CHAPTER 2: LITERATURE REVIEW

Image matching and registration is a very challenging area in the field of image processing. This work aims to study various algorithm related with image matching and registration. In this section, a comprehensive survey on various image matching and registration methods have been presented. Many feature based registration methods are available in literature. Authors of different papers try to find and match image features such as points, lines, corners and objects between reference and sensed images.

Many types of transformation is to be applied over the sensed image and transformed sensed image is matched with the reference image and some parameters needs to be calculated and then depending on the minimum and maximum value, optimizer method will decide the iteration and stopping criterion. Transformation can be rotation, scaling, translation, projective, shear and affine. Parameters which can be calculated may be normalized cross correlation, increment sign correlation, M estimator correlation and mutual information. Optimization techniques can be based on neural network, fuzzy, Nelder Mead or particle swarm. Optimization is the selection of a best alternate from the available alternatives. In other words, it consists of maximizing and minimizing a real function by choosing appropriate input values.

Figure 2.1 A survey of Image Matching and Registration

2.1 License Plate Detection, Segmentation and Registration

In previous years, many researchers have implemented several algorithms for License Plate (LP) Detection, Segmentation and Recognition in different conditions, but still search for robust algorithm is not over. There are many challenges in this area like images having complex background, noise, poor illumination, wrong capturing angle, low resolution camera, motion blur and LP of varying styles, where many algorithms fails to recognize the license plate.

Feature extraction for an offline license plate recognition system based on global features is suggested by Mohindra and Shukla [1]. They have tested their algorithm on 100 real time images captured at different times in a day. They have used edge detection method, connected component analysis and discrete radon transform method.
for image segmentation. The efficiency of recognizing each character is 95% and total time taken for one image recognition is 15 seconds.

Edge finding method and window filtering is suggested by Kranthi et al. [2]. Edge finding method is not giving significant output when image is having complex background because it is very sensitive to unwanted edges. So windowing method where intensity summation in horizontal and vertical direction is calculated that will provide proper detection of license plate. For object enhancement they have used scale down, sorting of intensity pixel and then multiplication of top 20% pixel by 2.55, which provide better result compared to histogram equalization technique.

Abbas et al. [5] suggested vertical edge based car license plate detection method. For low resolution image of size 352 * 288, in which they have used adaptive thresholding technique followed by vertical edge detection using 2 * 4 mask and proved that its performance is faster (47.7 ms) than the sobel and getting 91.4% correct detection. Advantages are like able to process low resolution image and with complex background, tolerance to lighting, tilt, varied sizes and designs of LPs. But their algorithm is highly complex.

Adaptive thresholding using the integral image described by Bradley and Roth [11], this type of thresholding is suitable in spatial variations in illumination. Integral image is summation of corresponding row and column and then different local threshold is calculated based on neighboring pixel values. It is highly robust when illumination changes occur in the image.

Multi features approach has been used by Zhang [17] for robust license plate detection. Features like mathematical morphology, rectangle features (Aspect ratio, Area, Density), edge statistics and character features. They have also used character feature verification algorithm to determine final detection.
Nipa et al. [18] proposed a case where LP’s are captured at an angle, in that skew correction is required and Harris corner detection is used to find features of license plate. For Harris corner detection thresholding, gradient, Gaussian smoothing and non maxima suppression is applied. Principle component analysis is used to extract only the meaningful data. It consists of mean, covariance matrix, Eigen value and Eigen vector calculation, finding maximum Eigen value and then corresponding Eigen vector is to be determined. Their algorithm works well for both positive and negative skewed images with less computation complexity. But the distance between camera and license plate image is very less.

Prathamesh et al. [19] describes Indian license plate localization approach based on features. They have used different binarization techniques like Otsu, Niblack’s and Adaptive thresholding method and then counting character pixels. In this approach, license plate at any corner of image can be detected. The main disadvantage of this method is when number plate is tilted then it is difficult to localize LP.

Intrinsic rules saliency is used for automatic license plate localization by Paunwala and Patnaik [21]. Firstly, candidate region identification using adaptive edge density and connected component analysis is done, then three functions are determined, one is density variance measurement for uniformity check, second is scan line variance for measurement of transition variance and third is edge angle analysis for identification of skewed text region. From these three functions, final F is calculated and if value of F is higher than some predefined threshold then it is considered as LP otherwise it is discarded. They got robust result on a mixed data set with high diversity. This method has advantages like detection box have high precision and it is independent of LP style, scaling, rotation and partial occlusion.

Paunwala and Patnaik [22] suggested a novel approach for multiple license plate extraction for the images captured at day and night
time. They are finding variance of the given image and if the variance is greater than some predefined threshold value then contrast enhancement is not required as that image was taken during daytime otherwise contrast enhancement is required which they have used is sigmoid function for the dark image. They got much higher success rate (around 99 %) because they have combined algorithm based on edge density, morphology, horizontal and vertical projection, Gaussian analysis, rectangluarity, aspect ratio analysis and finally plate companion filter is applied to detect only LP region and non LP region is discarded.

Epshtein et al. [23] discovered new algorithm for text detection using stroke width transform, in which edge detection, stroke width calculation, finding letter region, filtering and combining words are the steps followed for text detection in natural image. This algorithm is suitable for handwritten text detection, for the text varying in size, different language text and for the text having different orientation. Limitation of this algorithm is when the letters are very small and close to each other then in will be group together in stroke width labeling phase.

To detect text, Jin Lee [28] proposed an effective machine learning based algorithm for classifying two or more classes known as Adaboost. This algorithm constructs a strong classifier from a combination of weak classifiers. They have used 6 types of feature sets that are variance and expectation of XY derivatives, local energy of Gabor filter, statistical texture measure of image histogram, measurement of variance of wavelet coefficient, edge detection and edge interval calculation and connected component analysis. There are several boosting algorithms which will improve performance of Adaboost like real Adaboost, Gentle Adaboost and Modest Adaboost and performance of these methods are compared in terms of precision, recall and f-measure. Precision is the performance metrics – the
fraction of text window which are correctly classified as text. Recall – the fraction of all text windows which are correctly identified and f-measure is the harmonic mean of the precision and recall. For best performance, precision, recall and f-measure should be one.

License Plate detection algorithm using statistical features and LP templates are proposed by Kim et al. [30]. First, the entire image is searched and candidate areas based on the gradient features are located, the statistical features are calculated for region of interest (ROI), the candidates are examined to verify whether it contains LP and then LP templates are matched with ROI. Limitation of this algorithm is that it is difficult to construct general LP templates and it can work on a fixed scale.

Naito et al. [31] suggested Japanese license plate recognition method for passing vehicles under outside environment. They have used adaptive binarization, extraction of character region, hypothesis generation of registration numbers and then template matching for each hypothesis. Their algorithm is very efficient and able to process rotated and low contrast images. But the limitation of this method is license plate size should be fixed.

2.2 Feature based Image Matching and Registration Algorithms

Image features are unique image properties that can be used to establish correspondence between two images. Important features are points because their coordinates can be directly used to find the parameters of a transformation function which will register the images. In some images it may not be possible to detect point features. So lines or regions need to be determined in those cases.

Methods used for image matching and registration are classified into two broad categories which is Intensity based and feature based
method. Intensity based methods will compare intensity in two images and feature based methods will find correspondence between image features like points, contours and lines. Former category will register entire image or sub image while later category will establish correspondence between distinct points in image.

After comprehensive survey of existing feature based methods, there are essentially four steps that are feature selection, feature matching, determination of a transformation function and re-sampling. To register two images, a number of features are selected from the images and correspondence is established between them. The features used are corners, lines, curves, templates, regions and patches. The type of transformation function used should depend on the type of geometric difference between the images.

David Lowe [32] discovered a robust method for image matching. He established distinctive image features from scale invariant key points. The algorithm finds scale space extrema detection, key point localization, orientation assignment and key point descriptor. The method proposed by him is applicable for object recognition, matching and image mosaicing. This approach can robustly identify objects from clutter and occlusion and still giving near real time performance.

Mahesh and Subramanyam [35] show Scale Invariant Feature Transform (SIFT) algorithm for satellite images taken from Google earth. It detects features that are invariant to rotation, scaling, translation and illumination. Their algorithm finds scale invariant key points and Euclidean distance initial match to combine with Random Sampling Consensus (RANSAC) to achieve feature matching. The best candidate match for each key point is determined by its nearest Euclidean distance feature points from other images. They got very less root mean square error between the matched points after transformation.
Yan ke and Sukthankar [36] combined Principle Component Analysis (PCA) and Scale Invariant Feature Transform (SIFT) method, which is more distinctive representation for local image descriptors. They have evaluated their algorithm by calculating popular matrix called Receiver Operating Characteristics (ROC) and Recall and Precision. They have concluded that PCA is ill suited for representing the general class of image but it is very well suited for getting the variation in the gradient image of a key point that has been localized in scale space and orientation.

Krish et al. [38] suggested a new feature based image registration algorithm which finding rotation and scale invariant features and match them using generalized Hough transform and once feature correspondence has been established then estimated using non linear least squares and standard RANSAC algorithm.

One of the basic building blocks in any point based registration method is matching feature points which are extracted from a sensed image to their matching part in a reference image. Mount et al. [39] suggested solution for the fundamental problem of point matching. Given the two sets of points, find the affine transformation which transforms sensed image points so that its distance from the other reference image points is minimized. They measure distances using the partial Hausdorff distance. Point matching can be a computationally intensive task so they present two algorithms for point matching problem with an attempt to reduce computational complexity and still providing a good match. First algorithm is based on a branch and bound approach which show that by varying the approximation error bounds, it is possible to achieve a tradeoff between the quality of the match and the time required for the algorithm. Second method uses Monte Carlo method for accelerating the search process used in first algorithm.
Xiaolong Dai and Siamak Khorram [37] proposed a Feature-Based Image Registration Algorithm Using Improved Chain-Code Representation Combined with Invariant Moments. Their algorithm finds feature extraction using Thin and Robust zero crossing which includes selective zero crossing and edge refinement, contour search and sorting and region representation by chain codes and moments for both sensed image and reference image. Then initial image matching will start by finding minimum distance classifier based on combined criterion of chain code matching and invariant moment distance followed by initial transformation and consistency check in image space, control point extraction, transformation parameter estimation, and image re-sampling which results in registered image. The performance of this algorithm has been demonstrated by registering two multi temporal Landsat TM images taken in different years. In summary, the technique suggested by them is potentially powerful in terms of registration accuracy.

Guest et al. [84] a new correspondence algorithm is described called as CSM (Correspondence by Sensitivity to Movement) for accurate and robust correspondence calculations which is very important in medical and biological applications. This method is independent of the transformation and used in both 2D elastic registration algorithm for warping and a 3D rigid registration algorithm.

2.3 Intensity based Matching and Registration Algorithms

Intensity based techniques includes correlation based methods, Fourier methods and mutual information methods. Correlation based methods or template matching merge the feature detection and matching steps. Windows of predefined size or even entire image are used for correspondence estimation. The limitation of this algorithm is when the sensed image is deformed by more complex transformations;
this type of window is not able to cover the same parts of the scene. Area based methods may produce incorrect match in the smooth area of two images due to its non saliency. Normalized Cross correlation (NCC) method will match directly depending on image intensities so it is sensitive to change in intensities. The spatial coordinates where the maximum value of correlation is achieved is the matching position. Normalized cross correlation method is not suitable for multimodal images.

M – Estimator based image registration algorithm presented by K. V. Arya et al. [40] attempt to match template image with the target image. They used Huber’s estimator and Tukey’s bi square estimator for image registration and showed results for the images having occlusion and shadow. This method robustly registers images in the presence of noise and occlusion (artificially generated) up to 60%. It has been shown in that the proposed algorithm performs efficiently whereas other correlation based methods produce the incorrect registration. This allows one to get estimation procedures more robust to erroneous data. For instance, in the feature based approach, they are often false feature correspondence. This generally comes from a poor extraction of the features or from a bad matching of the features. M – Estimator method is not suitable when two input images are having high difference in rotation and scale.

Kaneko et al. [41] describes that ISC is formalized to be a binary or Gaussian distribution for a large image size through statistical analysis and modeling. They have made a statistic model with an analysis of its fundamental characteristics.

In [42], Fouad et al. focus on the geometric registration of images with disjoint local intensity shifts, especially due to large shadow differences when images captured at different time. Then in [43], Fouad author had extended their work of image registration under
illumination variations using region based confidence weighted M – Estimators.

Viola and Wells [50] suggested mutual information based image alignment. The method is based on mutual information between the two images. In this, entropy of two images, joint entropy and joint histogram is calculated. This is an intensity based method, it works well where edge and gradient based method fail to register and robust than correlation based method.

Chen et al. [52] proposed medical image registration for multi modal images (CT – MR Brain images) using generalized partial volume estimation (GPVE) technique. Their experimental results show that by properly choosing the kernel functions, the GPVE algorithm significantly reduces the interpolation induced artifacts and thus improving the registration accuracy.

Berthilsson [54] tried to register using cross correlation which is more global search method compared to others like steepest descent based methods. Algorithm is based on certain changes of coordinates in the image and Fast Fourier Transform that makes it faster.

Wolberg and Zokai [82] suggested hierarchical image registration where the affine parameters are computed iteratively using a variation of the Levenberg – Marquadt nonlinear least squares optimization method. In this paper, they are using Sum of Squared Difference (SSD) similarity measure for iterative estimation of perspective deformation. Every time, SSD value is compared with threshold, if SSD is higher than threshold then it is rejected and when SSD is minimum will be the best match. This method is less accurate than NCC but it is faster.

Huttenlocher et al. [83] described new method of similarity measure called Hausdorff Distance (HD). They registered binary images transformed by rotation and translation and compared HD based
method with classical CC based method and proved that HD is better than CC.

Mutual information (MI) based image registration can be done for multimodal images. MI which is originating from information theory is a measure of statistical dependency between two data sets. Ritter et al. [88] proposed hierarchical search strategy along with simulated annealing to find maximum mutual information.

### 2.4 Optimization Methods for Image Matching and Registration

Image matching and registration is an Optimization problem. Finding the maximum of similarity measure or minimum of dissimilarity measure is a multi-dimensional optimization problem. As number of dimension will increase, degree of freedom will decrease. Therefore it is becoming computationally demanding where exhaustive search over the entire image is required. Many optimization techniques are available such as Powell’s method, Downhill Simplex method, Brent’s method and series of one dimensional searches, Levenberg Marquardt optimization, Newton Raphson iteration, Stochastic search methods, Gradient descent methods, Genetic methods, Fuzzy methods, Simulated annealing, Geometric hashing, Nelder Mead Optimization, Neural Network based methods, Particle Swarm Optimization and Quasi exhaustive search methods and many more.

Meisen Pan et al. [59] suggested fuzzy based image registration in which Fuzzy C- Means (FCM) clustering is used and divided the coordinates of the pixels in the image into two clusters to fit a straight line, and then derived the slope of the line to compute the rotation angle of the image. Fuzzy method provides good result for mono modal and multi modal (with different sensors) images.
J. Senthilnath et al. [62] suggested multi objective function for fitness calculation. They have used distance condition and angle condition both because only distance condition cannot be effective for multi sensor images. These two conditions give better matching of the corner points between the two images. A set of three points between reference image and sensed image is required in the algorithm. In the past, RANSAC is used to match the points. Based on the results obtained to them, their algorithm is more efficient for multi sensor image registration. Registration accuracy is measured in terms of root mean square error.

Chen Lun Lin et al. [64] described Hybrid particle swarm optimization (HPSO) for medical image registration. HPSO includes two important concept subpopulation and crossover of genetic algorithm and conventional PSO and also provide comparison between HPSO, Genetic algorithm and conventional PSO.

Yong Fang Guo and Yi Cai Sun [66] introduced improved PSO for image matching and proved that this method is a simple but effective and reduce computation time. In [67], Yong Fang Guo have modified the method and calculated another objective function using distance transform and PSO and shown that it will again reduce time complexity of the algorithm.

Object detection can also be done using particle swarm optimization proved by Ankit Sharma and Singh in [70]. They have determined cross correlation coefficient between the images and applied PSO. Applications are in navigation, guidance, automatic surveillance and robot vision.

Finding point correspondence using Simulated Annealing (SA) explained by Starink and Baker [73] that tries to minimize a dissimilarity measure defined on point pairs. SA is a stochastic
optimization algorithm based on the physical analogy of annealing a system of molecules to its ground state.

Thevenaz et al. [74] explained iterative multi scale registration without landmarks in which they have combine Levenberg Marquardt method and sum of squared differences metric. They present an automatic sub pixel registration algorithm which minimizes the mean square difference of intensities between a reference image and sensed image. It uses coarse to fine pyramid strategy and performs minimization using Levenberg Marquardt method. They have used several images from PET and fMRI modalities and concluded that multi resolution refinement strategy is faster and more robust than single scale methods.