Chapter 2
Literature Review

2.1 Introduction

Summary of why it is important to adopt an accurate and affine invariant Image Registration method for building an Augmented Reality System: Robustness of Augmented Reality (AR) applications rolls around the accuracy of image registration procedure. Arbitrary viewpoints, different illumination conditions, image scale or even a wide baseline setup can hinder the appearance and detection of similar features in two or more images [Lee and Hollerer 2008]. Therefore, extracting affine invariant features in an image becomes the prior need of image registration. Many approaches have been proposed till date for extracting points of interest from an image that remain invariant under extreme changing conditions [Mikolajczyk and Schmid 2004, Lin et al. 2006, Li et al. 2012]. The approaches used for extracting these invariant features usually work in two modules, Module1: Feature Detector, which detects points of interest in an image that are covariant to a class of transformations. Two types of features that are extracted from image content are classified as global and local features. Global features treat an image as a whole and features like color and texture are used for interpreting a particular property (of interest) of all image pixels. On the other hand, local features tend to describe keypoints (points of interest) in an image. Module2: Feature Descriptor, in context of local extracted features, feature descriptor selects image characteristics around the extracted keypoint, describing shape, color, orientation, texture and more, to determine its appearance distinctively and are expected to be invariant to varying imaging conditions. However, the selection of an image registration method for a particular AR application also depends upon a number of other factors like desired accuracy, computational complexity, efficiency to perform adequately in unfavorable situations etc.

2.2 Literature Survey

Summary of Image Registration methods proposed till date: Detection of similar objects in different images is obstructed due to varying imaging conditions. In last decade, many approaches have been considered for excerpting invariant regions from an image such as Edge Based Region detector (EBR) and Intensity Based Region detector (IBR) proposed by Tuytelaars and Gool [Tuytelaars and Gool

2004], where EBR is a structure based detector which describes affine covariant regions in an image by identifying the curved and straight edges in the image as edges are considered to be one of the most stable features in an image that can be detected over a varied range of viewpoint, scale and changes in illumination conditions. Also, edge geometry diminishes the dimensionality of the problem [Tuytelaars and Gool 2004]. IBR, on the other hand identifies affine covariant regions in an image by detecting intensity extrema at multiple scales. The method explores image regions around these intensity extrema in a radial way, describing regions of arbitrary shape which are then replaced by ellipses [Tuytelaars and Gool 2004]. Another widely used feature detector is Scale Invariant Feature Transform (SIFT) proposed by Lowe [Lowe 2004], a fully scale invariant feature detector with few limitations with respect to computational complexity and robustness in case of extreme changing imaging conditions. To overcome some limitations of SIFT detector, few enhancements of the method has been proposed such as: Speeded Up Robust Features (SURF) [Bay et al. 2008], PCA-SIFT [Ke and Sukthankar 2004] and Affine-SIFT (ASIFT) [Yu and Morel 2011] feature detector.

Maximally Stable Extremal Regions (MSER) proposed by Matas et al. [Matas et al. 2004] is an efficient feature detector for extracting invariant regions of interest in an image and aims in overcoming the wide baseline stereo problem, i.e. the problem of identifying correspondences between an image pair describing a scene from different viewpoints. MSER is also termed as one of the most remarkable interest region detector because of its powerful properties like high repeatability, almost linear complexity, $O(nlog(logn))$, low computational demands and near frame rate applicability. The detector works efficiently for tasks like 3D pose estimation of weakly textured planar objects by constructing a perspective invariant frame on the closed contours, a method proposed by Donoser et al. [Donoser et al. 2011] where MSER detects closed contours in an image and then apply Ferns classifier for robust tracking of the detected MSERs in every frame. Similarly, Martedi et al. [Martedi et al. 2013] use MSER for region detection and tracking of arbitrary shapes where the image pixels outlining the MSERs are treated as sampling points to formulate the feature descriptor in order to estimation the orientation of the detected regions. Among many enhancements proposed for the MSER detector, Donoser and Bischof [Donoser and Bischof 2006] proposed a way for extracting MSERs in third dimension (Maximally Stable Volumes (MSV). Forssen [Forssen 2007] introduced MSER as a color based feature detector (Maximally Stable Color Regions (MSCR)) and Nistér and Stewénius [Nistér and Stewénius 2008] proposed a more modest version of MSER in terms of true-worst case linear time in order to make MSER detector more computationally inexpensive.

In last two decades many feature detectors [Harris and Stephens 1988, Mikolajczyk and Schmid 2002, Kadir et al. 2004, Matas et al. 2004] and feature descriptors [Freeman and Adelson 1991, Belongie et al. 2002, Lowe 2004, Ke and Sukthankar 2004,] have been proposed in various detector-descriptor combinations. In a comparative study performed by Moreels and Perona [Moreels and Perona 2007], the results showed that, Harris-affine detector with SIFT descriptor is more suited for changing illumination conditions of an image, whereas, Hessian-affine detector when combined with SIFT descriptor performs much well in different viewpoint changing imaging condition.

However, with the increasing amount of data in applications like object recognition, information retrieval, panorama stitching etc. the complexity of extracting invariant features becomes an obstruction. For example, storing high dimensional descriptors for extracted features requires substantial amount of memory and additionally computing correspondences between image pairs using such descriptor vectors within large datasets adds to the overall time consumption of the method. Therefore, binary features fulfills the necessity of extracting features that are quick in computation and are compact in their representation. In last one decade, several works propose binary feature detectors and descriptors, promising both increased performance as well as compact representation (Features from Accelerated Segment Test (FAST) [Rosten et al. 2010], Binary Robust Independent Elementary Features (BRIEF) [Calonder et al. 2010], Binary Robust Invariant Scalable Keypoints (BRISK) [Leutenegger et al. 2011], Oriented FAST and rotated BRIEF (ORB) [Rublee et al. 2011]). However, in a comparative study of binary detectors and descriptors conducted by Heinly et al. [Heinly et al. 2012], the authors show the performance comparison of binary features against conventional feature detection methods against non-geometric, affine and perspective transformations. Non-geometric transformations corresponds to image-capture dependent imaging conditions like illumination change, JPEG compression, blur change etc. On the other hand, affine and perspective transformations corresponds to image plane rotation/scaling and viewpoint change respectively. The result of their study shows that except for non-geometric transformations, SIFT outperforms all the binary detectors and descriptors.

Augmented Reality history and its evolution: AR has evolved as a potential technology for performing various tasks in practical fields of education, machine repair and maintenance, visualization, gaming etc. It upholds better understanding of knowledge by integrating the real physical world with visual and informative data in form of videos, graphics, sound etc. Several proposed works in literature have described the benefits of AR based systems. Sequentially moving through the period of time for noticing the evolution of AR in different fields, Schwald et al. [Schwald et al. 2001] designed a computer guided maintenance system for complex machinery using AR technology. They aimed in augmenting visual and audio elements to the working environment in order to demonstrate the right procedure to follow while handling a particular machine. Similarly, Wolfgang [Wolfgang 2002] proposed an AR based technique for designing applications in the field of development, production and services in the automotive and aerospace industries. The work of Zhong et al. [Zhong et al. 2003] described a prototype for distributed AR based tele-training system. They introduced a binary square marker for identifying real world objects and distributed AR enabled remote site collaboration between a local viewer and an expert. Lee and Rhee [Lee and Rhee 2008] presented an approach of ubiquitous computing using AR explaining how AR complements incisive and collaborative visualization in a three dimensional informative space embedded within the real world. Henderson and Feiner [Henderson and Feiner 2009] designed a prototype for supporting military mechanics tasks using AR techniques. By tracking a head-mounted display, the authors aimed in augmenting the users view with desired information. They tested their prototype against two baseline conditions, ensuring the applicability of AR in such a scenario. A maintenance collaborative system entitled CAMEKA is described in [Bottecchia 2010], which enabled real time collaboration between an expert and an operator having an access to AR display device fitted with a camera. This system allowed the expert to capture an image from the video flow, add instruction in form of graphics and text to the image and then send back the image to the operator's display device enabling him to access the needed information. Azpiazu et al. [Azpiazu et al. 2011] integrated AR techniques to design an application useful in railway sector. The application focused on minimizing the cost needed for railway management in form of ease integration of data useful for on-site workers to perform the operations. Kleiber and Alexander [Kleiber and Alexander 2011] provided a way for overcoming the drawback of asynchronous video exchange in tele-cooperation by providing synchronous shared visual context for the collaborators without a direct video link using AR techniques. They studied the effectiveness of their approach using engine camshaft maintenance. Benbelkacem et al. [Benbelkacem et al. 2011] and Fukayama et al. [Fukayama et al. 2012] focused on remote collaboration between technicians and experts to complete maintenance and repair tasks by giving augmented information on the user's field of view. Benbelkacem et al. [Benbelkacem et al. 2011] application design focused on service oriented architecture providing information transfer, exchange solutions and remote transfer of virtual objects in real time. Kleiber et al. [Kleiber et al. 2012] proposed an interactive AR enabling real time communication between a hand-held tablet (device displaying the augmented display) and the Virtual Reality (VR) system. Porcelli et al. [Porcelli et al. 2013] proposed an AR solution for providing better and satisfactory customer service. Ferrise et al. [Ferrise et al. 2013] on the other hand described an application that uses AR-VR technologies for supporting tasks related to the maintenance of industrial products. In a recent study, Oskiper et al. [Oskiper et al. 2015] proposed an approach to provide live augmentation of telescopic images using AR binoculars. Their approach aimed at achieving jitter free, robust and precise real-time augmented display using a wide and narrow field of view lenses for estimating the viewer's orientation and increasing the tracking efficiency respectively. Bai et al. [Bai et al. 2015] developed an interactive open-ended pretend play environment using AR ability to visually gestate the surrounding for children suffering from autism.

AR is also incorporated in varied fields of Education System and is widely used in understanding Medical phenomena [Chang et al. 2010], Environmental Science [Tsai et al. 2012] and for performing different Engineering tasks [Azuma et al. 1999, Behzadan 2015, Henderson and Feiner 2011]. Till date, an enormous amount of research has been done to analyze the usability and practical aspects of using AR in an educational system. There are different AR technologies proposed by researchers to demonstrate the limitations and advantages AR brings with it in different scenarios [Sergey et al. 2015, Ibáñez et al. 2014, Fleck and Simon 2013, Tsai et al. 2012, Cerqueira and Kirner 2012, Yeom 2011]. As an example, Sergey et al. [Sergey et al. 2015] proposed an approach to enable a student to access educational content in an interactive platform using AR and 3D visualization technologies. Ibáñez et al. [Ibáñez et al. 2014] designed an AR application for teaching concepts of electromagnetism by exploring the effects of magnetic field by experimentation. The application uses camera of a user's mobile device to recognize objects like cable, magnets, battery, etc. in the real environment and then superimposes information such as the electromagnetic forces and circuit behavior on these objects and the augmented scene is made visible on the user's device. Different aspects of AR in Educational settings are also reported by different studies [Chen et al. 2017, Santos et al. 2014, Radu 2014, Mathison and Gabriel 2012, Martin et al. 2011]. As an example, study conducted by Chen et al. [Chen et al. 2017] gives an overview of AR uses, features and practicality in educational environment. Mathison and Gabriel [Mathison and Gabriel 2012] provides various demonstrations to acknowledge the meaning of AR technology and how it can be used to make authentic learning environments even more engaging and meaningful. Martin et al. [Martin et al. 2011] discusses the evolution of AR trends in education. In addition, few studies also provide future suggestions for AR evolution by examining the current status, opportunities and challenges for AR in the field of education [Saidin et al. 2015, Wu et al. 2013, Cheng and Tsai 2013]. However, research in AR for education is still in its growing stage and is needed in order to address and discover the relevance and characteristics of AR in education by differentiating the AR technology from others. Therefore, extensive analysis of AR in different scenarios will allow designing unique and profitable learning environments based on AR [Wu et al. 2013, Cheng and Tsai 2013].

Summary of Image Registration methods appropriate for building an Augmented Reality System: Among many natural features that could be detected in an image scene, image corners were used by Yuan et al. [Yuan et al. 2006] who adopted Harris Corner detection algorithm and correlation matching method to match the Gray correlation value of corners for obtaining the initial set of matching corners. Their work had a low positioning accuracy as only one pixel level corners are detected. Harris algorithm was outperformed by SIFT [Chen et al. 2007, Li and Chen 2010] because SIFT descriptor incorporates properties like scale invariance and is robust to affine distortions. SIFT holds a number of promising properties for feature detection. However, computational inefficiency of SIFT lowers its applicability for image registration in AR. Li et al. [Li et al. 2012] proposed a real time registration method for AR where SURF is used as a feature detector as it inherits the advantages of SIFT and is computationally more efficient. Also, instead of detecting features for each frame, they made use of Lucas-Kanade optical flow algorithm (based on image pyramid) to trace the detected features down the frame lane which increased the image registration efficiency.

List of practical Augmented Reality works done using each method: Table 2.1. lists some of the methods for image registration process in an AR system. The table briefly explains the AR related work corresponding to each feature detector.

Table 2.1. AR Related Work

Feature Detector	Reference	AR Related Work
Harris Affine Hessain Affine	[Yuan et al. 2006]	A simple registration method based on projective reconstruction technique is proposed using natural features. Their approach processed in two modules: embedding and rendering. Embedding module involved harris affine and hessain affine features for specifying four interest points to compute the world coordinate system for specifying the position of virtual objects in the real scene. Rendering module involved the Kanade-Lucas-Tomasi (KLT) feature tracker to track the natural feature correspondences in a live video.
MSER	[Gomez and Karatas 2014]	Presented a hybrid algorithm for detection and tracking of text in natural scenes. MSER is used for asynchronous text detection and for tracking the detected text. The authors claimed that their proposed approach yeil real time video processing at high frame rates even on low-resource devices.
SIFT	[Chen et al. 2007	Proposed a system initialization algorithm for markerless AR using SIFT key points. Offline calibration of a small number of interest points from the real scene is done for identifying the target object in the video frame. These detected offline features are then matched to SIFT features extracted from the input image during online initialization. Their results demonstrated accurate and robust camera pose estimation.
	Li and Chen 2010]	Presented an approach to improvise E-Commerce technologies using AR techniques. The authors developed a markerless AR system based on SIFT features and demonstrated the efficiency and effectiveness of the method in E-Commerec applications.

Feature Detector	Reference	AR Related Work
ASIFT	[Ham and Golparvar-Fard 2013]	Based their proposed work on the fact that the energy performance simulation tools generally deviate from actual measurements. Therefore, analysing actual performance and computing deviations from simulated data in 3D would help to improvise simulation accuracy. Their approach made use of SIFT, ASIFT and SURF features for reconstructing 3D building thermal performance models from collections of unordered thermal images.
SURF	[Paz et al. 2012]	System named 'PortableAR' is proposed for diverse environments, like guided tourism visits or industrial maintenance tasks. The authors launched two versions of the system: First is based on AR tags. The overall performance of this version depicted low accuracy. The second version is based on SURF features and homography, overcoming the limaitation of first version and providing a successful AR experience.

Survey of various Image Registration methods which handles (to some extent or in some way) the varying imaging conditions: Starting in early 2000, researchers have taken into consideration different environments captured under varied imaging conditions for designing a markerless AR system. Some of the proposed approaches includes the work of Simon et al. [Simon et al. 2000] describing a markerless camera tracking system for AR in a multi-plane environment. Their approach presented an optical tracker for providing reliable results for uncalibrated plane tracking and camera recovery. Ferrari et al. [Ferrari et al. 2001] presented a system for planar AR by tracking local image patches in a completely unknown environment. Genc et al. [Genc et al. 2002] estimated the pose of a camera while observing the real scene using a two-stage process where features extracted from an unknown environment using an external tracker in the first stage are used for computing the pose of camera in real time. Simon and Berger [Simon and Berger 2002] determined the camera pose using piecewise planar structure in the scene, eliminating the need of markers or sensors to recover camera viewpoint. The results depicted comparable output in accuracy with full structure-and-motion methods but with better reliability for indoors and outdoors urban scenes. Lin et al. [Lin et al. 2006] introduced an image registration method based on the Fourier–Mellin transform for an outdoor AR system. Their approach tend to reduce the complex 3D registration problem to a 2D image registration by keeping the observer position fixed. The experimental results showed the applicability of Fourier–Mellin transform image registration algorithm [Ruanaidh and Pun 1998] for an outdoor AR system but the proposed system hold the limitation of fixed viewing position.

Survey of invariant Image Registration methods: Affine Invariance property of a detector is often correlated with its scale invariance characteristic. As in a generalized form, affine transformation could be defined as scaling in each direction of the scene captured in an image [Mikolajczyk and Schmid 2004]. However, scale invariant detectors tend to fail in case of significant affine transformation due to non-uniform scaling of image content. Research literature provides a number of feature detectors for extracting scale and affine invariant features from an image.

Scale Invariant Detectors: Approaches which are often invariant to significant scale changes in an image scene assume that the degree of scale change in image scene is similar in every direction. Existing methods search for local scale space extrema in the 3D representation of an image, extracting the features in vector form corresponding to (*x*-coordinate, *y*-coordinate and scale). This vector representation of extracted features is introduced way earlier by Crowley and Parker [Crowley and Parker 1984]. Their approach represented the scale pyramid formation which is computed using Difference-of-Gaussian (DoG) filters. Lindeberg [Lindeberg 1988] proposed the use Laplacian-of-Gaussian (LoG) and many other derivative based operators for finding 3D maxima of scale normalized differential operators. His findings proved the efficiency of LoG operator for detecting blobs in an image. Bretzner and Lindeberg [Bretzner and Lindeberg 1998] explored the scale invariance property of interest point detectors with automatic scale selection.

Lowe [Lowe 2004] proposed an efficient algorithm for extracting scale invariant features (SIFT) based on the evaluation of local 3D extrema in the scale-space pyramid designed using DoG filters. SIFT features are extracted from an image scene by following a four step process: 1) Scale space peak selection: to locate potential points of interest in an image. 2) Outlier rejection: to identify the keypoints. 3) Orientation assignment: rotation invariance is achieved for every extracted keypoint by determining its orientation based on local image properties.4) Image gradient key-point descriptor evaluation: to achieve invariance to scale and illumination changes, a descriptor vector is evaluated for each extracted keypoint. SIFT features are considered to act stable under different imaging conditions. However, the detector still holds certain limitations, for example, it tends to perform poorly under utmost viewpoint changes. Ke and Sukthankar [Ke and Sukthankar 2004] proposed an alternative to SIFT descriptor, claiming its improved image matching accuracy and lower computational complexity. Bay et al. [Bay et al. 2008] used integral images for attaining image convolutions and box filters for achieving scale space approximations (SURF features). Each detected feature is then associated with a 64 dimensional vector descriptor making it much faster when compared with a 128 dimensional SIFT descriptor. Rosten et al. [Rosten et al. 2010] implied the efficiency of corner detection to design a feature detection procedure that could be efficiently adopted for real time applications (FAST). Calonder et al. [Calonder et al. 2010] proposed an efficient keypoint descriptor using binary strings (BRIEF). Rublee et al. [Rublee et al. 2011] introduced an efficient alternative to SIFT and SURF detectors by combining the efficiencies of FAST keypoint detector and BRIEF descriptor and by overcoming some of their limitations.

Affine Invariant Detectors: An affine invariant algorithm for corner detection, proposed by Alvarez and Morales [Alvarez and Morales 1997] applied affine morphological multi-scale analysis to extract corners. Each extracted point is then associated with a chain of points detected at different scales having same local image structure. Finally, location and orientation of the detected corner is computed using the bisector line given by the chain of points. Lindeberg and Garding [Lindeberg and Garding 1997] developed a method for finding affine features representing a blob like structure using an iterative procedure in the context of shape from texture. Their approach extracted the maxima of a uniform scale-space representation and iteratively modified the scale and shape of points. Laptev and Lindeberg [Laptev and Lindeberg 2003] adopted the similar method to detect elliptical blobs in the context of hand tracking.

Tuytelaars and Gool [Tuytelaars and Gool 1999, 2000] proposed two approaches for detecting image features in an affine invariant way. The former approach extracted Harris points and used the nearby edge for defining a parallelogram region such that it systematically cover the same part of a surface in an image by automatically adopting different shapes for different viewpoints in different images. Several intensity based functions are used for determining such parallelogram regions in an image. However, for each extracted Harris point, two nearby edges are used which limited the number of potential features in an image. The latter approach initiated by extracting local intensity extrema and an ellipse is defined for the region determined by significant changes in the pixel intensity profiles. A similar approach based on local intensity extrema is introduced by Matas et al. [Matas et al. 2004]. The extracted regions of interest (MSER), defined as the connected components in an image having extremal regions either with higher (bright) or lower (dark) intensity properties than all the pixels on its outer boundary, are considered to be stable features because the ordering of pixels intensities is done under monotonic transformations.

Baumberg [Baumberg 2000] used affine shape estimation for matching and recognition. The proposed approach extracted interest points at several scales using the Harris detector and then adapted the shape of the point's neighborhood to the local image structure using the procedure proposed by Lindeberg and Garding [Lindeberg and Garding 1997]. Schaffalitzky and Zisserman [Schaffalitzky and Zisserman 2003] extended the Harris-Laplace detector [Mikolajczyk and Schmid, 2001] by affine normalization proposed by Baumberg [Baumberg 2000]. However, the location and scale of points are provided by the scale invariant Harris-Laplace detector [Mikolajczyk and Schmid, 2001], which is not invariant to significant affine transformations.

Yu and Morel [Yu and Morel 2011] proposed a fully affine invariant feature detector, ASIFT. The detector simulated the two camera axis orientation parameters: latitude and longitude angles and used SIFT to simulate scale parameter and to normalize the rotation and translation parameters. A comparative study presented by the authors stated that ASIFT is able to outperform other feature detectors under extreme viewpoint changing conditions. A study done by Mikolajczyk and Schmid [Mikolajczyk and Schmid 2004] took into consideration many of the above discussed image registration methods and advocated that no detector-descriptor combination performs well with viewpoint changes of more than 25−30°. This problem was later resolved by ASIFT detector which gave satisfactory results even at a viewpoint change of 60° between the reference and the matched image [Yu and Morel 2011].

Binary feature detectors such as FAST, BRIEF, BRISK and ORB are proven to execute faster than SIFT and SURF detectors, but in many cases this speed is achieved by reducing the number of sampling points around an extracted feature while designing the feature descriptor and thus compromising the matching quality [Heinly et al. 2012]. Also, these binary feature detectors are not rotation and scale invariant for larger parameters and hence achieving affine invariance properties using these detectors becomes questionable.

Survey of various Image Registration methods which handles Image Quality and noise in images: Another factor that effects the efficiency of image registration methods is the image quality. Although, there are not much reference articles present in the literature that directly discusses about image quality as a factor while designing an image registration method, however, the conditions taken into consideration often correlates with a few imaging conditions like illumination change [Simon and Berger 2002, Lin et al. 2006] when processing in an outdoor environment.

Summary of which Image Registration method is useful under what Imaging Condition: Table 2.2 summarizes the properties of some of the above listed feature detectors that are considered promising for identifying stable interest points in an image scene.

Survey of various metrics used for the comparative study of performance of Image Registration methods: There have been a number of comparative studies done in literature [Mikolajczyk and Schmid 2004, Mikolajczyk et al. 2005, Moreels and Perona 2007] to signify the performance of various feature detectors. The metrics generally used by these comparative studies for analyzing the performance of different methods involved: 1) repeatability score: defined as the ability of the detector to determine corresponding scene regions. 2) Computational complexity: serves as an important factor in cases when detectors are applied to image sequences or large image databases. This factor also plays a fundamental

Table 2.2. Properties of Feature Detectors

Technique	**Affine Invariance**	**Illumination Invariance**	**Robustness to viewpoint change**	**Step-wise Computational Complexity**
EBR	Partial	Partial	Limited to 25°-30°	Finding initial corner points and edges in the image = $O(n)$ Constructing the actual region =$O(pd)$ where p= number of corners and d= average number of edges nearby a corner
Harris-Affine	Partial	Partial	Limited to 25°-30°	Initial Point Detection (Harris Laplace)= $O(n)$ Automatic scale selection and shape adaption algorithm = $O((m+k)p)$ where m = number of inspected scales in the automatic scale space selection, k= number of iterations in the shape adaption algorithm and p= number of initial points.
Hessian-Affine	Partial	Partial	Limited to 25°-30°	Initial Point Detection (Harris Laplace)= $O(n)$ Automatic scale selection and shape adaption algorithm = $O((m+k)p)$ where m = number of inspected scales in the automatic scale space selection, k= number of iterations in the shape adaption algorithm and p= number of initial points
IBR	Partial	Partial	Limited to 25°-30°	Finding Intensity Extrema = $O(n)$ Constructing the actual region around the intensity extrema = $O(p)$ where p = number of intensity extrema
MSER	Partial	Partial	Limited to 25°-30°	Sorting Step = $O(n)$ (sort is implemented using binary sort) Union find algorithm = $O(n\ log \log n)$
SIFT	Partial	Partial	Limited to 25°-30°	Proportional to the number of keypoints detected. In matching step, the complexity increase because of high dimensionality of the descriptor
ASIFT	Fully Affine Invariant	Partial	Robust to a viewpoint change up to 60◦.	Proportional to input image area and the number of simulated tilts. Feature computation complexity is 1.5 times greater than single SIFT routine.
SURF	Partial	Partial	Limited to 25°-30°	Proportional to the number of keypoints detected. Matching complexity is much lower than SIFT because of the use of 64 dimensional vector descriptor.

role in implying the applicability of an image registration method in real time applications. 3) Efficiency in terms of correct correspondences between image pairs. 4) Ability of a detector to determine stable regions of interest under extreme changing imaging conditions [Yu and Morel 2011].

2.3 Problem Definition

Research gaps as emerged from the literature review: Literature survey presented in this chapter highlights a number of limitations of various image registration methods proposed till date, for example, SIFT and SURF feature detector's efficiency is limited for a viewpoint change of upto 30° and both the detectors are not fully affine invariant [Yu and Morel 2011]. ASIFT feature detector computational complexity is not appreciable for real time applications and binary feature detectors such as BRIEF, FAST, BRISK and ORB tend to compromise the feature matching efficiency in various scenarios for attaining the reduced time of processing [Heinly et al. 2012]. Therefore, to generalize the limitations of these image registration methods, following research gaps are revealed from the above literature survey keeping in mind their applicability in an AR system:

- Image registration results achieved in an environment where the scene is stationary, are quite accurate. The problem arises in case of a changing environment.
- Camera pose estimation becomes difficult and the system may breakout completely when majority of the surface texture has changed.
- Image registration methods for outdoor environment are not much accurate because these systems are sensitive to changes in illumination.
- Conventional image registration methods do not take into consideration images with varied image quality and imaging factors like blur and JPEG compression.
- Computational complexity of the methods make them unsuitable for real time AR applications.

Research gaps this work tries to reduce: To reduce some of the above listed research gaps, present research aimed at:

- Investigating the behavior of conventional image registration methods when dealing with varied image quality, changing imaging conditions and affine transformations.
- Designing and/or improving existing algorithms to obtain significant results in outdoor environments.
- Designing and/or improving existing algorithms to obtain significant results when dealing with varied image quality, changing imaging conditions and affine transformations.
- Achieving low computational complexity of designed and/or improved image registration method, increasing its applicability for AR applications.

Problem definition: To target above-mentioned aims, an intensive literature survey has been done on image registration methods from various journals, web sites and other sources. Image registration methods present till date are studied in a fashion that categorizes them according to their applicability, computation requirements and efficiency. A greater emphasis is given on the algorithms which can be directly used for markerless AR system buildup. However, as noticed in literature survey, there are only a few algorithms which can be used for designing an AR system in their present form. This is due to their limitations in terms of speed, affine invariance, robustness under varying imaging conditions and image quality. Solution to these limitations can be achieved either by developing new image registration methods or improving upon the existing methods so that their implication in an AR system achieves desired efficiency.

2.4 Objectives of the Research

The objective of the proposed research work is to build a robust image registration method for markerless AR such that it holds properties like: affine invariance, high accuracy, low computational complexity and works well with varying quality of images describing image scenes under changing imaging conditions. The objectives of the research here:

- Proposing and/or improving image registration techniques such that they are adaptive to scene illumination changes and other image changing conditions.
- Proposing novel image registration algorithm and/or improving existing algorithms in terms of computational complexity and robustness.
- Developing an image registration method appropriate and useful for AR applications.

2.5 Summary

This chapter presents the literature survey of different image registration methods proposed till date. The chapter briefly discusses the efficiency and limitations of these methods in different scenarios. The chapter also gives a brief history and evolution of AR in different fields of applications. Next chapter presents a brief overview of all the theoretical concepts directly related to the research work.