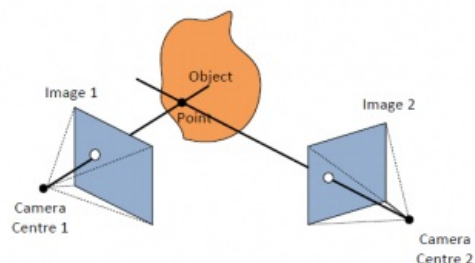


# Dense Image Matching



Point clouds are increasingly a prime data source for 3D information. For many years, Lidar systems have been the primary way to create point clouds. More recently, advances in the field of computer vision have allowed for the generation of detailed and reliable point clouds from images – not only from traditional aerial photographs but also from uncalibrated photos from consumer-grade cameras.

Dense image matching is the powerful technology underpinning this development.

(By Martin Kodde)

## Understanding photogrammetry

A good understanding of dense image matching requires insight into the way photogrammetry works. Photogrammetry in itself is not a new technology; it has been applied in practice for decades without many changes to its fundamental concepts. In photogrammetry, 3D geometry is obtained by creating images of the same object from different positions. This makes a single point on the object visible as a pixel in multiple

images. For each image, a straight line can be drawn from the camera centre through the pixel in the image. These lines will intersect at one point, which is the 3D location of the object point (Figure 1).

However, this requires the position and orientation of each image to be known. To this end, so-called tie points are used to link all the images together. Each tie point is a well-recognisable point that is identified in all images where it occurs. Sufficient tie points allow for the reconstruction of the relative position of all images. Additionally, known points or ground control points (GCPs) with 3D world coordinates should be added to obtain scale and absolute coordinates. Tie points and GCPs are combined in a bundle block adjustment, resulting in the 3D coordinates of all tie points and, more importantly, the position and orientation of each image.

## Finding corresponding points

In the old days of analogue aerial photogrammetry, tie points were physically identified by pinning small holes through the image at the tie-point location. When digital photogrammetry emerged, much of the manual labour was replaced by automated tie-point search software which can easily detect hundreds of reliable corresponding points in multiple images. Feature-based matching is often applied for this purpose. The algorithm attempts to detect well-recognisable features – such as a road marking, a building edge or any other strong change in contrast – in each individual image. Once all the features have been found, the algorithm proceeds to detect corresponding features in multiple images. This results in highly reliable corresponding points that are very suitable as tie points.

To obtain a dense point cloud, a corresponding point is needed for almost every pixel in the image. The feature-based matching approach is not suitable for this purpose since not every pixel in the image corresponds to a well-recognisable feature. Many pixels will represent a greyish surface of a road or pavement or a green patch of vegetation. These pixels cannot be automatically linked to a feature and will be missed by the feature-based matching approach.

## Searching row by row

Dense image matching takes an alternative approach to obtain a corresponding point for almost every pixel in the image. Rather than searching the entire image for features, it will compare two overlapping images row by row. Essentially, this reduces the problem to a much simpler one-dimensional search. This requires an image-rectification step before the matching starts. The images need to be warped in such a way that each row of pixels in one image corresponds exactly to one row in the other image, i.e. in technical terms the rows of the images should be parallel to the epipolar line. In the case of aerial images that are captured in long flight lines, there is usually a good correspondence between the rows so only a small correction is needed. Terrestrial and oblique images, however, may require significant adjustment to achieve this row-by-row property. From a computational point of view, this can be achieved with a simple perspective transformation. To the human eye, the resulting images can appear highly distorted (Figure 2).

Now, the algorithm can work row by row and pixel by pixel (Figure 3). For each pixel, it will search in the corresponding row for the pixel that is most likely to represent the same point in the real world. It will do this by comparing the colour or grey value of the pixel and its neighbours. At the same time, a constraint is set to ensure a certain amount of smoothness in the result. When a pixel is found in the second image that is a good match to the same pixel in the first image, the location of that pixel is stored. Once two corresponding pixels are known, traditional photogrammetric techniques can be used to compute the 3D intersection for the pixel.

# Semi-global Matching

The row-by-row approach to images is efficient but, since each row is handled entirely independently, there is a risk of a disconnect between the results. This effect is called streaking. To overcome this disadvantage, the semi-global matching approach was proposed. This method not only evaluates horizontally row by row but it also traverses the image in 16 different directions. This produces 16 matching results which are then combined in a weighted sum to achieve a final result that has much less noise. Furthermore, this approach may add other images that also have overlap for an even better result.

## Concluding remarks

There are many adaptations and alternatives to the matching approach presented in this article. Alternative implementations may improve memory efficiency, speed or reliability. Often, the algorithms do not store the corresponding pixels but rather the parallax between them as this is more memory efficient. Dense image matching is an essential technology for many recent innovations in the field of geoinformation. It is used to generate point clouds from aerial images, drone images, spherical images, etc. It can also increasingly be found in consumer-oriented products such as mobile apps for fast modelling of 3D objects.

*Figure 1, The principle of 3D point reconstruction from two images.*

*Figure 2, Two images rectified to create corresponding rows.*

*Figure 3, Searching row by row in two images.*