An overview of Image Registration/Matching System

- Introduction
- Pre-Processing
- Distortions and Errors
- Signal processing based Matching
- Artificial Intelligence based technique
- Change Detection
- Re-sampling
- SAR Image Registration and Matching

Introduction

Two images of the same region are said to be registered when equivalent geographic points of the scenes in the two images coincide. Image registration is a fundamental technique to match two images, which are acquitted at different time, different sensors, and different angles. The acquitting of image under different circumstances lead to misalignment of images, which needs to be corrected. Image registration is a very important image processing technique, which in general is independent of application of image processing. Some typical examples where image registration is required are as follows, tracking a target with real time image of the scene (target recognition), matching medical images acquired at different angles.

Image registration can be done by required combination of the following components

1) Feature Space	2) A search space
3) A search strategy	4) Similarity metric

The feature space extracts the information in the images that will be used for matching.
The search space is the class of transformations that is capable of aligning the images.
The search strategy decides how to choose the next transformation from this space, to be tested in the search for the optimal transformation.
The similarity metric determines the relative merit for each test. Search continues according to the search strategy until a transformation is found whose similarity measure is satisfactory.

Pre-Processing

Pre-processing is carried out on raw image so as to suppress the unwanted distortions and enhance the image features. Pre-processing techniques involve modifying an image to make it more like an ideal image. Pre-processing method involves three kinds of modification: **rectification, gray level modification, and sharpening/smoothing operations**.
Gray-level correction compensates for the non-uniformity of sensor sensitivity and contrast.
Sharpening and smoothing are important for edge detection and other types of feature extraction. This method of filtering the image is extremely useful when A.I based technique of image registration is used.
Rectification is a similarity matching process where geometry is made plan metric, i.e. where the two images of the same scene are transformed so that the size and shape of any object in one image is the same as the size and shape of that object in the other image. A reference image is provided by a reference platform, which in convention covers a larger area than the target area. The target image is generated by 2nd sensor on the on-board platform. The sensed image contains target or scene of interest only and is contained somewhere within the boundaries of the reference map. The problem is to rectify and correct the sensed image in such a way that accurate registration of all points within the sensed image (with respect to the reference map) is possible.

In real-time measurement, the data rate is very high, which forbids us to rectify every picture elements. In that case only certain fractions of sample are corrected fully geometrically and remaining samples are corrected using interpolation, re-sampling or statistical approximations.

Registration overview

The basic step for image registration is to match the sensed image with the reference image (usually covers a larger area). The sensed images generally have magnification or rotational error usually occurring from geometrical distortions. If a-priori geometrical distortion data or sensor attitude is known we can easily eliminate these errors.
The matching process begins with matching specific features of the sensed image with the reference image, namely lines, edge, vertices. The set of match points are found from the matching process which actually restricts the problems generated due to geometrical distortion, rotational errors.
Image matching algorithms includes correlation matching; symbolic matching and hybrid are used to locate the match point pairs on the sensed image. These match points are physical features and attributes who coordinates and detailing points are known with utmost precision.
After the match points are marked, mapping is done. The match points in general are distributed randomly in a mesh grid and they are used to develop a MAPPING MODEL. (Intuition – the norms and protocols of the model is used to do the mapping operation). The model in general is achieved through distortion functions which in general are piecewise polynomials.

Distortions

When an images of same scene are taken at different time, different viewing angles, sensor attitude data, different scanning conditions led to relative distortions in image. This distortion in general appears in form of geometrical and radiometric distortions.

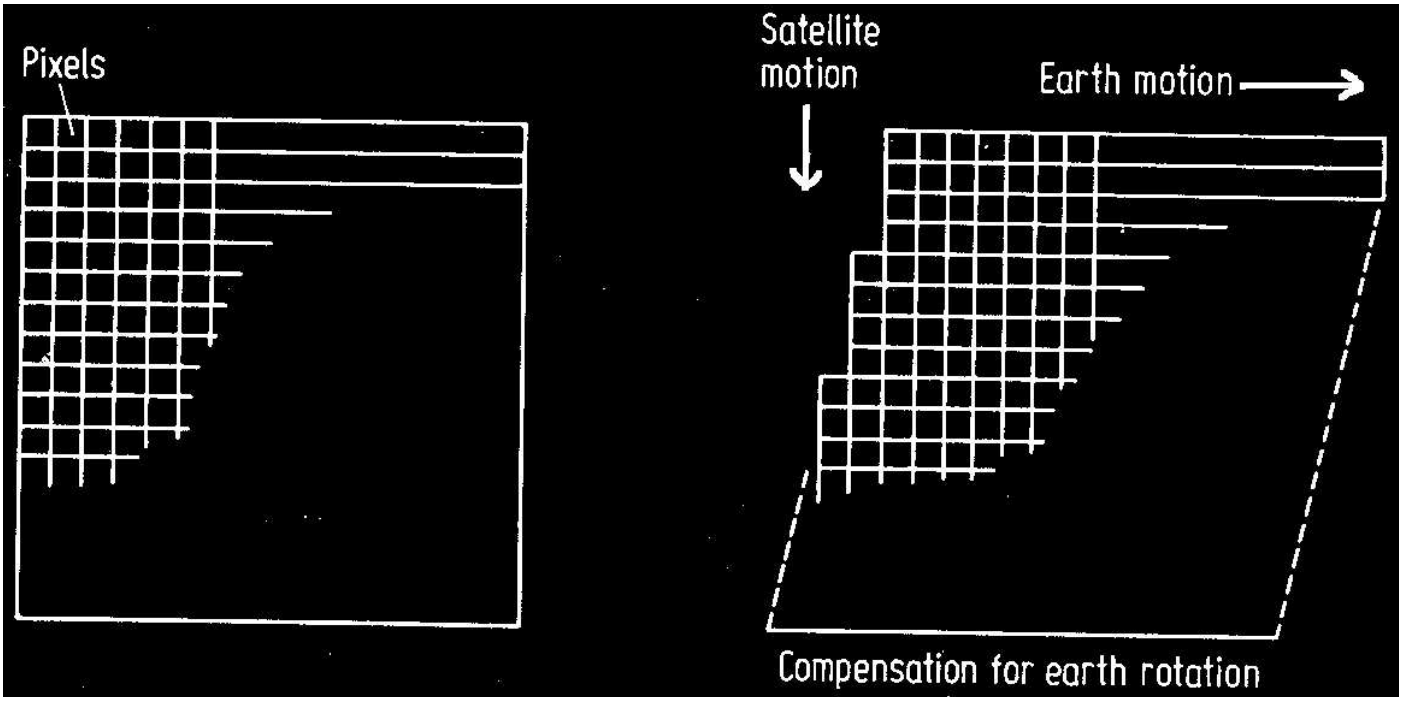
Radiometric distortion-
These distortions are caused by atmospheric and sensor induced filtering, sensor imperfections, camera or scanner shading effects, detector gain variations, and sensor detection gain errors. Radiometric distortions appear in the form of blemishes, horizontal stripes, shading, and non-uniform intensity distributions. Radiometric intensity corrections for non-uniform sensor sensitivity can be implemented in real -time using simple table lookup techniques.
Types of geometric distortion
1) Earth rotation 2) Panoramic distortion – Further affected by Earth curvature 3) Scan time skew
4) Platform variations of Height, velocity, attitude (roll, pitch, yaw).

Example of geometrical distortion (due to earth rotation)

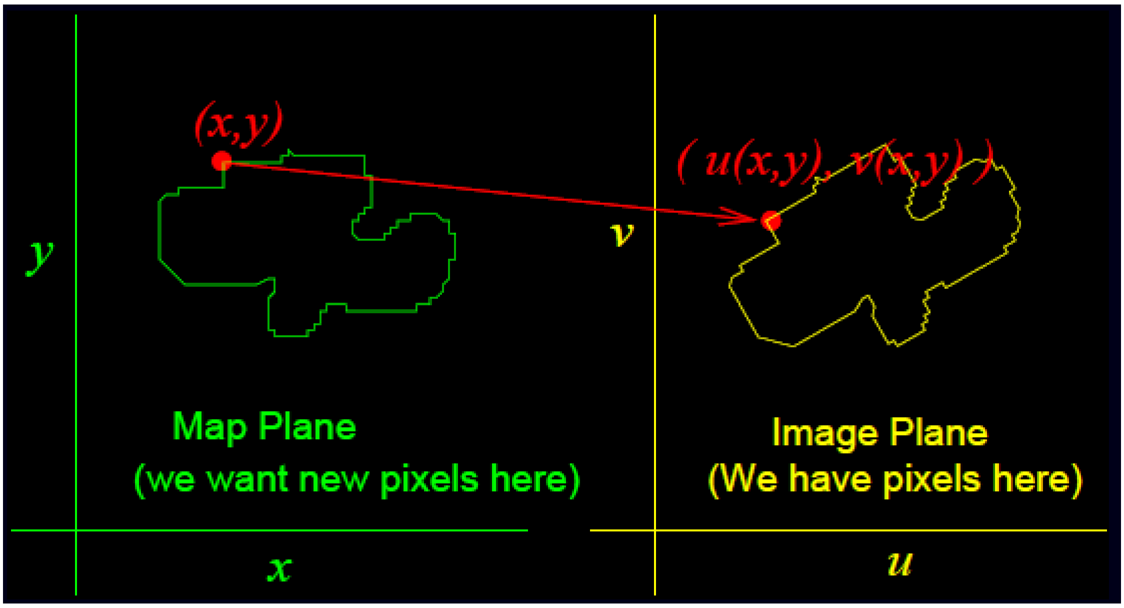
- Landsat MSS
 Standard image is 185 km^2 and consists of 2340 scan lines

Distortion of whole image is about 10 km.
 which is about 4.6 m per scan line

Earth Rotation



Distortion Correction



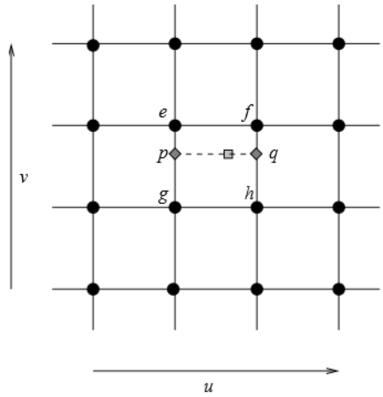
For each new pixel (x, y) in the map plane we find the equivalent point (u, v) in the image plane. To perform the operation we need to devise a mapping function.

We can perform the mapping function by interpolation or by strictly following a mapping function.

<u>**Interpolation**</u>

Nearest neighbour

In this method, the value from pixel nearest to (u, v) is marked into the figure where the new pixels are to be filled. Here the new pixels which are mapped from the original image is in (x, y) plane. This method is simple and fast.



Linear interpolation

In this method four pixels surrounding (u, v) are chosen and then interpolation is performed between 'e' and 'g' and considering it as p, and interpolating between 'f' and 'h' and considering it as 'q'. This methods brings on a smoothening effect on the image.

Bi-cubic splines

This method uses 16 pixels in the figure. A better result with much more smoothness is acquired. It has a problem of generating dark and spurious pixels.

Rectification algorithms

1) Signal Processing 2) A.I. Based 3) Hybrid method

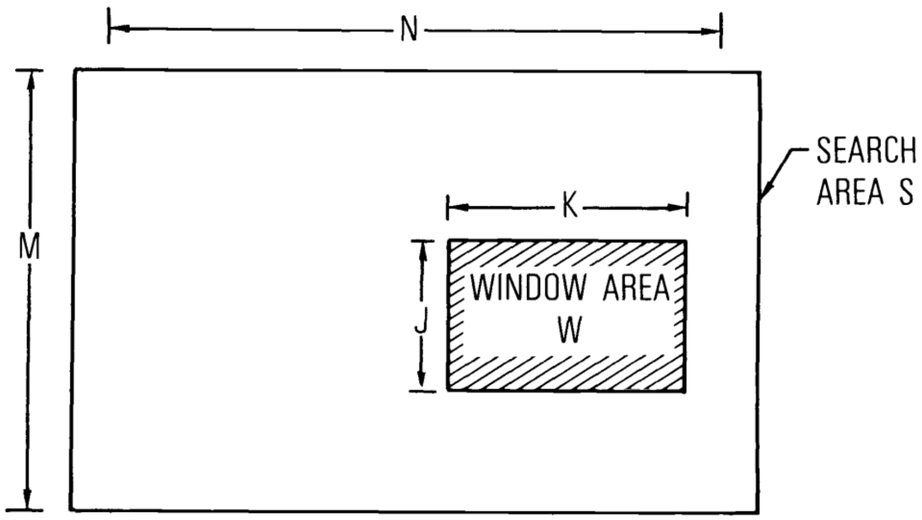
The signal processing based technique based on correlating the intensity value of the pixels of the sensed image and the reference image. The best matched image is a recorded on a basis which is recorded. A class of algorithms **SEQUENTIAL SIMILARITY DETECTION** provides a boost to conventional Correlation techniques in the task of finding the best image detection. One of the major advantages of using Correlation based method is that it doesn't fail even if substantial amount of noise is present in the image.This method possesses the property of noise suppression and find a good match provided the fact that relative magnification, distortion and rotational differences are not great.

In signal processing method, a very general technique of image registration is find correlation between the images using sliding window method. Attaching a pictorial illustration alongside.

The general formula for 2-d correlation coefficient measure is given by

$$R(m,n) = \frac{\sum_{j=1}^{J} \sum_{k=1}^{K} F_1(j,k)\, F_2(j-m,\ k-n)}{\left[\sum_{j=1}^{J} \sum_{k=1}^{K} F_1^2(j,k)\right]^{1/2} \left[\sum_{j=1}^{J} \sum_{k=1}^{K} F_2^2(j-m,k-n)\right]^{1/2}}$$

Referring to the image below,



F1(j,k) and F2(j,k) are two images to be registered. If both the images F1 and F2 match in same region inside the search areas, then images are considered to be registered. The window area starts from top left hand corner and slides row-wise and the size if the window area is j*k, where the search area is m*n.
Therefore total number of search windows created is (m-j+1)*(n-j+1). One of the search areas matches the sensed image. One of the major problems of correlation based technique is huge computation is required, i.e. in this case 2*(m-j+1)*(n-k+1)*j*k in worst case scenario. Another major problem is redundancy problem. Suppose the search area and the window area contains large region of uniform content and less feature, then matching will become hard to perform. Suppose the image of a barren field is to e matched and the search area contains uniform region over large area and the window region which needs to be matched is also an image of similar description. Then matching of image becomes very difficult.
In order to avoid this problem, a image mismatch checking method is used (derived from sequential similarity detection) is, in this method error minimisation is dictates whether the image is matched or not.

$$\varepsilon(m,n) = \sum_{j} \sum_{k} \left| F_1(j,k) - F_2(j-m,\ k-n) \right|$$

The window with minimum error explains whether the image is matched.

A.I. based technique

In the A.I. based technique, the importance is given to scene features like as edges, boundaries, erected structures, line markers, symbols. Then distance functions between sensed and reference image features are defined and a search for the best match is made. Graph matching, Linear and non linear optimization techniques are used to execute the search operation.
2 types of techniques used in AI based method is
1) Clustering features and feature groups
2) Symbolic matching.
In clustering technique, cluster of features are extracted from both reference image and sensed images, which may be edge, boundaries, patterns etc. No specific labeling of features are available in the image, so cluster matching is done on correspondence based on clustering of joint histogram of features in 2 images. The clustering unit is basically increased by 'n' units or sort of swiped across the search area at a certain angle and direction and groups values corresponding to the identification of two features
with one another.
In the structural/ Symbolic matching technique is applied to specific scenes which are predefined and structure interpretation from image boundaries. A data structure system has been created as part of the ACRONYM model-based vision system involving five levels of graphs: "context," "object," "observability," "interpretation," and "picture".

Modern techniques uses scene matching systems that extract line and vertex features derived from the scene. Feature weighting, based on match point location and modeling "feature contents", is used to define geometrical transformations between reference and sensed models that lead to highly accurate match point location. An important part of this work is the use of data structures for regions, lines, vertices, and connected groups that contain parameters and links to other data elements. The technique is experimentally highly tolerant to differences in scene contrast, scale, and orientation differences.

Change Detection Algorithms

when an image of same area is taken at different time intervals say t1 & t2 from same sensor, the images are not exact to each other.

The above images are captured by same sensor and it clearly illustrates the difference in images. A very basic idea for change detection is

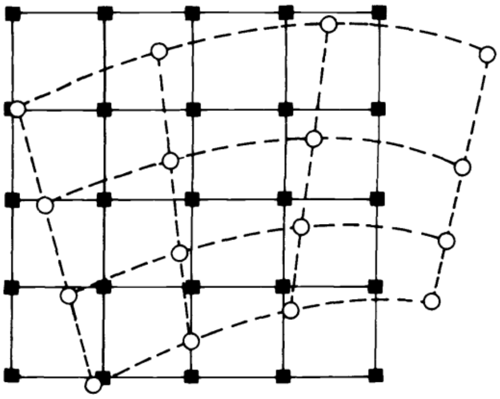
$$Id\ (x,\ y) = I1\,(x,\ y) - I2\,(x,\ y)$$

where I1 is the 1st image and I2 is the 2nd image. This change detection is applicable iff the images are spatially registered. Apart from pixel to pixel change detection, symbolic approaches are also available (in case of radiometric distortion). Symbolic techniques for change detection involves segmenting the images into regions or shapes with similar feature properties, and extracting symbolic descriptions of these regions. The symbolic descriptions are compared by Graph Matching, searching, etc. and it is similar to find the match point pairs. General problems faced by the change detection algorithms are differences in wavelength, illumination status, sensor orientation.

Re-sampling Technique

Image re-sampling is the process of transforming an image from one coordinate system to another. The two coordinate systems are related to each other by the mapping function of the spatial transformation. The inverse mapping function is

applied to the output sampling grid, projecting it onto the input. The result is a re-sampling grid, specifying the locations at which the input is to be re-sampled. The input image is sampled at these points and the values are assigned to their respective output pixels.
While registration of images, a problem arises because there were non-uniform spacing between the pixels. This non-uniformity arises along the scan line or across the scan line. suppose the image is scanned row-wise order , along the line is row-wise and across the scan line is column wise. This re-sampling method is necessary due to the fact of origin of radiometric and geometrical distortion. This results in bending of pixels, rather their spacing becomes non-linear. So it becomes necessary to re-shape the image in a linear order. An example of image which needs realignment of pixels into rectangular grid.



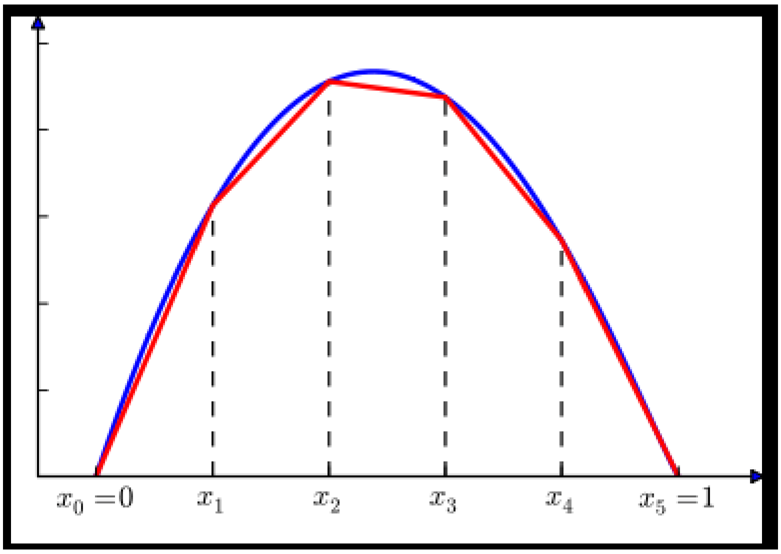
Re-sampling is done by copying intensity values of pixels (x, y) of distorted image to its ideal counterpart, keeping in mind the mapping is in accordance with certain protocols.
A pixel position (i,j) in an ideal image will usually not map to an integer coordinate (m,n)
the exact pixel position whose value is sampled in the distorted image but to a point (u,v) between pixel locations. There are two common solutions to this problem. One is to copy the value of the nearest neighbor to (u,v) into the ideal image. The other is to interpolate a value based on the values of pixels in a window around (u,v).

Resampling Techniques

Three techniques used for resampling 1-D or 2-d Data are 1) Nearest neighbour 2) Bilinear Interpolation

3) Cubic Interpolation

In the nearest neighbor resampling technique, the pixel value nearest to the resampled pixel is chosen for interpolation operation. This method is linear but it produces staircase response in the diagonals. The reason of this is because suppose pixel(1, 1) is to be resampled, then any of the nearest pixels (1, 2), (2. 1), (2, 2) will be used, and similar method is followed and carried on over rest of the image. There occurs a possibility of loss of pixel value is final image, hence giving a staircase response.
The bilinear resampling technique requires 4 closest pixels for evaluating the value of the pixel which is required to resampled. It uses a weighted average of the four nearest cell centres. The closer an input cell centre is to the output cell centre, the higher the influence of its value is on the output cell value. This means that the output value could be different than the nearest input, but is always within the same range of values as the input. The bilinear transformation brings a smoothened effect on the picture but many edge marking and highlight points are lost in the process. Instead, it should be used for continuous data like elevation and raw slope values.
Spline convolution is an approximation of resampling method by fitting piecewise polynomial model. An example of piecewise model is non-continuous plot drawn on a set of points.



Curve in the blue is the continuous plot and the one in red is a linear piecewise approximation. (Cubic piecewise model can be implemented instead of linear piece-wise model, but the complexity of implementation increases). Here, though, fitting cubic piecewise model is necessary for approximation.

The contribution of the spline kernel to the approximation method is twofold: (a) reduction of the summation term in the ideal interpolation from an infinite number to summation over four input sample pixels, and (b) optimization in terms of zero mean error and minimum error variance.

For 2-d image processing, resampling can be done in 2 ways, either along the scan line, or across the scan line. The difference is only either it is done row-wise or done column wise. The advantages of nearest neighbor technique and bilinear convolution over cubic interpolation is overshooting of data and exaggeration of peaks. Unpleasant artifacts and other effects still exists in spline convolution. Still, cubic convolution is preferred over them due to certain factors like computing time, statistical factors, frequency content.

Errors

Errors exist in unprocessed images and they contribute to existing distortions when compared to the refined reference image. In general four categories of error exist and they are namely global, regional, local and unstructured errors.

Global errors affect the image in overall, affecting all the points and elements of the image and can be removed by general rectification method. These errors are generally some sort of bias added on to image.

Regional error affects certain region of the image uniformly. These are generally scattered throughout the image and errors appear in a clustered form. An example of this error is contrast reversal, such that relative order of contrast between two images to be compared is not preserved. This can occur in IR images acquired at different times of the day or in cases where a SAR image is compared with a visible image of the same scene. A significant variation in contrast may lead to grossly incorrect matches. Edge matching and cluster reward algorithms play an important part to remove this sort of error.

Local errors affects pixels independently each pixels within a small region. An example of this is speckle and additive noise, which constructs this error. The spatial size of the speckle is generally on the order of the limiting resolution of the coherent SAR imaging system, and this signal- dependent coherent noise makes rectification and restoration of SAR images more difficult.

The last group of errors is non-structured errors which constitute the errors which are unprecedented and randomized and their occurrences are inevitable. For example cloud shading and chemical content of atmosphere constitutes unstructured error.

Things not covered

1) Distortion modelling
2) SAR Image Registration and Matching (Almost similar to what explained above)