINEERING, Vol. 9, No. 1, 121-130, February 2012

[2] Turk, M., and Pentland, A.," Face Recognition using Eigenfaces", Conference on Computer Vision and Pattern Recognition , , Maui, HI , USA, pp. 586 – 591. 3 – 6 June 1991

[3] Kim, K.," Face Recognition using Principal Component Analysis", National Institute of Technology, Rourkela, India, 2008.

[4] Gupta, S.,Sahoo, O.P., Goel, A.,and Gupta, R. ,"A New Optimized Approach to Face Recognition using Eigen Faces", Global Journal of Computer Science and Technology, Vol. 10, No. 1, pp. 15 – 17,April 2010.

[5] BALLARD, D.; BROWN, C., Computer Vision. New Jersey: Prentice-Hall, 1982, pp. 252- 253, 2007.

[6] Trefethen, Lloyd N., and David Bau., "Numerical linear algebra". Philadelphia, PA: Society for Industrial and Applied Mathematics, 1997.

[7] http://www.mathcs.gordon.edu/~senning/presentations/HomogeneousCoordinates.pdf.

[8] https://www.cs.mtu.edu/~shene/COURSES/cs3621/NOTES/geometry/homo-coor.html

[9] https://www.ecse.rpi.edu/Homepages/qji/CV/3dvision_intro.pdf

[10] Sturm, P., Ramalingam, S., Tardif, J-P, Gasparini, S., Barreto, J.," *Camera Models and Fundamental Concepts Used in Geometric Computer Vision*", Foundations and Trends@ in Computer Graphics and Vision, Vol. 6, Nos. 1–2, pp. 1–183, 2010

[11] Szeliski, R., "Computer *vision algorithms and applications*". London: Springer. ; 2010

[12] Shah , M.," *Fundamentals of Computer Vision*", University of Central Florida, Orlando FL 32816, pp.1-133,1997.

[13] Will,H.," Linear Algebra for Computer Vision", Attribution-ShareAlike 4.0 International (CC BY-SA 4.0),pp. 1-14,2014

BIOGRAPHY

1. **Anshul Garg** is an Assistant Professor in the Department of Computer Application, Chandigarh Group of Colleges, Landran, Mohali. She received M.Tech (CSE) degree in 2008 from Kurukshetra University, Haryana,India. Her research interests are Cloud Computing, Big Data, Data Mining, Software Engineering etc.

2. **Kiran Kaundal** is an Assistant Professor in the Department of Computer Application, Chandigarh Group of Colleges, Landran, Mohali. She received M.Phil (Mathematics) degree in 2009 from Himachal Pradesh University, Shimla, India. Her research interests are Linear Algebra, Discrete Mathematics, Operational Research etc