IMAGE FORMATION PROCESS

Imaging process maps three-dimensional scene points onto a two-dimensional image plane.

Imaging (Perspective) Transformation

3D-Scene ===============> 2D-Image

Perspective Transformation is key to the Imaging Process. It projects 3D-points onto a plane.

An analytical model of Perspective Transformation can be found by establishing:

- A world co-ordinate system representing 3D scene.
- Camera co-ordinate system \((x, y, z)\) which has the image plane.
IMAGING PROCESS MODEL

For the camera coordinate system \((x, y, z)\) assume:

- Its xy-plane is coincident with the image plane.
- Its z-axis is along the optical axis of camera or line of sight of the observer.
- Centre of camera lens is at \((0, 0, \lambda)\) where \(\lambda\) is the focal length of lens.

If \((X, Y, Z)\) is the world co-ordinates of any point in a 3-D scene and \((x, y)\) is its projection onto the image plane.

Perspective transformation is the relation between \((X, Y, Z)\) and \((x, y)\) which can be obtained by using similar triangles.
PERSPECTIVE TRANSFORMATION

\[ \frac{x}{\lambda} = -\frac{X}{(Z - \lambda)} = \frac{X}{(\lambda - Z)} \]
\[ \frac{y}{\lambda} = -\frac{Y}{(Z - \lambda)} = \frac{Y}{(\lambda - Z)} \]

The image plane co-ordinates of the projected 3-D point are:

\[ x = \frac{\lambda X}{(\lambda - Z)} \]
\[ y = \frac{\lambda Y}{(\lambda - Z)} \]

These are non-linear equations. It is often convenient to express them in linear matrix form.

A point in cartesian world co-ordinate system can be represented as a vector in cartesian and homogeneous co-ordinates.

\[ w = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \text{or} \quad w_h = \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix} \]
PERSPECTIVE TRANSFORMATION
(cont. 1)

Perspective matrix

\[
P = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & -1/\lambda & 1
\end{bmatrix}
\]

\[
\mathbf{c}_h = P \mathbf{w}_h
\]

\[
= \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & -1/\lambda & 1
\end{bmatrix}
\begin{bmatrix}
kX \\
kY \\
kZ \\
-kZ/\lambda + k
\end{bmatrix}
\]

\[
= \begin{bmatrix}
kX \\
kY \\
kZ \\
-kZ/\lambda + k
\end{bmatrix}
\]

\(\mathbf{c}_h\) elements are the camera coordinates in homogeneous form.
The cartesian coordinates of any point in the camera coordinates system are:

\[
\begin{align*}
| x | &= | \lambda X/\lambda - Z | \\
| y | &= | \lambda Y/\lambda - Z | \\
| z | &= | \lambda Z/\lambda - Z |
\end{align*}
\]

The \textit{z} component is of no interest in terms of image formation model.

It acts as a free variable in the inverse perspective transformation.

The \textit{Inverse Perspective Transformation} maps an image point back into 3-D

\[
\mathbf{w}_h = \mathbf{P}^{-1} \mathbf{c}_h
\]
PERSPECTIVE TRANSFORMATION
(cont. 3)

\[
\mathbf{P}^{-1} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 1/\lambda & 1 \\
\end{bmatrix}
\]

An image point has co-ordinates \((x_0, y_0, 0)\) where image plane is located at \(z = 0\).

This point in homogeneous vector \(\mathbf{c_h}\)

\[
\mathbf{c_h} = \begin{bmatrix}
kx_0 \\
k y_0 \\
0 \\
k \\
\end{bmatrix}
\]

\[
\mathbf{w_h} = \mathbf{P}^{-1} \mathbf{c_h}
\]

\[
= \begin{bmatrix}
kx_0 \\
k y_0 \\
0 \\
k \\
\end{bmatrix}
\]
PERSPECTIVE TRANSFORMATION
(cont. 4)

In cartesian coordinates

\[
\begin{vmatrix}
X \\
Y \\
Z
\end{vmatrix}
= \begin{vmatrix}
x_0 \\
y_0 \\
0
\end{vmatrix}
\]

It gives \( Z = 0 \) for any 3-D point

☛ Problem

● The problem here is caused by mapping a 3-D scene onto the image plane.

● It is a many-to-one transformation.

● The image point \((x_0, y_0)\) corresponds to all the collinear 3-D points that lie on the line passing through \((x_0, y_0, 0)\) and \((0, 0, \lambda)\)
The equations of the line joining \((x_0, y_0, 0)\) and \((0, 0, \lambda)\) are:

\[
X = \frac{x_0}{\lambda}(\lambda - Z) \\
Y = \frac{y_0}{\lambda}(\lambda - Z)
\]

They show unless some-thing is known about the 3-D scene point, it is not possible to completely recover the 3-D point from its image.

» This is the basic computer vision problem which every one in this area is researching on.

» There are some limited solutions.
COMPUTER VISION

The purpose of computer vision is to enable a computer to understand its environment from visual information i.e. images using (grey level, colour, shades, texture etc)

Computer vision is also known as *Machine Vision* and related to:

- Image Processing and Image Analysis
- Image Understanding
- Pattern Recognition
COMPUTER VISION PROBLEM

Given a 2-D (two-dimensional) image, infer the objects that produce it.

Image I(x, y) is a function of Gray level at each point in the image.

\[ I(x, y) \rightarrow \text{Object Information's} \]
\[ 2-D \rightarrow 3-D \text{ (three dimensional)} \]

- It is a transformation from 2D to 3D
- A many-to-one transformation. Many solutions.
- How to identify the correct solution?
- Third dimension (depth) is missing
- We can not completely recover the third dimension from a 2-D image.
COMPUTER VISION PROCESSES

● Sensing

How to get a digital image?
Image Sampling and Quantization

  ● Video or a TV camera
    (Monochrome and Colour Images)
  ● IR camera (Thermal Image)
  ● Range finder/Laser etc.
    (Range Images)

● Preprocessing

  ● Noise reduction
  ● Filtering
  ● Smoothing
  ● Enhancement etc.

■ Frequency Domain Methods
  (also known as Image Processing)
■ Spatial Domain Methods

● Segmentation

  Partitioning an Image into Objects of Interest.
COMPUTER VISION PROCESSES
(cont. 1)

● Shape Extraction
  Computation of Depth Information by
  ● Shape from Contours
  ● Shape from Shading
  ● Stereoscopic Imaging
  ● Shape from Motion
  ● Shape from Texture

● Description
  Computation of Features

● Recognition
  Identifying the Objects

● Interpretation
  Assigning meaning to a group of recognised objects
CLASSES OF VISUAL PROCESSING

Three Levels of Visual Processing

They are based on the sophistication of implementation and widely known as

Low, Intermediate and High level Vision

● Low Level Vision

It involves preprocessing and low level feature extraction

● Requires no intelligence

● Involves only the data driven processes
CLASSES OF VISUAL PROCESSING
(cont. 1)

● Intermediate Level Vision

It involves image segmentation and image analysis type processing. Raw shape feature extraction processes are also its part.

  ● Very important processing stage
  ● It combines the data driven and goal driven processes

● High Level Vision

It involves the goal driven processes consisting of model and knowledge based vision.

  ● Recognition
  ● Interpretation
IMAGE SEGMENTATION

The Image Segmentation Process segment an image.

Segmentation:

* Subdivide an image into its constituents parts/objects.

* The level to which this subdivision is carried out depends on the problem being solved.

Constituent Parts:

* Object Boundaries

* Uniform Intensity or colour regions

* Lines, Points and Curves etc.

Image Segmentation is a very important part of Computer Vision