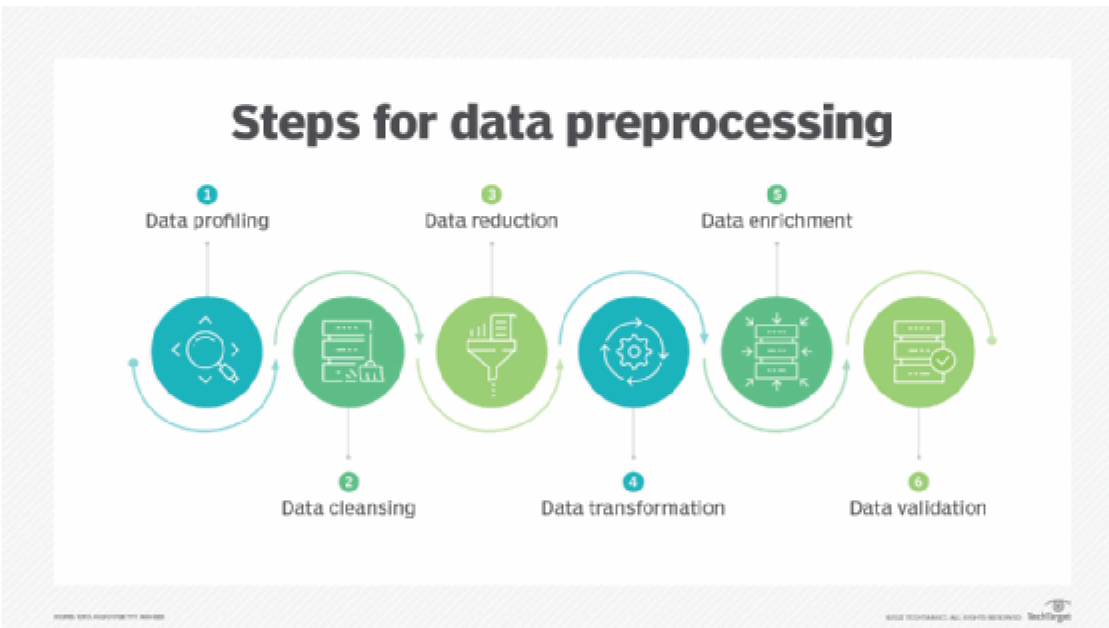
UNIT-3 Image Processing

1. **Image Pre Processing**

Image preprocessing is quite useful to improve the quality of images and thus boost them for analysis and further processing. Some powerful image preprocessing techniques include noise reduction, contrast enhancement, image resizing, color correction, segmentation, feature extraction, etc.

Image preprocessing is the steps taken to format images before they are used by model training and inference. Image preprocessing may also decrease model training time and increase model inference speed.

Methods of Image Pre-Processing



Introduction to Image Pre-Processing

Pixel brightness transformations(PBT)

Gamma Correction

Histogram equalization

Sigmoid stretching

Geometric Transformations

Image Filtering and Segmentation

Image Segmentation

Types of image processing methods

Common image processing include image enhancement, restoration, encoding, and compression.

Data preprocessing techniques might include:

For example, converting your colored images to grayscale images. This is because in many objects, color isn't necessary to recognize and interpret an image. Grayscale can be good enough for recognizing certain objects.

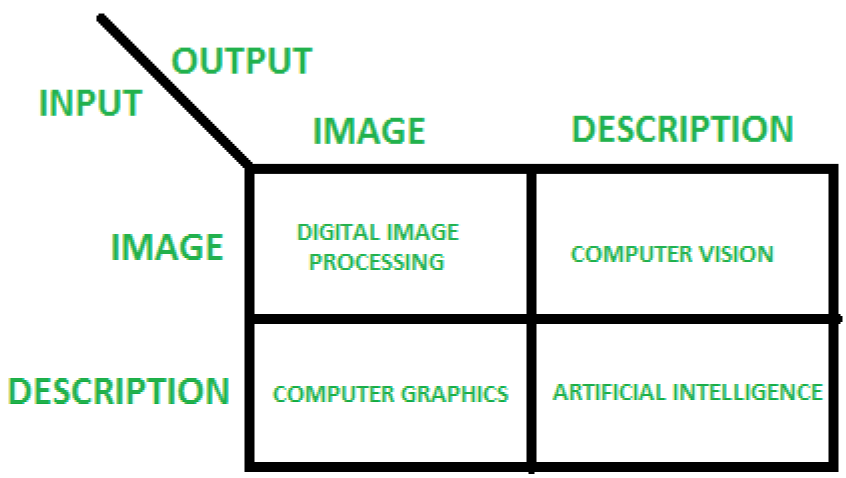
Image processing basically includes the following three steps:

Importing the image via image acquisition tools;

Analysing and manipulating the image;

Output in which result can be altered image or report that is based on image analysis.

Improved image quality: Digital image processing algorithms can improve the visual quality of images, making them clearer, sharper, and more informative. Automated image-based tasks: Digital image processing can automate many image-based tasks, such as object recognition, pattern detection, and measurement



The widely used algorithms in this context include denoising, region growing, edge detection, etc. The contrast equalization is often performed in image-processing and contrast limited adaptive histogram equalization (CLAHE) is a very popular method as a preprocessing step to do it.

Image processing techniques for machine learning.

a) Image Restoration.

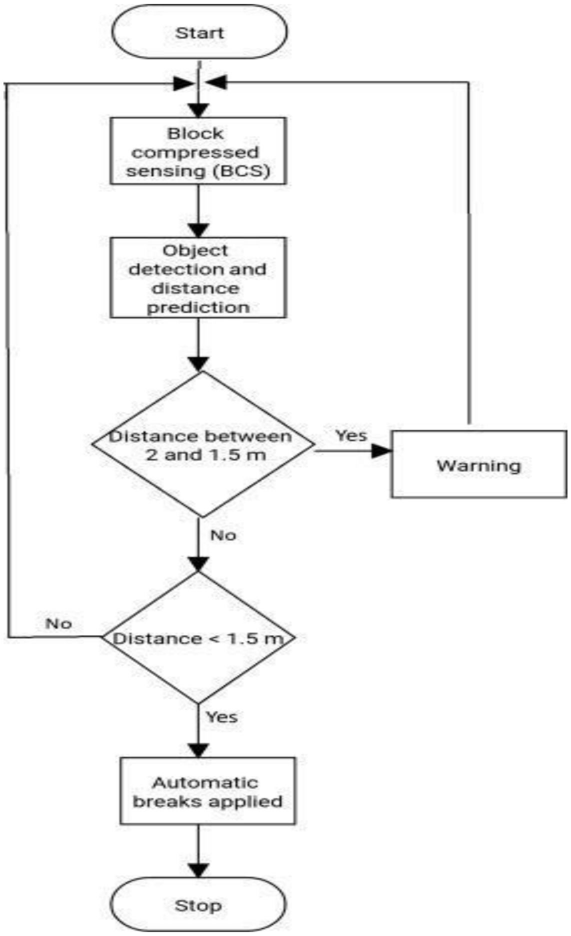
b) Linear Filtering.

c) Independent Component Analysis.

d) Pixelation.

e) Template Matching.

f) Image Generation Technique (GAN)



Feature plays a very important role in the area of image processing. Before getting features, various image preprocessing techniques like binarization, thresholding, resizing, normalization etc. are applied on the sampled image.

Image Representations

Images can be represented as a function of two variables, X and Y, which define a two dimensional area. Digital images are made of a grid of pixels. The Pixel is the raw building block of an image. The representation of an image can take many forms. Most of the time, it refers to the way that the conveyed information, such as color, is coded digitally and how the image is stored. Methods proposed for image representation range from color histogram to feature statistics, from spatial frequency to region-based, and from color-based to topology- based. The statistical properties in physical level usually grasp semantics in difficulty. Bitmap image (or raster image): can represented by our image model I(r, c), where we have pixel data and corresponding brightness values stored in some file format. **2.** Vector images: refer to the methods of representing lines, curves shapes by storing only the key points.

Continuous & Discrete Representations

Discrete data is a numerical type of data that includes whole, concrete numbers with specific and fixed data values determined by counting. Continuous data includes complex numbers and varying data values measured over a particular time interval.

Edge Detection

Edge detection works by detecting discontinuities in brightness. It is used for image segmentation and data extraction in areas such as image processing, computer vision, and machine vision. Common edge detection algorithms include Sobel, Canny, Prewitt, Roberts, and fuzzy logic methods.

Edge detection is a technique of image processing used to identify points in a digital image with discontinuities, simply to say, sharp changes in the image brightness. These points where the image brightness varies sharply are called the edges (or boundaries) of the image. The four steps are Image smoothing, Enhancement, Detection and Localization. Type edge-detection approaches can be broadly classified First derivative / Gradient based edge detectors, Second derivative / Zero crossing (Laplacian) based edge detectors and optimal edge-detector.

The Canny Edge Detection algorithm is a widely used edge detection algorithm in today's image processing applications. It works in multiple stages as shown in fig 1.2. Canny edge detection algorithm produces smoother, thinner, and cleaner images than Sobel and Prewitt filters.

This technique is employed after the image has been filtered for noise (using median, Gaussian filter etc.), the edge operator has been applied (like the ones described above, Canny or Sobel) to detect the edges and after the edges have been smoothed using an appropriate threshold value.

Kirsch Compass Masks

Kirsch Compass Mask is also a derivative mask which is used for finding edges. Kirsch mask is also used for calculating edges in all the directions.

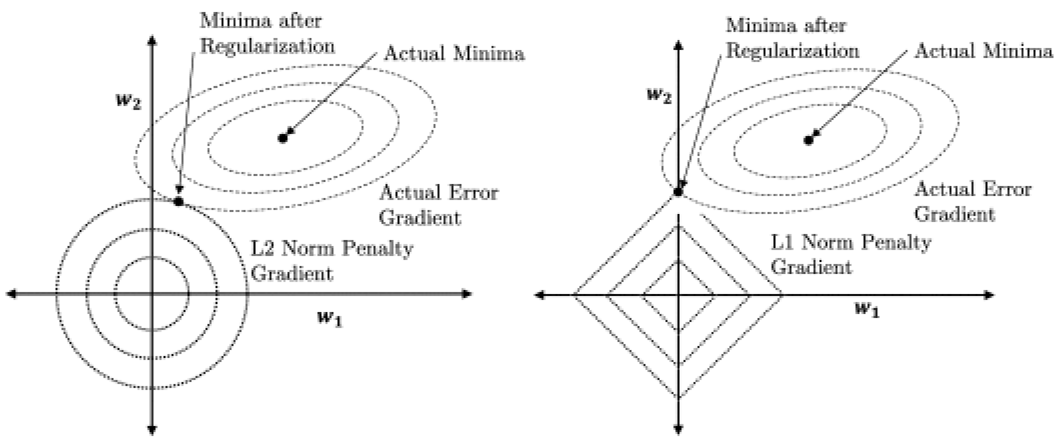
The most common method for edge detection is to calculate the differentiation of an image. The first-order derivatives in an image are computed using the gradient, and the second-order derivatives are obtained using the Laplacian. Another method for edge detection uses Hilbert Transform. Of these kernels, Sobel convolution kernels are used for horizontal and vertical edge detection. They are listed as follows: Horizontal Sobel edge detector: (13.12) − 1 − 2 − 1 0 0 0 1 2 1 .

Depending on the type of transition, there are three different types of edge detection: rising edge: when the input signal is transitioning from a low state (e.g. 0) to a high state (e.g. 1) falling edge: when the input signal is transitioning from a high state (e.g. 1) to a low state (e.g. 0)

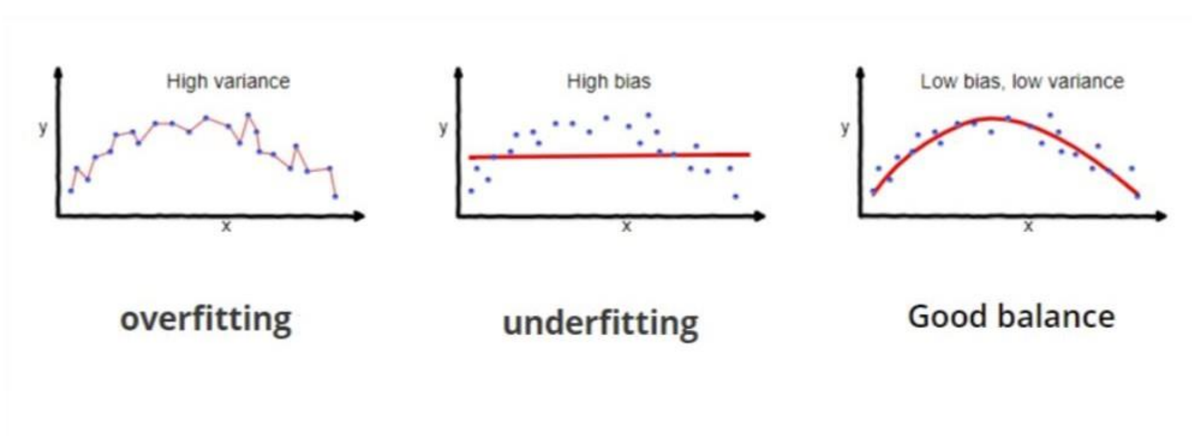
Commonly used metrics are MSE, PSNR and SSIM. MSE is basically a weighted function of deviations in images, or square difference between compared images. The Sobel operator performs a 2-D spatial gradient measurement on an image and so emphasizes regions of high spatial frequency that correspond to edges. Typically it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image. The algorithm used for mask detection is the Residual Neural Network (ResNet 50) algorithm for feature extraction. The three types of edges in digital images processing, based on intensity changes. Namely, step edges, ramp edges and edges noise. An edge is defined as a set of pixels where there is an abrupt change in colour intensity over distance. Laplacian filter is a second-order derivative filter used in edge detection, in digital image processing. In 1st order derivative filters, we detect the edge along with horizontal and vertical directions separately and then combine both. But using the Laplacian filter we detect the edges in the whole image at once.

Regularization Theory

Regularization theory studies methods for the solution to ill-posed problems (i.e., problems for which at least one of the conditions of uniqueness, existence, or continuous dependence of the solution on the data is not ensured). Regularization is a technique that helps prevent overfitting by penalizing a model for having large weights. Essentially, a model has large weights when it isn't fitting appropriately on the input data. During the L2 regularization the loss function of the neural network as extended by a so-called regularization term, which is called here Ω. The regularization term Ω is defined as the Euclidean Norm (or L2 norm) of the weight matrices, which is the sum over all squared weight values of a weight matrix. L2 regularization takes the square of the weights, so the cost of outliers present in the data increases exponentially. L1 regularization takes the absolute values of the weights, so the cost only increases linearly. LASSO regression, also known as L1 regularization, is a popular technique used in statistical modeling and machine learning to estimate the relationships between variables and make predictions. LASSO stands for Least Absolute Shrinkage and Selection Operator.



The difference between ridge and lasso regression is that it tends to make coefficients to absolute zero as compared to Ridge which never sets the value of coefficient to absolute zero.

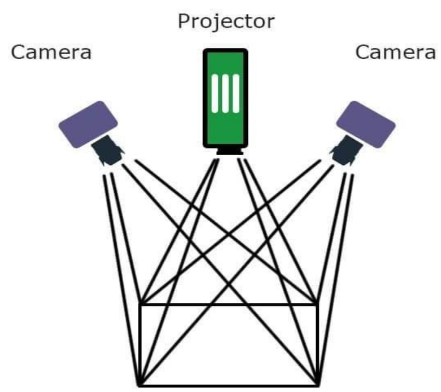


Optical Computation

Optical computing (also known as optoelectronic computing and photonic computing) is a computation paradigm that uses photons (small packets of light energy) produced by laser/ diodes for digital computation. Photons have proved to give us a higher bandwidth than the electrons we use in conventional computer systems. OC may be divided into two types: digital optical computing (DOC) and analog optical computing (AOC). Optical computing makes use of photons to utilize the wave interference pattern as well as wave propagation to identify outputs. By this method, optical computing enables immediate, error-free calculating. As the data is transferred, it gets processed. The theoretical advantages of optical computing are numerous, including greater energy efficiency, potentially much higher computing speeds, and greater information storage.

Stereo Vision

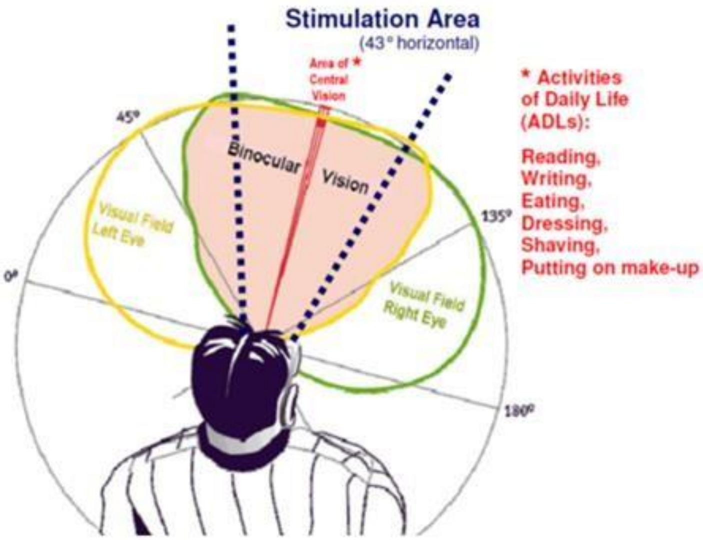
Stereo vision is the process of extracting 3D information from multiple 2D views of a scene. Stereo vision is used in applications such as advanced driver assistance systems (ADAS) and robot navigation where stereo vision is used to estimate the actual distance or range of objects of interest from the camera. A stereo camera is a type of camera with two or more image sensors. This allows the camera to simulate human binocular vision and therefore gives it the ability to perceive depth.



There are two types of stereoscopic vision: monocular and binocular. Monocular vision is when each eye sees a separate image. This is what people with normal vision use. Binocular vision is when each eye sees the same image.

Advantages of Stereoscopic Vision

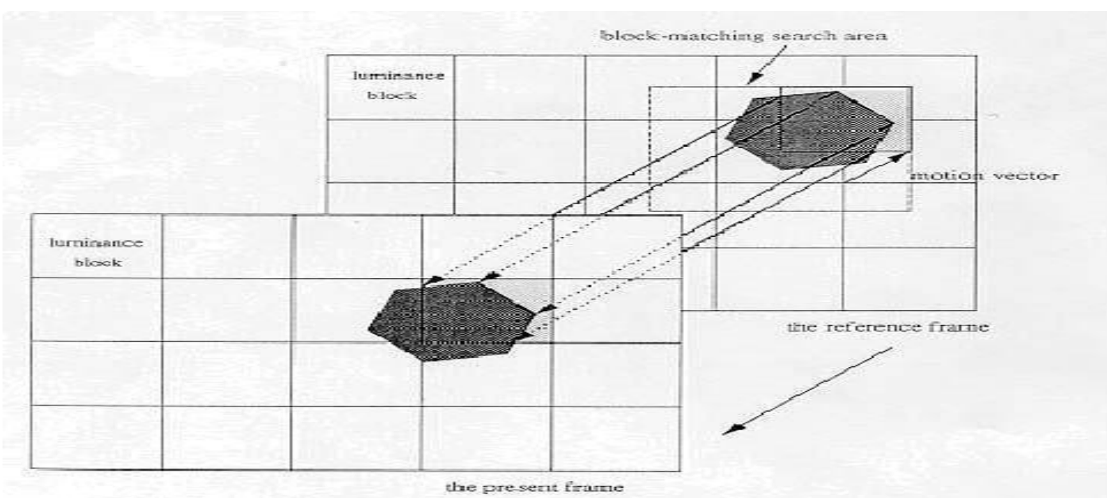
With the help of stereoscopic vision, humans can manage to handle small objects. It helps to reciprocate threats and react accordingly. Provides a deep sense of perception. It helps to achieve accuracy in various profiles like the manufacturing industry. It is also used in industrial automation and 3D machine vision applications to perform tasks such as bin picking, volume measurement, automotive part measurement and 3D object location and identification. Depth perception – the ability to see our environment three-dimensionally – is based on “Stereoscopic vision”, which requires two intact eyes that are aligned and move synchronously. Try the following: place a round patterned object, e.g. a painted coffee cup, at arm's length on the table in front of you. Stereopsis is the visual perception that deals with the sensation of depth from the two slightly different projections on human eyes. The differences in the two images (for each eye) are called horizontal disparity.



Motion Estimation

Affine motion estimation is a technique used in computer vision and image processing to estimate the motion between two images or frames. It assumes that the motion can be modeled as an affine transformation (translation + rotation + zooming), which is a linear transformation followed by a translation.

Successive video frames may contain the same objects (still or moving). Motion estimation examines the movement of objects in an image sequence to try to obtain vectors representing the estimated motion. Motion compensation uses the knowledge of object motion so obtained to achieve data compression.



Motion estimation is an important part of image analysis. Estimated motion vectors in image sequences may be used for e.g. motion detection, tracking, identification, segmentation, and 3d reconstruction. Moreover motion vectors may be used for motion analysis, e.g. detection of abnormal behavior. Basic Motion Estimation Techniques

Motion estimation is an important process in a wide range of disciplines and applications, such as image sequence analysis, computer vision, target tracking, and video coding. The methods for finding motion vectors can be categorised into pixel based methods ("direct") and feature based methods ("indirect").

The algorithm compares the values of two pixels occupying the same position in adjacent frames to determine the moving object area, sets a spatial sample set for each pixel, and defines the spatial sample difference consensus (SSDC), which represents the change of stable spatial relationship. They represent two different approaches to motion estimation (pel-based and block-based, respectively).

Block matching can be considered as the most popular method for practical motion estimation due to its lesser hardware complexity. As a result it is widely available in VLSI, and almost all H. 261 and MPEG 1-2 codecs are utilizing block matching for motion estimation.

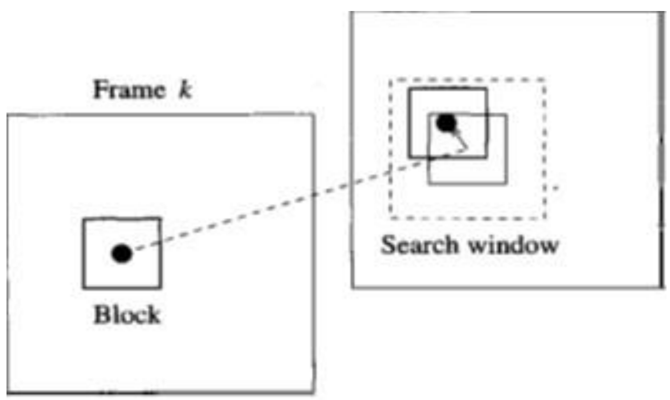


Figure 3 Block matching

3 types of movement in photography

Suspended movement. Perhaps the most obvious type of movement in photography, suspended movement illustrates one of the camera's most remarkable attributes: the ability to freeze a literal split second, to capture details imperceptible to the human eye.

Motion blur

Visual flow

The types of motion are:

Oscillatory motion.

Rotational motion.

Translational motion.

Uniform motion.

Non- uniform motion.

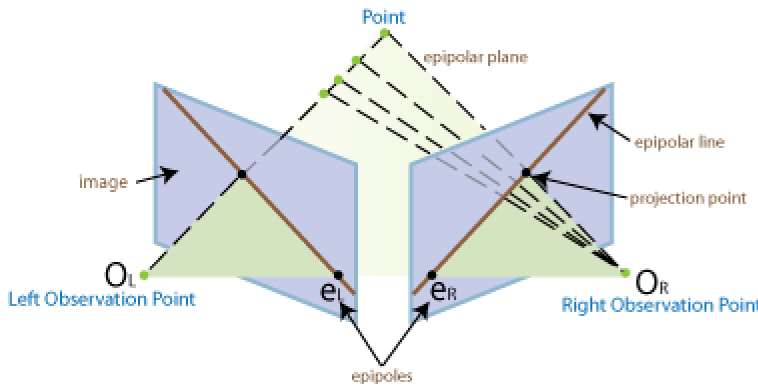
Periodic motion.

Circular motion.

Linear motion.

Structure from Motion

Structure from motion (SfM) is the process of estimating the 3-D structure of a scene from a set of 2-D images. SfM is used in many applications, such as 3-D scanning , augmented reality, and visual simultaneous localization and mapping (vSLAM). SfM can be computed in many different ways.



Structure from motion (SfM) is a photogrammetric range imaging technique for estimating three-dimensional structures from two-dimensional image sequences that may be coupled with local motion signals. It is studied in the fields of computer vision and visual perception. The structure from motion (SfM) problem in computer vision is the problem of recovering the three-dimensional (3D) structure of a stationary scene from a set of projective measurements, represented as a collection of two-dimensional (2D) images, via estimation of motion of the cameras corresponding to these images.

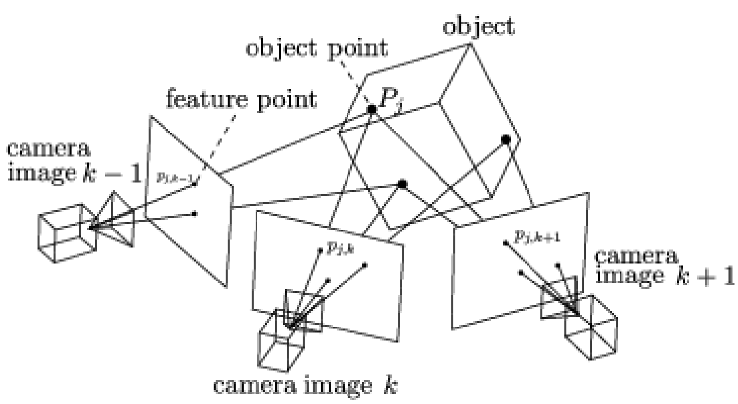
Motion is mainly described in terms of the following terms:

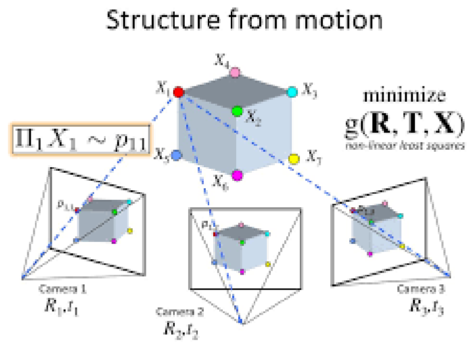
Distance.

Displacement.

Speed.

Time.





The main disadvantages of SfM include the possible deformation of the modeled topography, its over-smoothing effect, the necessity for optimal conditions during data acquisition and the necessity of a ground control point.