# NOVEL POSE ESTIMATION ALGORITHM FOR MOBILE AUGMENTED REALITY BASED ON INERTIAL SENSOR FUSION

BY

MIR SUHAIL ALAM

A dissertation submitted in fulfilment of the requirement for the degree of Master of Science (Computer and Information Engineering)

Kulliyyah of Engineering
International Islamic University Malaysia

JANUARY 2022

# ABSTRACT

Augmented Reality (AR) applications have become increasingly ubiquitous as it integrates virtual information such as images, 3D objects, video and more to the real world, which further enhances the real environment. AR functions as an interface by superimposing virtual world information on top of the real environment. Moreover, image registration is critical in computer vision. Image registration is widely employed in a variety of applications, including *image matching, change detection, 3D reconstruction, mobile robots* and more. However, the main concern in the augmented reality is the registration of virtual information, or how to stack the virtual information generated by a computer into a real environment based on its surroundings. Vision-based pose estimation in augmented reality application has been widely investigated. Though, many of the earlier techniques relied on markers, and vision-based registration on mobile devices is currently too computationally expensive and imprecise. The accuracy of vision-based pose estimation approach is degraded due to illumination changes from frame-to-frame image sequence. This may cause jitter in the estimated pose. Many researchers have investigated the pose estimation and augmentation of the 3D virtual object in the physical environment. However, certain loopholes exist in the existing system while estimating the object's pose, making it inaccurate for Mobile Augmented Reality (MAR) applications. This study proposes to estimate the pose of an object by blending both vision-based technique and MEMS sensor (gyroscope) to minimize the jitter problem in MAR. The algorithm used for feature detection and description is Oriented-FAST Rotated-BRIEF (ORB), whereas to evaluate the homography for pose estimation, Random Sample Consensus (RANSAC) is used. We evaluated the performance of augmenting the 3D object using the both the techniques which includes vision data only and incorporating the sensor data with the vision data. After extensive experiments, the validity of the proposed method was superior to the existing vision-based pose estimation algorithms. The proposed method has proven to be successful in overcoming the problem of jitter in the existing system. The proposed algorithm was benchmarked with two other algorithms, where some shortcoming in the latter were addressed such as computational cost, however, the issues of gyroscope like drift remains limitation.

# خلاصة البحث

أصبحت تطبيقات الواقع المعزز (AR) منتشرة على نطاق واسع لأنها تدمج العناصر الافتراضية مثل الصور والأشكال ثلاثية الأبعاد والفيديو إلى العالم الحقيقي ، وبالتالي فإن الواقع المعزز يثري التفاعل مع البيئة. يعمل الواقع المعزز كواجهة ، حيث يقوم بتركيب معلومات العالم الافتراضي فوق البيئة الحقيقية. علاوة على ذلك ، يعد تسجيل الصور أمرًا بالغ الأهمية في رؤية الكمبيوتر. يتم استخدام تسجيل الصور على نطاق واسع في مجموعة متنوعة من التطبيقات ، بما في ذلك مطابقة الصور ، واكتشاف التغيير ، وإعادة البناء ثلاثي الأبعاد ، و الروبوتات المتنقلة. ومع ذلك ، فإن العقبة الرئيسية في الواقع المعزز هي تسجيل المعلومات الافتراضية ، بمعنى آخر كيفية تكديس المعلومات الافتراضية التي تم إنشاؤها بواسطة الكمبيوتر في محيط البيئة الحقيقية. تمت دراسة تقدير الوضع القائم على الرؤية في تطبيقات الواقع المعزز على نطاق واسع. على الرغم من أن العديد من التقنيات السابقة كانت تعتمد على العلامات ، إلا أن التسجيل المستند إلى الرؤية على الأجهزة المحمولة يعد حاليًا مكلفًا للغاية وغير دقيق من الناحية الحسابية. تتدهور دقة نهج تقدير الوضع القائم على الرؤية نتيجة لتغيرات الإضاءة من تسلسل صورة من إطار إلى إطار. وهذا بدوره قد يسبب توترًا في الوضع المقدر. درس العديد من الباحثين تقدير الوضع والتكبير البصري للعناصر الافتراضية ثلاثية الأبعاد في البيئة المادية. ومع ذلك ، توجد بعض الثغرات في النظام الحالي أثناء تقدير وضع الكائن ، مما يجعلها غير دقيقة في تطبيقات الواقع المعزز للجوال (MAR). تقترح هذه الدراسة تقدير وضعية المجسم عن طريق دمج مستشعر MEMS (جيروسكوب) بتقنية قائمة على الرؤية لتقليل مشكلة الاهتزاز في MAR. الخوارزمية المستخدمة لاكتشاف الميزات ووصفها هي موجَّهة - سريعة الدوران - موجز (ORB). بينما لتقييم التجانس لتقدير الوضع ، تم استخدام توافق العينة العشوائية (RANSAC). قمنا بتقييم أداء الزيادة المرئية للعنصر ثلاثي الأبعاد باستخدام كل من التقنيات القائمة على الرؤية وبعد دمج بيانات المستشعر مع بيانات الرؤية. بعد تجارب مكثفة ، أصبحت الطريقة المقترحة متفوقة على خوارزميات تقدير الوضع القائمة على الرؤية الحالية. أثبتت الطريقة المقترحة نجاحها في التغلب على مشكلة الاهتزاز في النظام الحالي. تم قياس الخوارزمية المقترحة مقابل خوارزميتين أخريين ، حيث تمت معالجة بعض أوجه .القصور في الخوارزميتين مثل التكلفة الحسابية ، ومع ذلك ، لا تزال مشاكل الجيروسكوب مثل الانجراف دون حل.

# APPROVAL PAGE

I certify that I have supervised and read this study and that in my opinion, it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Master of Science (Computer and Information Engineering)

…………………………………..
Malik Arman Morshidi
Supervisor

…………………………………..
Teddy Surya Gunawan
Co-Supervisor

…………………………………..
Rashidah Funke Olanrewaju
Co-Supervisor

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Master of Science (Computer and Information Engineering)

…………………………………..
Suriza Ahmad Zabidi
Internal Examiner

…………………………………..
Farah Diyana Abdul Rahman
Internal Examiner

This dissertation was submitted to the Department of Electrical and Computer Engineering and is accepted as a fulfilment of the requirement for the degree of Master of Science (Computer and Information Engineering)

…………………………………..
Md. Rafiqul Islam
Head, Department of Electrical and Computer Engineering

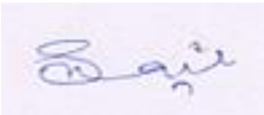
This dissertation was submitted to the Kulliyyah of Engineering and is accepted as a fulfilment of the requirement for the degree of Master of Science (Computer and Information Engineering)

…………………………………..
Sany Izan Ihsan
Dean, Kulliyyah of Engineering

# DECLARATION

I hereby declare that this dissertation is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Mir Suhail Alam

Signature  Date: 11-01-2022

**INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA**

**DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH**

**NOVEL POSE ESTIMATION ALGORITHM FOR MOBILE AUGMENTED REALITY BASED ON INERTIAL SENSOR FUSION**

I declare that the copyright holders of this dissertation are jointly owned by the student and IIUM.



No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

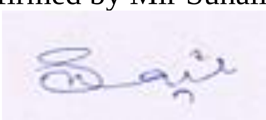
1. Any material contained in or derived from this unpublished research may be used by others in their writing with due acknowledgement.

2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purposes.

3. The IIUM library will have the right to make, store in a retrieved system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Mir Suhail Alam



11-01-2022

……………………………… ………………………..

Signature Date

# ACKNOWLEDGEMENTS

All glory is due to Allah, the Almighty, whose Grace and Mercy have been with me during my programme. Although it has been difficult, His Mercies and Blessings on me have made the enormous effort of finishing my thesis easier.

I am very grateful to my supervisor, Asst. Prof. Dr Malik Arman Morshidi, whose patient demeanour, politeness, promptitude, thoroughness, and friendship enabled me to complete my task successfully. I would want to take this opportunity to thank him for his detailed remarks, helpful recommendations, and thought-provoking questions, which have significantly improved this thesis. His exceptional understanding of the purpose and content of this work resulted in intelligent remarks, recommendations, and questions that aided me. Despite his commitments, he made time to listen to and attend to me whenever I desired. The moral support he offered to me was without a doubt a boost that aided in the development and writing of the draught of this research effort. I am also grateful to my co-supervisors, Prof. Dr. Teddy Surya Gunawan and Assoc. Prof. Dr. Rashidah Funke Olanrewaju, whose assistance and cooperation aided in completing this study.

Finally, I want to thank my beloved parents and family for their prayers, understanding, and patience while I was away.

Once again, I thank Allah for His inexhaustible generosity to me, one of which is allowing me to complete writing this thesis.

Alhamdulillah

# TABLE OF CONTENTS

Abstract .......... ii
Abstrsact in Arabic .......... iii
Approval Page .......... iv
Declaration .......... v
Copyright Page .......... vi
Acknowledgements .......... vii
Table of Contents .......... viii
List of Tables .......... x
List of Figures .......... xi
List of Symbols .......... xiii
List of Abbreviation .......... xiv

**CHAPTER ONE: INTRODUCTION .......... 1**
1.1 Background of the study .......... 1
1.2 Problem Statement .......... 4
1.3 Research Objectives .......... 4
1.4 Research Methodology .......... 4
1.5 Research Scope .......... 6
1.6 Thesis Organization .......... 6

**CHAPTER TWO: LITERATURE REVIEW .......... 7**
2.1 Introduction .......... 7
2.2 Mobile Nature of AR Applications .......... 10
2.3 Major Steps involved in the Operation of AR system .......... 12
2.4 Feature Detection and Description .......... 13
2.4.1 Physical Characteristics of Feature Detectors .......... 16
2.5 Algorithm Used for Feature Detection and Description .......... 17
2.5.1 Feature Detector Algorithms .......... 18
2.5.1.1 Harris Detector .......... 20
2.5.1.2 SUSAN Detector .......... 21
2.5.1.3 FAST Detector .......... 22
2.5.1.4 Hessian Detector .......... 23
2.5.2 Feature Descriptor Algorithm .......... 24
2.5.2.1 Scale Invariant Feature Transform (SIFT) .......... 24
2.5.2.2 Speeded-Up Robust Features Descriptor (SURF) .......... 27
2.5.2.3 Graditent Location Orientation Histogram(GLOH) .......... 30
2.5.2.4 Binary Robust Independent Elementary Features (BRIEF) .......... 31
2.5.2.5 Oriented FAST Rotated BRIEF .......... 32
2.5.3 Feature Matching .......... 37
2.6 Homography Estimation .......... 40
2.6.1 RANSAC Algorithm .......... 41
2.7 Inertial Sensors .......... 43
2.7.1 Gyroscopes .......... 44
2.7.2 Accelerometers .......... 46

2.8 Principle of Augmented Reality ........ 48
2.9 Related Work ........ 48
2.9.1 Vision based methods. ........ 49
2.9.2 Inertial sensor-based methods. ........ 50
2.9.3 Fusion of Sensor-Based Inertial-Vision Methods ........ 51
2.10 Application Fields of Augmented Reality ........ 54
2.10.1 Tourism and Navigation ........ 54
2.10.2 Training and Education ........ 55
2.10.3 Entertainment and Advertisement ........ 56
2.11 Summary ........ 60

**CHAPTER THREE: RESEARCH METHODOLOGY ........ 61**
3.1 Introduction ........ 61
3.2 Proposed Method ........ 61
3.2.1 Video Acquisition ........ 62
3.2.2 Target Surface Recognition ........ 63
3.2.2.1 Feature Detection ........ 64
3.2.2.2 Feature Description ........ 65
3.2.2.3 Feature Matching ........ 65
3.2.2.4 Homography Estimation ........ 65
3.3 Pose Estimation in Augmented Reality ........ 66
3.4 Inertial Sensor (gyroscope) ........ 68
3.5 Summary ........ 69

**CHAPTER FOUR: RESULTS AND DISCUSSION ........ 70**
4.1 Introduction ........ 70
4.2 Experimental Setup ........ 70
4.2.1 Hardware Setup ........ 70
4.2.2 Software Setup ........ 71
4.3 Comparative Results of different feature detection algorithms ........ 71
4.4 Quantitative Analysis of Feature detection Algorithms ........ 73
4.5 Feature Matching Time based on different parameters ........ 77
4.6 Experimental Results of Pose Estimation and 3D Augmentation ........ 80
4.7 Quantitative Analysis ........ 81
4.7.1 Ground Truth ........ 81
4.7.2 Vision Based Analysis ........ 82
4.8 Incorporation of sensor data ........ 86
4.9 Qualitative Analysis ........ 89
4.10 Summary ........ 92

**CHAPTER FIVE: CONCLUSION AND FUTURE WORK ........ 93**
5.1 Conclusion ........ 93
5.2 Future Work ........ 94

**BIBLIOGRAPHY ........ 95**

**LIST OF PUBLICATION ........ 108**
**APPENDIX A: MATLAB CODES ........ 109**

# LIST OF TABLES

Table 2.1 Comparison Table of Feature Extraction Algorithms 35

Table 2.2 Some Typical MAR applications 58

Table 4.1 Hardware configuration 71

Table 4.2 Comparison of Different Feature Extraction Algorithms Based on Different Parameters 74

# LIST OF FIGURES

Figure 1.1 Overview of the stages involved in the pose estimation ..........................5

Figure 2.1 From reality (left) to virtuality (right) .....................................................8

Figure 2.2 Abstract representation of the MAR system components ........................12

Figure 2.3 Corner Feature Charateristics .................................................................19

Figure 2.4 Detecting feature in an image patch using FAST detector .......................23

Figure 2.5 SIFT descriptor representation for a 16x16 pixel patch and 4x4 descriptor array.........................................................................................................26

Figure 2.6 Gaussian second order derivatives and their approximations ..................28

Figure 2.7 Dividing the interest region into 4 × 4 sub-regions for computing the SURF descriptor..........................................................................................29

Figure 2.8 GLOH algorithm schematic representation using log-polar bins.............30

Figure 2.9 Feature matching between images...........................................................37

Figure 2.10 Homography evaluation for Pose Estimation........................................40

Figure 2.11 Under various situations, a mass-and-spring accelerometer: (a) stationary or in uniform motion; (b) accelerating; (c) stationary..............................47

Figure 3. 1 Overview of the stages associated with vision and sensor-based pose estimation. ..................................................................................................62

Figure 3.2 Reference coordinate system ...................................................................67

Figure 4.1 Stages of image registration ....................................................................72

Figure 4.2 Different transformations of the image ...................................................73

Figure 4. 3 Comparison of image rotation and illumination matching time..............78

Figure 4.4 Comparison of affine and blur matching time..........................................78

Figure 4.5 Comparison of scale up and scale down matching time...........................79

Figure 4.6 Key points of the target (left), 5 strong key points in each frame (right) .82

Figure 4.7 Ground truth of all video frames .............................................................82

Figure 4.8 Feature Matching ........................................................................ 84

Figure 4.9 Homography using RANSAC ................................................................ 84

Figure 4.10 Matching Key-points versus frame number .......................................... 85

Figure 4.11 Key-point Matching based on set threshold. .......................................... 86

Figure 4.12 Sensor data of x, y, z axis ........................................................................ 87

Figure 4.13 Combined gyroscopic data along x, y, z axis. ......................................... 87

Figure 4.14 Augmentation based on the vision data of different frames. .................. 90

Figure 4.15 Augmentation based on vision data from frame 428-431 ...................... 91

Figure 4.16 Augmentation after incorporating the sensor data with the vision data of the same frames. ........................................................................ 91

# LIST OF SYMBOLS

| | |
|---|---|
| $I_x$ and $I_y$ | Local image derivatives in the *x* and *y* directions respectively. |
| $\lambda_1, \lambda_2$ | Eigenvalues of the auto-correlation matrix. |
| K | Adjustable parameter |
| $\sigma$ | Standard deviation |
| $v$ | Feature vector |
| $C_L^I$ | Rotation matrix |

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AR | Augmented Reality |
| AI | Artificial Intelligence |
| VR | Virtual Reality |
| MAR | Mobile Augmented Reality |
| MEMS | Micro-electromechanical system |
| ORB | Oriented FAST Rotated BRIEF |
| SURF | Speeded Up Robust Feature |
| BRIEF | Binary Robust Independent Elementary Features |
| SIFT | Scale Invariant Feature Transform |
| RANSAC | Random Sample Consensus |

# CHAPTER ONE
# INTRODUCTION

## 1.1 BACKGROUND OF THE STUDY

With the passing of time, technology is evolving rapidly and as we step into the future, intelligent or smart machines in many areas are also increasingly changing and upgrading human capacity. Augmented reality (AR) and Artificial intelligence (AI) are separate technologies; however, they can be combined to produce novel experiences. In order for digital and physical items to coexist in AR, a 3D representation of the world must be created. Visual data, as well as accelerometers and gyroscopes, are utilised to create a world map and track movement on it. Most of these activities are still performed using classic computer vision approaches that do not employ machine learning. However, AI models have proved tremendously successful in creating many of the necessary immersive AR experiences on their own. Deep neural networks can recognise vertical and horizontal planes, estimate depth and segment pictures for realistic occlusion, and infer 3D positions of objects in real time. Because of these capabilities, AI models are replacing some of the more traditional methods to computers that support augmented reality experiences. AI a subfield of computer engineering that refers to the intelligence shown by computers or programming ventures. AI is currently a potential distinct benefit alongside machine learning throughout the whole of the computing supported by solid information investigation. AI research has quickly influenced the rise of keen advancements that have a significant impact on our daily lives. Science, engineering, business, and medicine have become more intelligent, with forecasting abilities to make our lives easier in our daily activities. In the coming years, artificial machines will augment or replace human capabilities in a variety of areas.

AR is an experience that combines the physical and digital environments; it has sparked a lot of interest in the academic community due to its widespread applicability. AR functions as an interface, superimposing virtual world information on top of the real environment. In other words, it alters the real world with computer-generated items. Unlike Virtual Reality (VR), where the entire environment is produced by a computer. AR is a combination of tangible and virtual things that can register real and virtual objects and work together in real time (Caarls, 2009; W. Zhang et al., 2017). The major goal is to blend real-world environments with computer-generated VR elements such as 3D models, music, text, video, and more. (R. Azuma et al., 2001) described AR in a novel way that includes real-time object augmentation. In the actual world, virtual and physical items must be mathematically aligned. AR technology also incorporates Real-Time Tracking and Registration, Multimedia, Sensing, and Intelligent Interaction (Y. Chen et al., 2019). Furthermore, with the assistance of software development kits (SDK) such as Vuforia and ARToolKit (Khan et al., 2018; X. Liu et al., 2018) (ARKit for Apple and ARCore for Android), various AR games have been produced. AR applications in smartphones employ computer vision to find various objects of interest in their surroundings by using the camera as input. There are four forms of AR technology: marker-based AR, markerless AR, projection-based AR, and superimposition-based AR. Although the objective is identical, that is, to improve the surrounding view using computer-based graphics. However, each type of technology has unique characteristics that set it apart from the others.

Image registration is critical in computer vision. Image registration is widely employed in a variety of applications, including *image matching* (Han et al., 2020;

Jiayuan Li et al., 2020; J. Ma et al., 2021; Q. Wu et al., 2020), *change detection* (B. Chen et al., 2020; Jia et al., 2019), *3D reconstruction* (Hu et al., 2020; López-Torres et al., 2020; Lv et al., 2020), *guidance* (Lee et al., 2020; Marquardt et al., 2020), *mapping sciences* (Parmehr et al., 2014), and *mobile robots* (Yuanzhe Wang et al., 2021; Yule Wang & Yang, 2018). Also, the main concern in the augmented reality is the registration of virtual information, or how to stack the virtual information generated by a computer into a real environment based on its surroundings (Blanco-Novoa et al., 2018; Juan Li et al., 2017; Maimone et al., 2017). Vision-based pose estimation in augmented reality application has been widely investigated. However, many of the earlier techniques relied on markers, and vision-based registration on mobile devices is currently too computationally expensive and imprecise. The accuracy of vision-based pose estimation approach is degraded due to illumination changes from frame-to-frame image sequence. This may cause jitter in the estimated pose. So, to improve the accuracy of the pose and reduce this research gap, there is need of algorithm to solve the various issues of mobile augmented reality (MAR) applications. To address this issue, we suggest using inertial sensor fusion to align vision-based pose estimation with sensor-based pose estimation. Even though a number of pose estimation systems have been offered, providing a low-cost, real-time, accurate, and easy-to-deploy solution remains a major challenge. This research will investigate how the built-in inertial sensors in mobile phones such as gyroscope and accelerometers can be incorporated with the vision data to improve the performance accuracy of pose estimation. The outcome of this research would be a robust and smooth pose estimation algorithm that utilise the low-cost inertial sensor (gyroscope) that has been there in almost every mobile devices.

## 1.2 PROBLEM STATEMENT

The objective of the pose estimation is to estimate precisely the relative position of the camera, with the primary aim of placing virtual objects into the real video stream. The virtual object must be kept intact without jitters on the real object. However, many existing techniques based on vision data only suffer from a tracking error, and the inaccuracy is caused mostly due to the illumination changes in the entire video sequence. This project will explore how built-in inertial sensors present in the mobile phones such as a gyroscope can aid the performance accuracy of vision-based pose estimation that can result in better augmentation of the virtual objects in the real environment.

## 1.3 RESEARCH OBJECTIVES

This study will primarily investigate the existing methods of pose estimation in the vision-based pose estimation and will attempt to reduce shortcomings in the previous studies by incorporating the sensor data with vision data to enhance the accuracy of the pose estimation in augmented reality. The objectives of the study are listed below:

1. To investigate and analyze the feature detection and matching techniques used in the MAR applications.

2. To propose an algorithm for enhancement of pose estimation accuracy in MAR applications based on inertial sensors and vision data.

3. To evaluate the performance parameters of the proposed algorithm on pose estimation based on inertial sensor.

## 1.4 RESEARCH METHODOLOGY

This study looks at how integrating the inertial sensor data can improve pose estimate accuracy on mobile devices. Sensor fusion will be used to refine the pose information derived from vision processing. The basic architecture of our study is shown in Figure 1.1. The method begins by recognising the features in the first frame of the video sequence. We will study several keypoint detectors and descriptors, including SIFT, SURF, and ORB and many more, to establish the trade-off between accuracy and speed. Enough descriptors are required to match and determine feature correspondences between frames. Secondly, we will investigate feature tracking between video frames. Finally, we will incorporate inertial data with the vision data to increase the efficiency and trajectory of pose estimation. When there are major changes in the interframe patch and the vision-based approach fails to estimate the pose accurately, the system pauses a few frames before determining that tracking has been lost and proceeding with the re-initialisation by incorporating the gyroscope sensor data with the vision data to enhance the accuracy of the pose estimation.

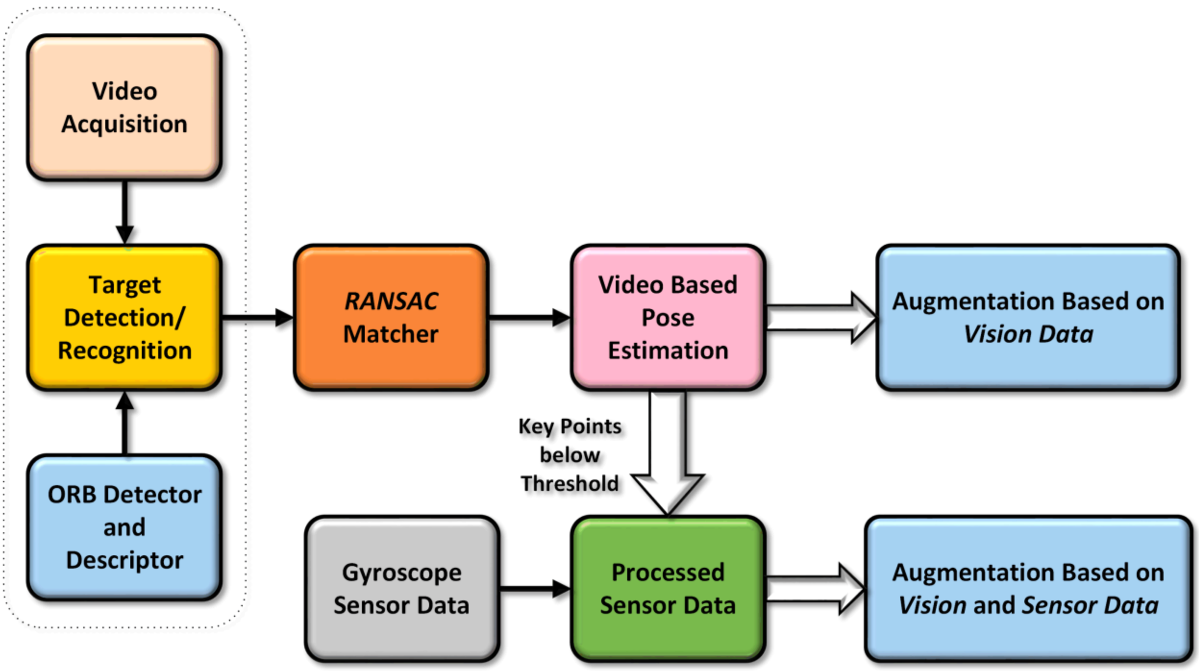



Figure 1.1 Overview of the stages involved in the pose estimation.

## 1.5 RESEARCH SCOPE

This study will investigate how the built-in MEMS sensors in mobile devices help to improve the accuracy of pose estimation for MAR applications. Furthermore, only a limited number of robust feature extraction techniques, such as SIFT, SURF, and ORB, will be investigated. However, difficulties concerning sensors, such as the influence of drift on the gyroscope and accuracy issues in fast motion, will not be addressed in this study.

## 1.6 THESIS ORGANIZATION

This dissertation's flow is classified as follows. The second chapter comprises literature review and addresses research on pose estimation based on vision data and sensors in the domain of augmented reality. The methodology and implementation of the study are covered in Chapter 3. Chapter 4 expands on the methodology and implementation. Finally, in Chapter 5, the conclusion, benchmarking, and future recommendations are presented.

# CHAPTER TWO
# LITERATURE REVIEW

## 2.1 INTRODUCTION

Augmented Reality (AR) is technology that has been evolved from virtual reality. Based on information provided by the computer system, this method largely improves the user's perspective of the real world, and reality can be augmented by superimposing computer-generated virtual things, scenarios, and system prompts over the real surroundings. It mostly assists humans in depicting inaccessible scenes in the real world. AR is a technology that seamlessly combines real-world and virtual data. Using computing technology, physical information (sound, visual information and more) that is difficult to experience in the real world at a certain time and space is replicated and superimposed. Humans interpret virtual information when it is applied to the real environment. As a result, the experience surpasses reality. In real time, virtual and physical objects are superimposed on the same screen or place. With the advancement of science and technology, AR has increasingly entered the people’s field of vision. AR augments virtual information with the real-world environments to provide individuals with more meaningful information (Khairnar et al., 2015; Rehman & Cao, 2017; Santos et al., 2017).

Mobile augmented reality (MAR) has captivated the curiosity of both industry and academia during the last decade. MAR augments a mobile user's real world with computer-generated virtual content. In the case of intense visuals that touch the mobile user's original vision, the intensity of the virtual contents and their influence on the mobile user's vision determine realism or virtuality. Figure 2.1 displays the classification of several types of reality and virtuality. Real Reality refers to the user's environment as it exists without the help of technology, whereas Virtual Reality refers to the reality that people perceive that is unrelated to their surroundings and is entirely manufactured by a computer. AR on mobile devices is now a realistic due to advancements in mobile technology such as built-in cameras, processing capabilities, sensors and mobile cloud computing.

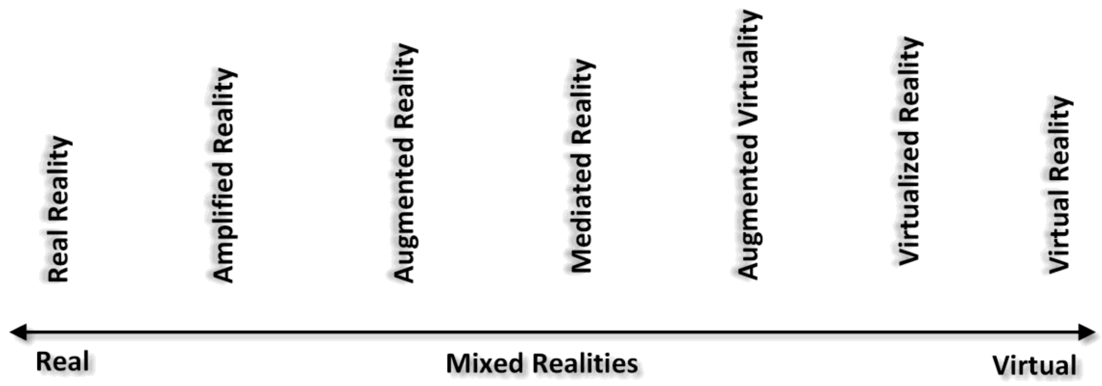



Figure 2.1 From reality (left) to virtuality (right)

Human-computer interaction interfaces, mobile cloud computing, mobile computing, computer vision, device-to-device communications and network caching breakthroughs have enabled new user experiences that improve how we acquire, interact with, and display information in our surroundings. In previously inconceivable ways, we may now merge information from our senses and mobile devices. Cloud infrastructure and service providers continues to provide innovative services to encourage the development of new MAR applications. The MAR market is expected to grow at a 31.1 percent compound annual growth rate (CAGR) from 2020 to 2025, rising from USD 7.6 billion in 2020 to USD 29.5 billion by 2025. (Mobile Augmented Reality Market by Device Type, Offering Application; COVID-19 Impact Analysis ;

MarketsandMarkets™, n.d.). The growing use of MAR in entertainment, e-commerce, and a variety of other industries has fuelled industry growth. One of the main drivers for the growth of the MAR market is the rising interest of large technology businesses in mobile AR apps. Applications that utilise AR have the ability to influence user opinions and purchasing behaviour. It scientifically explores user experience and its effects on customer perceptions of hedonic and utilitarian pleasure, informativeness, and ease of use. The study done by (Qin et al., 2021) reveals that satisfaction and informativeness are significantly related to MAR app use. When consumers evaluated the MAR application as, useful, enjoyable, and beneficial to knowledge attainment, they had a larger positive affective reaction. In practise, it gives information into MAR usage that benefit both customers and MAR app developers, as well as guides firm marketing strategy. Considering all of the definitions from previous scholars (R. Azuma et al., 2001; R. T. Azuma, 1997; Liao, 2019), we may conclude that MAR: a) "integrates real and virtual things in the real world; b) responds in real time; c) registers and aligns real and virtual objects; and d) executes and/or shows the augmented view on a mobile device".

MAR system is any system that possesses all of the aforementioned features. A successful MAR system allows users to focus on the application rather than the implementation (Papagiannakis et al., 2008). Many case-specific MAR apps have been developed in recent years, with the majority of them focusing on education, culture, tourism, although MAR games are currently gaining popularity. For example, Pokemon GO is a popular MAR software which offers a mobile AR game experience based on locations. Pokemon GO reveals many similarities with a previous similar MAR programme, Ingress, and though it was extremely popular in the days following its announcement, generating over $2 million USD in revenue per day, it is already losing momentum. Many researchers have investigated the effects of an augmented reality system on students enthusiasm for a visual arts course, as well as location-based MAR applications (Geiger et al., 2014).

## 2.2 MOBILE NATURE OF AR APPLICATIONS

Since most MAR apps are mobile in nature, they run on mobile/wearable devices such as tablets, smartphones, smart glasses, and, in rare cases PCs. A mobile application is classified as a MAR application if it includes the following features:

- **Input:** It takes into account the smartphone's numerous sensors (camera, microphone, gyroscope, GPS), and any associated device.
- **Processing:** It defines the kind of info that will appear on the screen of the mobile device. This may demand access to data on the device or in a remote database.
- **Output:** It displays its output on the mobile device's screen alongside the user's current view (i.e., It augments the virtual object into the real environment of the user).

Since the projected information is quickly placed on the actual world, AR glasses are the greatest option for ubiquitous mobile AR. Also, their computer capacity is limited, and most applications remain rudimentary as a result. Smartphones are also an ideal option due to their higher computing power and portability, however, to benefit from AR apps, the user must "point and hold." Tablets, PCs, and laptops become cumbersome and limited in their capabilities.

MAR faces additional challenges because of specific mobile platform needs, such as computational power and energy constraints. To work in new situations, it is