Combining Image Processing Techniques and Mobile Sensor Information for Marker-less Augmented Reality Based Reconstruction

I. S. Weerakkody

Index No: 14001527

Supervisor: Prof. N. D. Kodikara

January 2019

Submitted in partial fulfillment of the requirements of the B.Sc in Computer Science Final Year Project (SCS4124)

Combining Image Processing Techniques and Mobile Sensor Information for Marker-less Augmented Reality Based Reconstruction

I. S. Weerakkody

Declaration

I certify that this dissertation does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any university and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, be made available for photocopying and for interlibrary loans, and for the title and abstract to be made available to outside organizations.

Candidate Name: I. S. Weerakkody

………………………………………………

Signature of Candidate Date:

This is to certify that this dissertation is based on the work of
Mr. I. S. Weerakkody
Under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Prof. N. D. Kodikara

………………………………………………

Signature of Supervisor Date:

Dr. K. D. Sandaruwan

………………………………………………

Signature of Supervisor Date:

Abstract

Marker-less augmented reality based reconstruction using mobile devices, is a near impossible task. This largely due to the lack of processing power in mobile devices when considering vision based tracking for localization and to the lack of accuracy in mobile GPS when considering mobile sensor based approach for localization.

In order to address this problem this research presents a novel approach which combines image processing techniques and mobile sensor information which can be used to perform precise position localization in order to perform augmented reality based reconstruction using mobile devices. The core of this proposed methodology is tightly bound with the image processing technique which is used to identify the object scale in a given user image. Use of mobile sensor information was to classify the most optimal locations for a given particular user location.

This proposed methodology has been evaluated against the results obtained using 10cm accurate RTK device and against the results obtained using only the A-GPS chips in mobile device. Though this proposed methodology require more processing time than A-GPS chips, the accuracy level of this proposed methodology is outperforms that of A-GPS chips. And the results of the experiments carried out further convince that this proposed methodology facilitates improving the accuracy of position localization for augmented reality based reconstruction using mobile devices.

Keywords: augmented reality based reconstruction, position localization, image processing, mobile sensor information

Preface

This research presents a novel approach for position localization for augmented reality based reconstruction using mobile devices. The implemented algorithm which is used to calculate the object size in a given image have seen the use of an existing OpenCV template matching algorithm which is TM_CCOEFF_NORMED. All the other components in the research design in Chapter 3 are my own work. Furthermore the data of all the carried out experiments and their analysis which is produced in Chapter 5 are entirely my own work.

Acknowledgement

This dissertation marks the conclusion of my Bachelor's Degree in Computer Science at the University of Colombo School of Computing. It has been a fascinating four years which I learned much during this enjoyable time. First and foremost, I have to thank my research supervisors, Prof. N. D. Kodikara and Dr. K. D. Sandaruwan. Without their assistance and dedicated involvement in every step throughout the year, this study would have never been accomplished.

I would like to thank Dr. G. D. S. P. Wimalaratne, senior lecturer of University of Colombo School of Computing and Dr. K. L. Jayaratne, senior lecturer of University of Colombo School of Computing for providing feedbacks on research proposal and interim report to improve the study.

Thank you Mr. S. Jayasinghe instructor at Physical Department, University of Colombo for providing the university ground premises on numerous occasions to evaluate this research.

Thank you Mr. A. Weerasinghe Technical Manager at Retail Infomation Technologies (Pvt) Ltd for providing me with the RTK device which was used throughout this research.

Finally thank you Mr. E. M. K. De Silva and Mr. T. M. A. Dabare and all my friends for helping me build the object that was used for evaluations.

I. S. weerakkody
January 2019

Table of Contents

List of Figures

List of Tables

List of Acronyms

A-GPS	Assisted Global Position System
D-GPS	Differential Global Position System
AR	Augmented Reality
HMD	Head Mounted Display
RTK	Real-Time Kinematic
SURF	Speeded Up Robust Features
SIFT	Scale-Invariant Feature Transform
ORB	Oriented FAST and rotated BRIEF
DoG	Difference of Gaussian

Chapter 1 - Introduction

1.1 Background to the Research

Augmented reality is the integration of digital information with the user's environment in real time. Unlike virtual reality, which creates a totally artificial environment, augmented reality uses the existing environment and overlays new digital information on top of it. This information is mostly animations or contextual digital information.

Figure 1.1 depicts Paul Milgram's reality-virtuality continuum which further explains the concepts of augmented reality and virtual reality.

Figure 1.1: Simplified representation of a "virtuality continuum".

The area between the two extremes, where both the real and the virtual are mixed, is called mixed reality. This in turn is said to consist of both augmented reality, where the virtual augments the real, and augmented virtuality, where the real augments the virtual (Milgram & Kishino, 1994).

When considering the term augmented reality, augmented reality has two main genres such as marker-less augmented reality (Marker-less AR) and marker-based augmented reality (Marker-based AR).

"Marker-less AR" is a term used to denote an augmented reality application that does not require any pre-knowledge of a user's environment to overlay 3D content into a scene and hold it to a fixed point in space. This approach uses technologies such as vision based tracking or mobile sensor information to do the appropriate augmentation.

Until recently, most augmented reality fell under the category of “Marker-based AR” which required the user to place a “tracker”, an image encoded with information that is translated by complex software to produce a 3D object that maintains spatial orientation within a scene in order to achieve the desired effect.

This research mainly concerns augmented reality based reconstruction for relatively large objects such as partially destroyed ancient monuments. With this proposed solution users would have the opportunity to see ancient monuments in their full structure at their original location. Hence, users will have the chance to relive the glory and beauty of these ancient monuments. When considering augmented reality reconstruction for such large objects, aspects such as position accuracy, stability and hardware support are important facts.

Comparing the marker-less augmented reality approach against marker-based augmented reality approach, when considering an aspect such as position accuracy maker-based approach looks more accurate. But when considering an aspect such as stability, maker-less approach looks more promising. When considering an aspect such as hardware support marker-less augmented reality is usually not supported for desktops. When considering mobile devices mobile devices usually supports both these approaches (Jack C.P. Cheng, et al., 2014).

When using a marker-based approach for augmented reality based reconstruction one of the main drawbacks is that using markers would clutter the physical environment and also the task of placing markers in appropriate positions is a time consuming one which will result in an additional overhead. And also there is a huge possibility that the markers could be destroyed or removed with time which makes it impossible to track. Hence, using a marker-based approach for augmented reality based reconstruction may not be the most suitable option.

When considering marker-less augmented reality there are two main approaches. One is using vision based tracking and the other one is by using sensor information. Even though some of the drawbacks of marker-based approach can be somehow controlled using a marker-less approach, position accuracy is relatively low when using a marker-less approach. This position accuracy is solely dependent on the accuracy of the localization technique. When considering an aspect such as localization accuracy, hardware specification of the target devices become a resounding factor.

For an instance when using sensor based approach for the localization, accuracy of a typical GPS (Global Positioning System) receiver is approximately 5 meters. But by using devices such as D-GPS (Differential GPS) receivers it is possible to achieve an accuracy level of 1 meter. But mobile devices are equipped with A-GPS (Assisted GPS) which has an accuracy of approximately 5-8 meters (Paul A. Zandbergen & Sean J. Barbeau, 2011). A-GPS has relatively low accuracy compared to a typical GPS receiver or a D-GPS receiver. Since mobile devices (smartphones) usually have relatively low cost GPS chips such as A-GPS and since the ancient monuments which are to be reconstructed using augmented reality are mostly in rural or forest areas, accurate localization using mobile devices become a difficult task. In addition, the accuracy of GPS varies depending on the number of GPS satellites and is reduced in GPS interfering spots. Hence, it is not possible to do augmented reality based reconstruction by only using GPS localization approach for ancient monuments (Soyoung Hwang & Donghui Yu, 2012).

Another localization method is to use vision based localization approach. This basically deals with a set of reference images that are being used to calculate the location and the viewpoint of the user (Didier Stricker & Thomas Kettenbach, 2001). Using this technique the user’s position will be calculated according to the matched features. In this sort of an approach the number of reference images or the number of features are proportional to the accuracy of the location value given. It is easy to assume that increasing the number of reference images is the most optimal solution to achieving an increased accuracy, but as the number of reference images increases the number of comparisons will increase as well. Depending on the hardware specification of the device, doing high number of comparisons will be a time consuming operation. And also devices with higher processing power will be needed in order to obtain a timely feedback. As the processing power of a mobile device is relatively low it is difficult to achieve a higher accuracy level in a less amount of time by only using this approach.

Even though at present there are many ongoing researches and studies on augmented reality, precise augmentation using marker-less approach has been a constant issue. As mobile devices has a relatively low processing power and as they are equipped with relatively low cost GPS chips, doing precise augmentation using the maker-less approach for mobile devices has become an even bigger issue in recent past.

In this research it is expected to combine the two approaches of mobile sensor information based augmented reality and vision based tracking for outdoor augmented reality to overcome the disadvantages of using a single approach. Also this research will result in a new method of position localization for mobile devices which uses both image processing techniques and mobile sensor information.

1.2 Research Question

The topic of augmented reality based reconstruction has been approached in the past using highly sophisticated equipment such as D-GPS receivers (Gerhard Schall, et al., 2009), HMD (Head Mounted Device) and Laptops (Vassilios Vlahakis, et al., 2001) which are relatively expensive compared to a mobile device. Rather than these sophisticated equipment if this is to be done using mobile devices (smartphones), then the following issues would have to be solved.

Considering mobile sensor based augmented reality

- When using A-GPS, its precision and update rate are not sufficient for precise and accurate tracking. May require additional hardware equipment to improve the precision (Gerhard Schall, et al., 2009).
- When using inertial tracking (IMU (Inertial Measurement Unit) sensors such as accelerometers, gyroscopes and magnetometers) unfortunately, inertial sensors are very susceptible to drift over time for both position and orientation (Mark Billinghurst, et al., 2014).
- GPS accuracy is not consistent in all areas. Forest and around building the GPS accuracy is relatively low (Soyoung Hwang & Donghui Yu, 2012).

Considering vision based approach

- Marker-less vision based augmented reality requires relatively high processing power for high level of accuracy (Didier Stricker & Thomas Kettenbach, 2001).
- In a marker based approach, markers in the environment typically leads to more accurate localization results, but it requires intrusive and accurate positioning of markers which result in an additional overhead and also it will clutter the physical environment (Ludovico Carozza, et al., 2012). Since markers could be fully or partially destroyed with time, accuracy level of the localization would decrease with time.

How to do marker-less augmented reality based reconstruction with increased efficiency and accuracy using image processing techniques and mobile sensor information?

1.3 Aims and Objectives

As mentioned in the above section the main problem that arises when considering augmented reality based reconstruction using mobile devices for large objects is that the position localization accuracy achieved using the A-GPS chips in the mobile device is not accurate enough to perform precise augmented reality based reconstruction. Hence, the definitive goal of this research project is to combine image processing techniques and mobile sensor information in order to achieve less time consuming and accurate position localization using mobile devices for augmented reality based reconstructions of large objects. The following objectives will also be achieved throughout this research.

- Develop an algorithm using existing template matching algorithms. This new algorithm has the capability to identify the size of an object in an image.
- Improve and train this new algorithm using a set of pre taken images to perform accurate position localization.
- Evaluate the proposed position localization method against a RTK device.

1.4 Justification for the research

Augmented reality is becoming mainstream and have become a worthwhile topic in many industries. Augmented reality development can be applied towards gaming, entertainment, marketing, education, fashion, art, and so much more. This is an exciting new technology that is being improved every day. One of the main reasons for this rise of augmented reality is the rise of mobile devices (smartphones). The number of smartphone users have surpassed 2 billion in 2016 and is expected to increase to 2.86 billion by 2020 (Number of smartphone users worldwide from 2014 to 2020 (in billions)., 2016).

Even though there are several augmented reality based applications that are being constantly used with mobile devices, augmented reality based reconstruction is a topic that has not been constantly addressed using mobile devices. This is largely due to the lack of GPS localization accuracy of the mobile device.

This research will aim to provide a new method of position localization for mobile devices which can be used for augmented reality based reconstruction purposes of ancient monuments.

1.5 Methodology

This research will combine image processing techniques and mobile sensor information in order to do precise augmented reality based reconstruction for large objects. In this research the main use of mobile sensor information would be to reduce the image processing time. The following steps illustrates the methodology in which this research was conducted.

- Select a set of known locations from the site that is going to be reconstructed (number of locations depends on the scale of the object and the accuracy level).
- Obtain a reference image from each of the selected locations.
- Geo tag the obtained reference images using a RTK device (accurate up to 10cm).
- Associate a bearing value with each of these selected location. Bearing value is calculated using the direct angle from the location to the object.
- For each of the selected set of locations, implement a functions where it has the scale of the object as the co domain and distance from the object to the user location as the domain. Hence, this function represent the scale deviation of the object against the deviation of the distance from the user to the object.
- Once a user enters to the site that is to be reconstructed, obtain a geo tagged user image which contains the full structure of the object when it is seen from that particular side.
- Based on the geo tag of the user image and the geo tags of the reference images, obtain the set of most optimal locations.
- Using the user image and the set of reference images taken from the most optimal locations do the relevant template matching and obtain the best matching location.
- Identify the scale of the object in the user image.

- With the use of the function at the best matching location and the scale of the object in the user image, obtain the distance to the object from the user's location.
- Calculate the accurate GPS value of the user using the distance from the user to the object, the geo tag of the best matching reference image and the bearing associated with that best matching location.
- Do the appropriate augmentation based on that obtained accurate GPS value. And keep track of the augmented object using extended tracking feature. Below figure (Figure 1.2) further illustrates these steps.

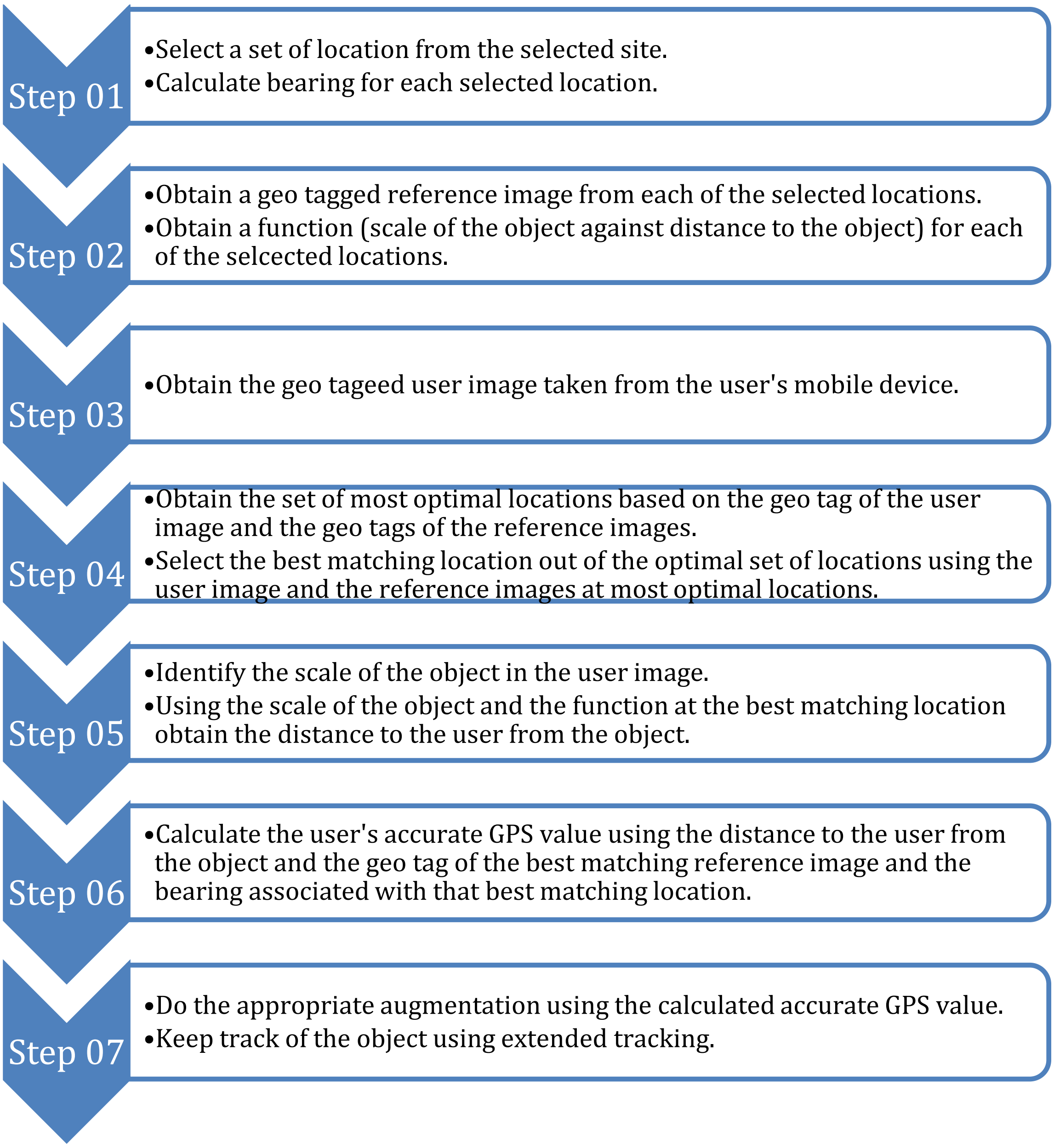


Figure 1.2: Research methodology.

1.6 Outline of the Dissertation

The rest of this thesis is organized as follows: The second chapter is dedicated for the literature review to discuss related work on different technologies which are used to perform augmented reality based reconstruction. The third chapter elaborates the design. The ways in which this research was conducted have been discussed in this chapter. The fourth chapter discusses the implementation phase of this project. This chapter describes about research challenges, implementation strategies, proposed solutions, etc. Evaluation phase comes under the fifth chapter. It consists of the main evaluation strategies used in the study and evaluation results. The last chapter presents the conclusion and future work. It mentions the future enhancements possible with the results of this study.

1.7 Delimitations of Scope

1.7.1 Scope

The following section explains the content that will be covered and to what extent they will be covered in this research. Though this research answers a typical question in the area of augmented reality this research will also provide an improvement to the location values obtained by only using A-GPS receivers of a mobile device. This improvement will be achieved by combining image processing techniques together with mobile sensor information. Using this new approach for localization using mobile devices, this research will aim to achieve a higher level of precision for marker-less augmented reality. It is expected to use this improved marker-less augmented reality approach for augmented reality based reconstruction using mobile devices for partially destroyed ancient monuments.

1.7.2 Delimitations

Following delimitations were set as boundaries when conducting this research. Since this research proposes a solution for outdoor augmented reality based reconstructions for large objects, some of the delimitations are based on environmental factors. And also since this research uses image processing techniques some of the following delimitations are based on the physical attributes of the object that is going to be reconstructed.

- This proposed solution will only be for outdoor usage using mobile devices.
- When considering improvement on mobile GPS localization, this research will only concentrate on use cases where there is augmented reality based reconstruction. Hence, in this research it is assumed that the user of the mobile device is facing the site that is going to be reconstructed and the full structure of the monument is included in the user image.
- Even though this proposed approach is a marker-less augmented reality approach this proposed solution will require some sort of pre knowledge about the environment. This knowledge will basically include the set of geo tagged reference images and the deviation of the scale of the object in the user image against the distance to the object from all angles.
- When considering vision based tracking different lighting conditions will not be considered in this research. Hence, this proposed approach will only be for day time usage.
- When considering mapping and augmenting the appropriate object, occlusion handling will not be considered in this research.
- When selecting an appropriate use case for this augmented reality based reconstruction, for evaluation purposes a physically existing site is selected.
- The proposed methodology will not be applicable to sites that are fully destroyed since this research uses vision base tracking.

1.8 Conclusion

This chapter laid the foundation for this thesis. It introduce the research problem with a thorough explanation of the background and then it moved on and justified the research problem. A brief explanation of the methodology followed by the scope outlines the areas which will be covered in this research. Delimitations which were taken into consideration during this research was then explained. On these foundations, the dissertation can proceed with a detailed description of the research.

Chapter 2 - Literature Review

2.1 Introduction

The previous chapter described an overview of this research. It includes the background, scope, delimitations, research question, aims and objectives, justification and a brief explanation of the methodology. This chapter contains an in detailed description about several existing augmented reality approaches, image processing techniques, Linear Regression method and the great circle distance method with Vincenty equation.

2.2 Marker-less augmented reality using sensor information

According to a research done on global pose estimation using multi sensor fusion, outdoor augmented reality typically requires tracking in unprepared environments. For global registration, Global Positioning System (GPS) is currently the best sensing technology, but its precision and update rate are not sufficient for high quality tracking. In this research the researchers have presented a system that uses Kalman filtering for fusion of Differential GPS (DGPS) or Real-Time Kinematic (RTK) based GPS with barometric heights and also for an inertial measurement unit with gyroscopes, magnetometers and accelerometers to improve the transient oscillation. They have developed a hardware tracking module using Differential GPS (DGPS) or Real-Time Kinematic (RTK) based GPS. This hardware tracking module is suited for use with handheld augmented reality devices due to its small weight and form factor (Gerhard Schall, et al., 2009). A typical DGPS receiver is accurate for less than one meter (Vassilios Vlahakis, et al., 2001). Hence, with the hardware module these researchers were able to obtain higher accuracy rate for position localization.

But when considering a mobile phone, mobiles phones are not equipped with facilities such as DGPS. Mobile phone has three main positioning technologies Assisted GPS (A-GPS), WiFi positioning and cellular network positioning. When considering an iPhone, the accuracy levels using these three technologies would be 8meters, 74 meters, 600 meters respectively (Zandbergen, 2009). This level of accuracy is not enough for a precise and accurate augmented reality based reconstruction.

And also this accuracy is highly dependent on environmental factors. In addition, the accuracy of GPS varies depending on the number of GPS satellites and is reduced in GPS interfering spots such as in a forest or around buildings (Soyoung Hwang & Donghui Yu, 2012).

In order to increase the accuracy of the location data given when only using the GPS in smartphones, in 2009 Paul A Zandbergen carried out a research which used other sensors information such as gyroscope data and compass data which are built in to smartphones together with the GPS information. And in this research it is proven that this proposed system achieves a reasonable accuracy level in GPS Interfering Area as well as open space areas when compared to a system that only uses GPS (Zandbergen, 2009).

There have been hybrid approaches which combined the data of the compass, accelerometer and the GPS together in order to calculate where the object should be augmented in the field of view without any actual processing of the real image. With the compass it is possible to tell the direction which the device is pointing at and the accelerometer is used to calculate orientation of the device using gravitation to its advantage (Qing Hong Gao, et al., 2017).

2.3 Augmented reality using Vision based tracking approaches

Vision based tracking for augmented reality has become increasingly popular in recent times due to the minimal hardware requirements and the ubiquity of mobile devices such as smartphones and tablets which feature both a camera and screen, making them ideal platforms for augmented reality technologies. Even though mobile devices (smartphones) are not equipped with the same computational power as laptops or desktops, the considerable computation power they possess had made vision based tracking through mobile devices even more popular and interesting approach for marker-less augmented reality. The optical sensors used for vision based tracking can be divided into three major categories such as Infrared Sensors, Visible Light Sensors, and 3D Structure Sensors. Out of the above three categories considering Visible Light Sensors, which is the most common category, cameras that are suitable for this method of tracking can be found in devices ranging from laptops to smartphones and tablet computers, and even wearable devices. For an augmented reality based reconstruction system which is a video see through system Visible Light Sensors are really useful as they can be used both for registering the virtual object in the real world and for the real world video background which is shown to the user. This visual based tracking is also divided into three major types Fiducial tracking, Natural Feature tracking, and Model Based tracking (Mark Billinghurst, et al., 2014).

2.4 Marker-based augmented reality using vision based approaches

Fiducial markers can be introduced and localized in the environment, so that online localization can be achieved by simply recognizing them using an appropriate sensing pipeline. Fiducial markers typically consists of small colored LEDs or pieces of paper. This tracking method may lead to more accurate localization results than the maker less approaches. But it requires intrusive and accurate positioning of markers within the environment and markers could clutter the physical environment as well (Ludovico Carozza, David Tingdahl, et al., 2012). And also there is a certain possibility that these markers could be fully or partially destroyed with the time due to environmental effects. Marker less systems, on the other hand, are not invasive and they won’t clutter the environment.

2.5 Marker-less augmented reality using vision based approaches

As the computational power of devices used for augmented reality applications improved, it became possible to register the pose of the camera, in real time, using features which already exist in the natural environment. Complicated image processing algorithms are used to detect features in the captured images which are unique in their surroundings, such as points, corners, and the intersections of lines. In a research done in 2014, for each of these features, a unique "descriptor" is calculated which allows for identification and differentiation of each feature. By matching features detected in the scene with those detected in the object to be tracked, the pose can be computed using similar algorithms as those used in the Fiducial marker techniques (Mark Billinghurst, Adrian Clark, & Gun Lee, 2014). Some of the more common natural feature detection and description algorithms include SIFT, SURF, BRIEF, ORB, BRISK and FREAK.

When considering vision based tracking for outdoor augmented reality, performance is a critical issue. To increase the accuracy level of a vision based tracking system the number of reference images should be increased. When increasing the number of reference images, then the number of comparisons will increase as well. Hence, the tracking time would increase rapidly. Since mobile phones are not equipped with high computational power (with respect to laptops and other devices used for augmented reality) this is a major problem that needs to be solved. And also the when considering an outdoor environment lighting condition is also a fact to be concerned with (Didier Stricker & Thomas Kettenbach, 2001).

Also Marker-less Augmented Reality (AR) registration using the Standard Homography Matrix has a low accuracy when it is used for image based registration (Ebrahim Karami, st al., 2015).

2.6 Comparison of description algorithms

When considering vision based tracking image matching becomes an important aspect that should be considered. As proposed by Lowe in 2004, SIFT, SURF, and ORB, work against different kinds of transformations and deformations such as scaling, rotation, noise, fisheye distortion, and shearing. These three types of description algorithms have been evaluated using three sets of parameters. These parameter include number of key points in images, the matching rate, and the execution time required for each algorithm. Also he have shown that, ORB is the fastest algorithm while SIFT performs the best in the most scenarios. For special case when the angle of rotation is proportional to 90 degrees, ORB and SURF outperforms SIFT and in the noisy images, ORB and SIFT show almost similar performances. In ORB, the features are mostly concentrated in objects at the center of the image while in SURF, SIFT and FAST key point detectors are distributed over the image. The researcher who did this research have proven their conclusions by producing descriptive results of each and every experiment (Lowe, 2004).

Scale Invariant Feature Transform also known as SIFT was first proposed by Lowe which solves the image rotation, affine transformations, intensity, and viewpoint change in matching features. The SIFT algorithm has 4 basic steps. First is to estimate a scale space extrema using the Difference of Gaussian (DoG). Secondly, a key point localization where the key point candidates are localized and refined by eliminating the low contrast points. Thirdly, a key point orientation assignment based on local image gradient and lastly a descriptor generator to compute the local image descriptor for each key point based on image gradient magnitude and orientation (Yan Ke & Rahul Sukthankar, 2004).

Speed up robust feature also known as SURF approximates the DoG with box filters. Instead of Gaussian averaging the image, squares are used for approximation since the convolution with square is much faster if the integral image is used. Also this can be done in parallel for different scales. The SURF uses a BLOB detector which is based on the Hessian matrix to find the points of interest. For orientation assignment, it uses wavelet responses in both horizontal and vertical directions by applying adequate Gaussian weights. For feature description also SURF uses the wavelet responses. A neighborhood around the key point is selected and divided into sub-regions and then for each sub-region the wavelet responses are taken and represented to get SURF feature descriptor. The sign of Laplacian which is already computed in the detection is used for underlying interest points. The sign of the Laplacian distinguishes bright blobs on dark backgrounds from the reverse case. In case of matching the features are compared only if they have same type of contrast (based on sign) which allows faster matching (JongBae Kim & HeeSung Jun, 2008).

2.7 Location positioning for an indoor environment

In a research done in 2010 by Gerhard Reitmayr, et al. the researchers have presented a vision-based location positioning system using augmented reality technique for indoor navigation. The method proposed in this paper automatically recognizes a location from image sequences taken of indoor environments, and it realizes augmented reality by seamlessly overlaying the user's view with location information. In order to obtain these location positions in this research, the researchers have used a pre-constructed an image database and location model, which consists of locations and paths between locations, of an indoor environment. Location is recognized by using prior knowledge about the layout of the indoor environment. To carry out the experiments of this research the researchers have used highly sophisticated equipment such as wireless camera, mobile pc has been used together with a Head Mounted Display. In this research the researcher haven't used GPS-based Location Positioning systems to a great extent as GPS radio signals have difficulty in penetrating building walls. Since this proposed approach is for indoor usage and indoor environments usually consists of high number of walls the accuracy level given by a GPS-based Location Positioning systems will be relatively low. And results of the experiments carried out in this research have shown that this proposed system produces an average location recognition success rate of 89% in an indoor environment (Gerhard Reitmayr, et al., 2010).

2.8 Simultaneous Localization and Mapping for Augmented Reality

According to a research done in 2002 by Welch and E. Foxlin, in order to provide accurate registration of augmented visuals using an augmented reality system, these systems have to deal with two fundamental technical challenges. Given below are the above mentioned two challenges.

1. The current view of the real world that needs to be augmented.
2. The virtual object geometry and its accurate registration with the real world.

Generally the first challenge is commonly referred as tracking problem (Brunelli, 2009) and the latter challenge is referred as the authoring problem.

Even though Simultaneous Localization and Mapping (SLAM) provides an inherent tracking solution, it does not provide any reference to a known, global location. Therefore information that is referenced to such a real location, for example through a GPS position, cannot easily be rendered in a purely SLAM-based system. According to this research the researchers have developed a panoramic mapping and tracking approach that is integrated with other sensors to provide global registration. Hence, the method proposed in this paper is based on a simultaneous mapping and tracking approach, operating on cylindrical panoramic images (Welch & E. Foxlin, 2002).

2.9 Template matching

In 2009, Brunelli introduced template matching technique as a digital image processing technique. In template matching basically there are two images. One is referred as the template image (T) and the other one is commonly referred as the source image(S). With the use of an appropriate template matcher it is possible to find small parts of a source image which matches with the template image. When considering a typical template matching algorithm they are variant to geometrical transformations. This means that if the source image has undergone a scale transformation or a rotation transformation it makes it difficult to find the matching part of template image in the source image. And also when considering a typical template matching algorithm they are variant to brightness and contrast as well (Brunelli, 2009).

In 2010, a pair of researchers were able to come up with an algorithm name Ciratefi which is scale, rotation, brightness and contrast invariant. Though this algorithm has many advantages compared to other existing template matching algorithms, even the researchers who came up with this algorithm has admitted that the proposed template matching algorithm named Ciratefi is slow when compared to other template matching algorithms (Sidnei Alves de Araújo & Hae Yong Kim, 2010). Hence, it is difficult for a user to get a timely feedback using this algorithm which makes it difficult to use such algorithm for localization. As users won't stay in a certain location for considerable amount of time.

2.10 Linear Regression

Regression is a method of modelling a target value based on independent predictors. This method is mostly used for forecasting and finding out cause and effect relationship between variables. Regression techniques mostly differ based on the number of independent variables and the type of relationship between the independent and dependent variables.

Simple linear regression is a type of regression analysis where the number of independent variables is one and there is a linear relationship between the independent(x) and dependent(y) variable. To perform this linear regression there are several different techniques. Some of them include Least Square Estimation, Generalized Least Square Estimation, Maximum Likelihood Estimation and Bayesian Estimation. When considering the Least Square Estimation technique following are the list of applied formulas in Least Square Estimation technique.

$$a = \frac{(n\sum xy) - (\sum x \sum y)}{(n\sum x^2) - (\sum x)^2} \tag{1}$$

$$b = \frac{(\sum y \sum x^2) - (\sum x \sum xy)}{(n\sum x^2) - (\sum x)^2} \tag{2}$$

$$y = ax + b \tag{3}$$

Equation (3) represents the linear regression function. Variable 'a' and 'b' in that function are calculated using equation (1) and equation (2). In equation (1) and equation (2) the variable 'n' represents the number of matching x and y pairs that were used (A. F. Seber & J. Lee, 2003).

2.10.1 Use of linear regression in this research

Linear regression is used in this research to calculate the distance to the user from the object. This calculation is done using the user image sent by the user. As previously mentioned in the methodology, first step after selecting a site to reconstruct is to select a set of locations around the object in that site in order to obtain reference images. For each of these selected locations a linear regression is calculated by using a set of user images taken near to that particular location together with the Linear Square Estimation technique. For a single location, obtain all the user images taken near to that particular location. For a single user image, distance to the object from the place where the user image was taken is considered as 'x' value and the identified scale of the object in that particular user image is taken as 'y' value. Take all the pairs of 'x' and 'y' values for a single location using all the user images taken near to that particular location and obtain the linear regression by using equation (1), equation (2) and equation (3). In this research only three user images were taken near a particular location. Hence, value of 'n' will be three. Since, there will only be three pairs of 'x' and 'y' values for a single location. Similarly it is done for all the locations around the selected object that is going to be reconstructed.

2.11 The great circle distance method with Vincenty equation

Great Circle Distance method with Vincenty equation was used throughout this research to accurately calculate the GPS location values of the user. Three input variables were used as inputs to this equation. They were the Longitude, Latitude values of the best matching location, distance to the user from the best matching location and the bearing associated with that best matching location. Compared to other equations Vincenty geodesic equation takes into account the distance geometry approach ellipsoidal. While the distance geodesic equation Haversine uses an approach perfectly spherical geometry (sphere). When considering an aspect such as over speeding detection Vincenty have better performance in compared to Haversine equation (Kifana & Abdurohman , 2012).

2.12 Summary

This chapter provided an in detailed description of the background studies which are related to this research study. Next chapter will describe on how those related studies were combined to produce the research design which was used in throughout this research

Chapter 3 - Design

3.1 Introduction

This chapter will serve as a baseline description of the research design used throughout this research. Moreover, this chapter will provide a high level description of the research design which illustrates how all the steps in the research design is combined to produce the final output for a given user image.

3.2 Obtaining reference images

Initially a set of locations will be selected from the site that is to be reconstructed. Below figure (Figure 3.1) depicts how those initial locations were selected. The number of locations and the distance from a location to the object (*d* meters as depicted in the figure) is highly dependent on the scale of the object.

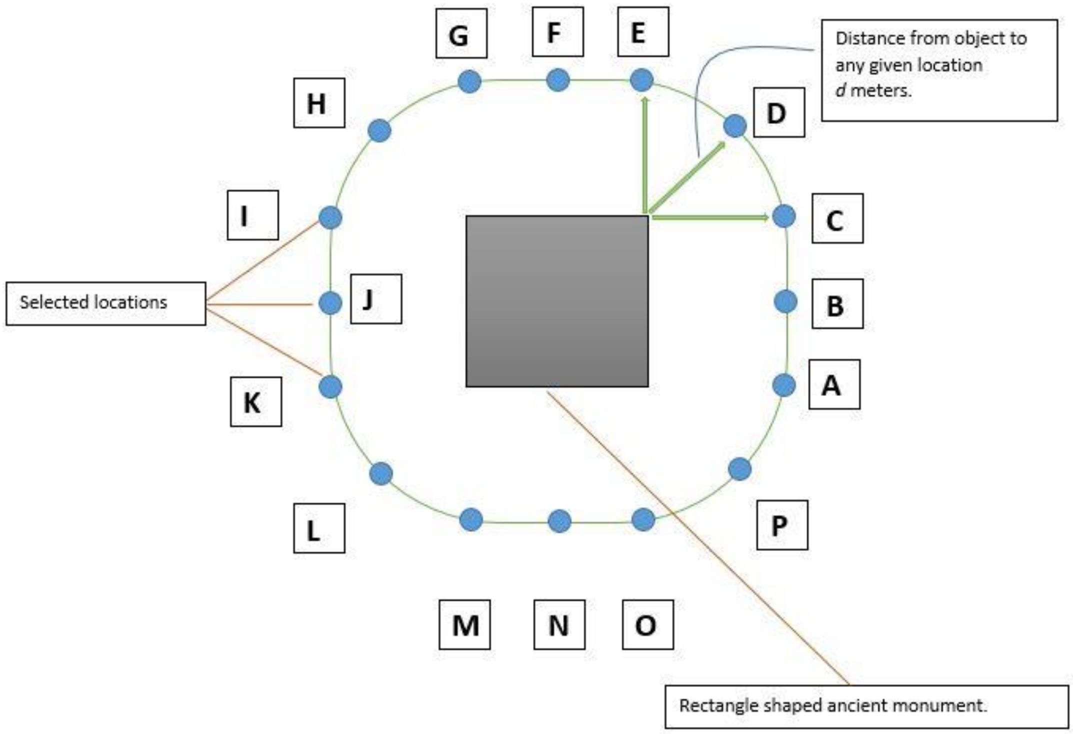


Figure 3.1: Location selection to obtain reference images.

From every one of these locations, a single reference image will be captured. All of these reference images are geo tagged and all of these locations will have a bearing value associated to them. For each location this bearing value is calculated based on two points. One point is the point at which the reference image was taken and the second point is the closest point from the first point to the object itself. Longitude and Latitude values of these locations will be measures using an accurate device (RTK device). When considering these reference images, these images will have the full portion of the object when it is seen from their respective angle and also it will be zoomed into the object so that reference images won’t contain much of the surrounding environment.

3.3 Deriving functions for each location

Once we have selected the set of locations for reference image capturing, out of those locations select a single location. For that selected location obtain a set of three images in the following order.

- First image – *d* meters away from the object.
- Second image – *(d+1)* meters away from the object.
- Third image – *(d+2)* meters away from the object.

Below figure (Figure 3.2) depicts how the set of three images were captured for a certain reference location in the site that was used for evaluation of the proposed methodology.

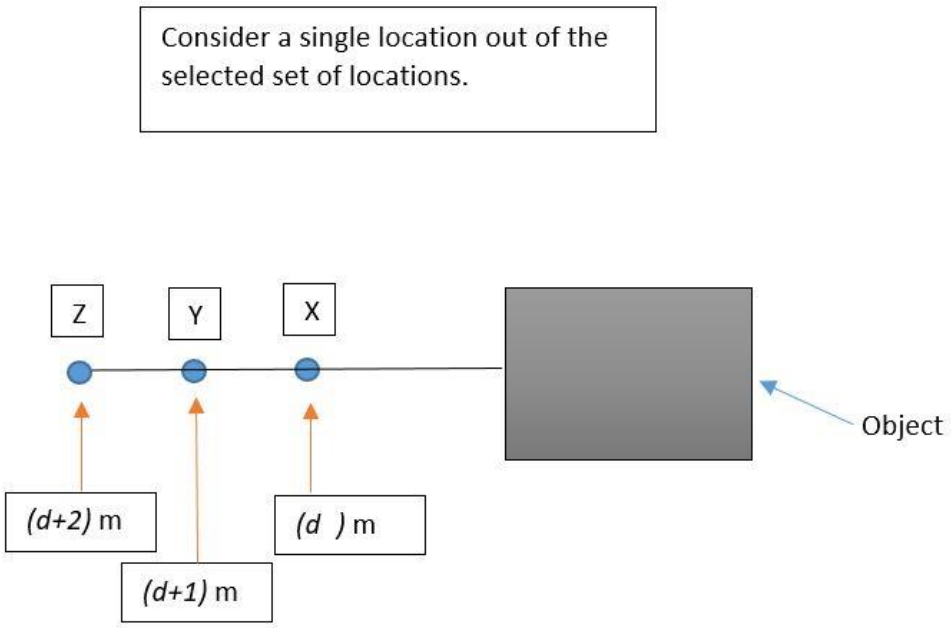


Figure 3.2: Obtaining images to derive functions.

After obtaining those set of images, identify the scale of the object (S_1, S_2 and S3) in each of the above three images using the implemented new algorithm. Since we are using linear regression to obtain the function at a certain location, use these three values and their respective distance from the object to obtain the function. When doing so, the three point will be taken as (d, S_1), ($d+1$.S_2) and ($d+2$, S_3).Using the above three points and Least Square Estimation obtain the function for that selected location. Similarly do this for every other selected location and obtain functions for each of those locations.

3.4 Filtering the reference images to find the most optimal set of locations

In this type of a research which interacts with the user in real time, getting a timely feedback is essential. And most of that time that is taken to provide the feedback depends on the vision based tracking system that is being used. This research uses a modified version of existing template matching techniques for vision based tracking and below figure (Figure 3.3) depicts the filtering mechanism which was used in this research to reduce the time taken to do the vision based tracking.

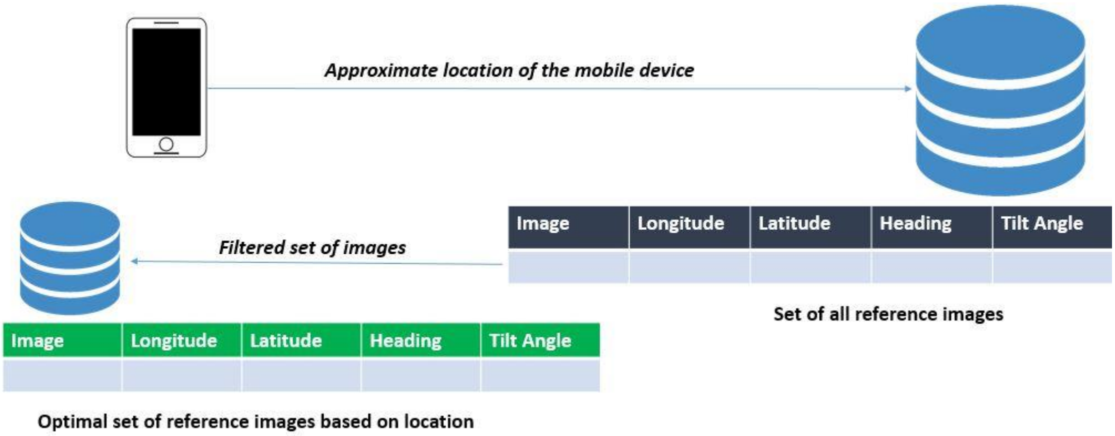


Figure 3.3: Process of filtering the reference images.

Once a user arrives with a mobile device to the site that is going to be reconstructed, obtain the approximate location values (Longitude and Latitude) of the user using A-GPS in the user's mobile device. This GPS value can be extracted from the geo tagged user image of the user. Using this approximate location value of the user and the accurate geo tags of the reference images filter the set of all reference images and obtain the most optimal set of reference images based on the user's current location. Rather than comparing all the reference images with the user image, in this approach only few reference images are compared with the user image. Hence, this will reduce the performance time of this implemented solution.

3.5 Using template matching to find the best matching location

As briefed in the literature, to find the best matching location based on a user image and a particular set of pre taken reference images there at two main image processing techniques that are capable of this task. One is by using a description algorithms such as SIFT, SURF, ORB and etc. And the other technique would be to use a template matching algorithm.

Considering the first approach which uses description algorithms, as mention above these description algorithms are highly based on the features in the reference image and the user image. In order to use such algorithm to find the best matching location there should be high number of features in the object itself rather than the background for any given image of the object since the features in the background are subject to change with the time. But when considering the features in an image with the object most of the features detected using these algorithms are outside of the object rather than in it. Hence, such approach to use description algorithms to find the best matching location would be inappropriate.

Below figure (Figure 3.4) depicts the features that are detected when using description algorithms.

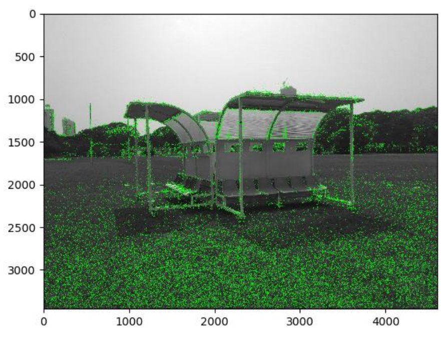
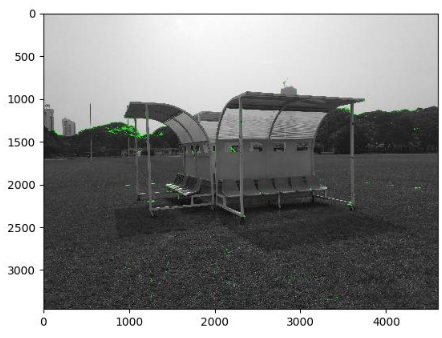
SURF Features	**ORB Features**
	

Figure 3.4: Detected features using description algorithms.

When considering template matching technique, template matching algorithms don't have these disadvantage which are there in description algorithms. But as described in the literature review typical template matching algorithms are not scale invariant. Hence, it is difficult to perform template matching when the scale of the object in the source image does not match with the scale of the object in the template. Template images are the set of reference images. Since taking reference images is a onetime process object scale in those template images are fixed. Source images are the images taken by the user. Hence, the scale of the object in the source images varies depending on the distance from the user to the object. To perform template matching in this kind of a situation, existing template matching algorithms were modified in the following way.

1. Obtain the user image and set it as the source image.
2. Scale down the source image to a predefined value in order to reduce the number of pixels in the user image and improve the performance of this algorithm,
3. Obtain the most optimal set of reference images based on the user's current location.
4. Scale down all the optimal reference images to a certain value (this value depend on the scale of the object) and set them as template images for the initial iteration.
5. Perform typical template matching algorithm using user image and each of the scaled down template images and compute a correlation score.
6. Increase the scale of the template images minimally and use them as template images for the next iteration and repeat step 4.
7. Until the scale of template image reaches a predefined value (this value depend on the scale of the object) repeat step 4 and step 5.
8. Obtain the highest 10 correlation scores of each template and select the template which has the largest sum over its 10 highest correlation scores and set it as the best matching template.

Figure 3.5 depicts how the modified version of the template matching algorithm is performed on a single optimal image.

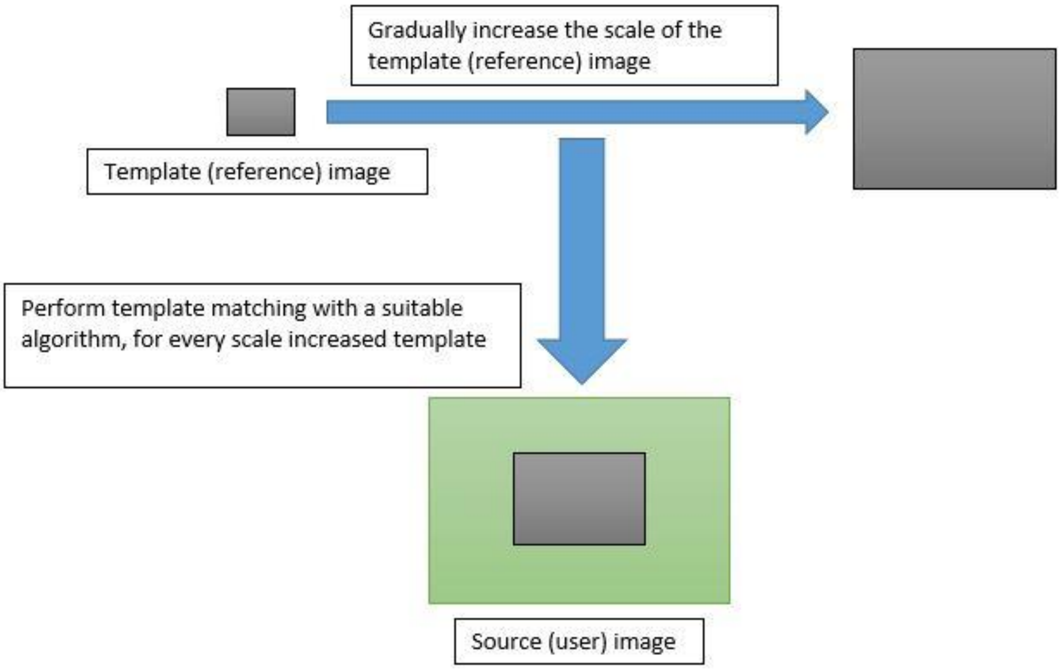


Figure 3.5: Implemented template matching algorithm.

Set the best matching location, as the geo tag of the best matching template (reference) image.

3.6 Obtaining the scale of the object in the user image

After obtaining the template image that best matches with the user image, select 10 scale values of that template image which have the highest correlation score. Calculate the average of those 10 scale values and set it as the scale of the object in the user image. Since over 90% of the template image contained the object and the full portion of the object was included in the template image it can be assumed that these two values are approximately the same.

3.7 Obtaining user's accurate location value

Obtain the function at the best matching location which represents the deviation of the scale of the object in the user image against the deviation of the distance from the user to the object. Since the scale of the object in the user image is known as it was produced in the last step, using that scale value solve the function to obtain the distance from the user to the object. Using that distance from the user to the object and the distance from the best matching location to the object obtain the distance from the user to the object.

Using the obtained distance to the user from the best matching location and the geo tag of the best matching reference image and the bearing value associated with the best matching reference image calculate the user's accurate longitude and latitude values using The Great Circle Distance method with Vincenty equation.

3.8 Augmentation

By using the accurate location values obtained in the previous step, perform the appropriate augmentation. Furthermore after the initial augmentation to keep track of the object when the user is moving around using extended tracking.

3.9 High level description of the research design

The following figure (Figure 3.6) depicts the high level description of the research design which further illustrates how to calculate user's accurate GPS location by using a geo tagged user image together with the proposed methodology.

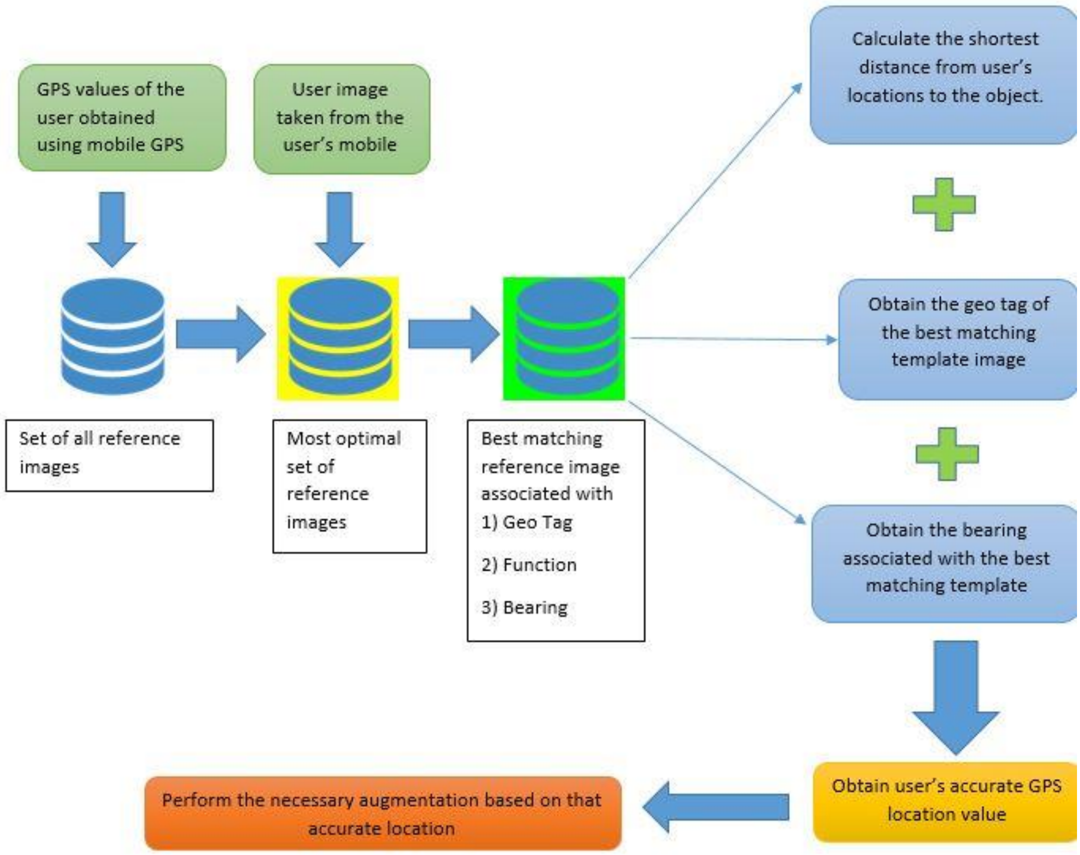


Figure 3.6: Research design.

3.10 Summary

This chapter provided an overall in detailed picture of the research design used throughout this research study. Following chapter will describe on how this proposed research design was used to conduct the implementations of the experiments.

Chapter 4 - Implementation

4.1 Introduction

This chapter mainly focuses on the implementation of the experiments which were done according to the research design described in the previous chapter. In addition to the implementation, different tools and existing algorithms used in this research will also be discussed. Moreover, information of data set which was used to carry out these experiments will be described further.

4.2 Obtaining reference images

When obtaining reference images, as described in section 3.1 a set of 32 locations around the object were selected and reference images were taken from those set of selected locations. Below figures (Figure 4.1 and Figure 4.2) depicts how the locations were selected in the site that was used for evaluation.

Figure 4.1: Obtaining 32 reference images.

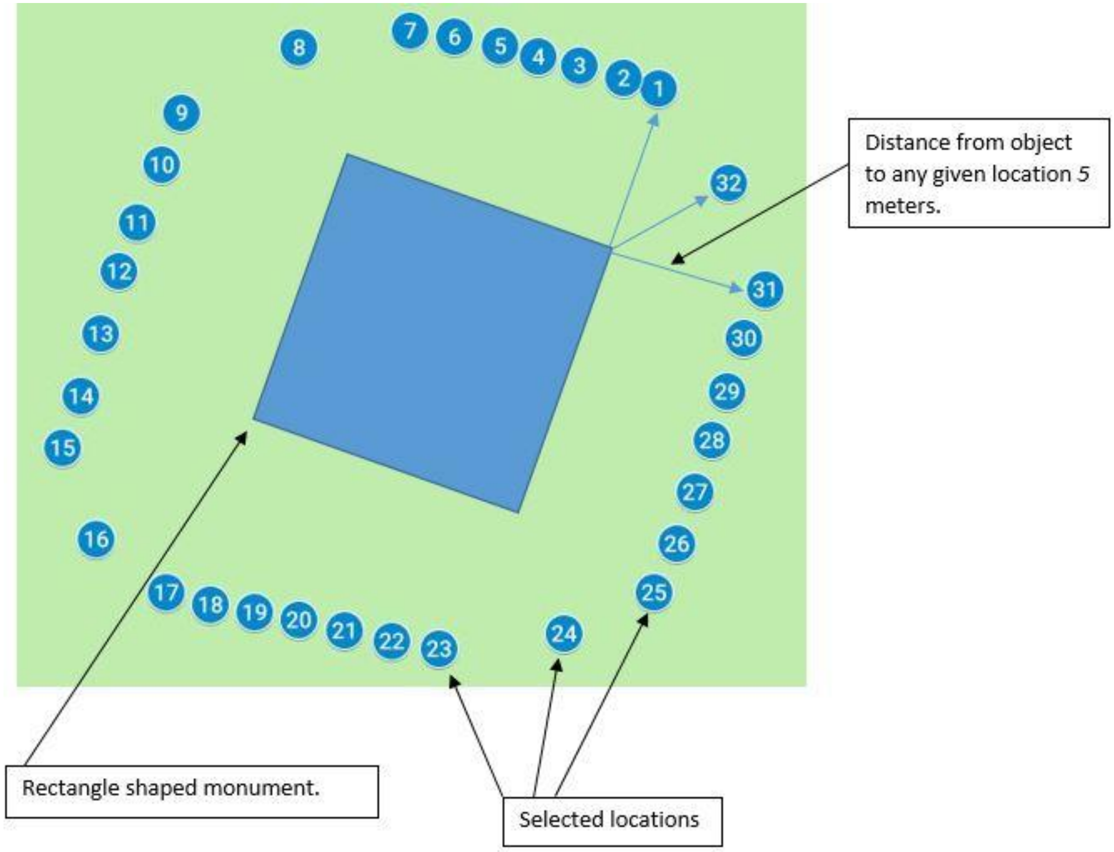


Figure 4.2: Detailed view of the 32 selected locations.

- Approximate physical scale of the object used for evaluation – length 5 m, width 5m, height 2.5m
- Distance from a selected location to the object - 5 meters
- Approximate distance between two selected locations - 1 meters
- Number of locations around the object - 32
- Device used to geo tag the obtained reference images - RTK Device (navcom SF340)
- Device used to obtain reference images - Canon EOS 1300D DSLR Camera

As mentioned in section 3.1 each of these selected locations are associated with a highly accurate GPS location value which was obtained using a highly accurate (accurate up to 10cm) RTK device. Below table (Table 4.1) depicts the GPS values associated with each of the selected locations.

Location	**Latitude**	**Longitude**
1	6.9008566	79.8608584
2	6.9008588	79.8608493
3	6.9008615	79.8608396
4	6.9008634	79.8608304
5	6.9008658	79.860822
6	6.9008679	79.860812
7	6.9008692	79.8608022
8	6.900865647	79.86077773
9	6.9008513	79.8607522
10	6.90084	79.8607476
11	6.9008273	79.8607422
12	6.9008165	79.8607383
13	6.9008029	79.8607344
14	6.9007893	79.8607299
15	6.900778155	79.860726
16	6.9007579	79.8607334
17	6.9007463	79.8607487
18	6.9007436	79.8607584
19	6.900742	79.8607681
20	6.9007403	79.8607778
21	6.9007379	79.8607879
22	6.9007358	79.8607984
23	6.9007342	79.8608087
24	6.9007374	79.8608362
25	6.9007465	79.8608559
26	6.9007569	79.8608611
27	6.9007683	79.8608649
28	6.9007795	79.8608687
29	6.9007904	79.8608721
30	6.9008018	79.8608759
31	6.9008127	79.8608806
32	6.900835773	79.86087253

Table 4.1: 10cm accurate GPS data of all the selected locations.

As mentioned in section 3.1 each of these selected locations are also associated with a bearing value. Below table (Table 4.1) depicts the bearing values associated with each of the selected locations.

Location	**Bearing (Degrees)**
1	021^o
2	021^o
3	021^o
4	021^o
5	021^o
6	021^o
7	021^o
8	336^o
9	291^o
10	291^o
11	291^o
12	291^o
13	291^o
14	291^o
15	291^o
16	246^o
17	201^o
18	201^o
19	201^o
20	201^o
21	201^o
22	201^o
23	201^o
24	156^o
25	111^o
26	111^o
27	111^o
28	111^o
29	111^o
30	111^o
31	111^o
32	066^o

Table 4.2: Bearing associated with each selected location.

As described in this section a set of 32 locations were selected as reference locations to capture reference images this number of locations are highly dependent on the scale of the object. For this object increasing the number of locations for a value more than 32 locations will result in having reference images that are similar to their nearby reference images. Hence, it will reduce the accuracy of the implemented template matching algorithm. Decreasing the number of reference locations for a value less than 32 locations will reduce the accuracy of the implemented template matching algorithm for user images taken in between two reference locations. Hence, the most optimal number of locations for this object is 32.

4.3 Deriving a function for each location

As described in section 3.2, Linear Regression was used to derive the function at each of the selected locations. When deriving the functions

1. Select a single location out of the selected set of location.
2. For that selected location obtain three set of images.
    a. First image - 5 meters away from the object.
    b. Second image - 6 meters away from the object.
    c. Third image - 7 meters away from the object.
3. Identify the scale of the object (S_1, S_2 and S3 respectively) in each of these obtained three images. This task was done using the same algorithm which was used to identify the scale of the object in the user image (explained in section 3.5). In this step original images were used while increasing the image size by 0.01% in each iteration (this will be further explained in section 3.5).
4. To obtain the Linear regression function for this selected location use the points (5, S_1), (6, S_2) and (7, S_3).
5. Do this for all the selected locations.

Below table (Table 4.1) depicts the scale of the object and the accuracy of matching in the set of three images when the image is taken 5 meters away from the object, 6 meters away from the object and 7 meters away from the object.

Location	**Matching Accuracy**	**Scale(5m)**	**Matching Accuracy**	**Scale(6m)**	**Matching Accuracy**	**Scale(7m)**
1	82	0.10476	82	0.09245	81	0.08266
2	84	0.10785	81	0.09625	79	0.08349
3	85	0.11175	84	0.09686	82	0.08336
4	83	0.10815	81	0.09385	78	0.08182
5	85	0.10325	83	0.08965	81	0.07855
6	82	0.10435	79	0.08815	79	0.07765
7	85	0.09735	79	0.08675	75	0.07565
8	85	0.10625	80	0.09443	75	0.08456
9	83	0.10225	80	0.09325	77	0.08295
10	82	0.10805	78	0.09565	74	0.08553
11	82	0.12015	76	0.10165	73	0.09696
12	79	0.12025	75	0.10605	69	0.09605
13	82	0.11985	74	0.10651	69	0.10355
14	80	0.11091	75	0.09755	69	0.08795
15	79	0.10325	70	0.09475	62	0.08416
16	85	0.10395	79	0.09105	76	0.08304
17	85	0.10186	83	0.08893	82	0.07736
18	84	0.11011	83	0.09505	80	0.08535
19	84	0.12115	81	0.10215	79	0.08941
20	86	0.12085	87	0.09945	86	0.08625
21	85	0.11895	85	0.10155	83	0.08766
22	85	0.11285	83	0.09735	81	0.08544
23	86	0.10535	83	0.09365	76	0.08568
24	82	0.10355	76	0.09225	72	0.08108
25	85	0.10485	77	0.09164	69	0.08055
26	84	0.10995	76	0.09545	71	0.08155
27	84	0.11515	79	0.09966	75	0.08559
28	85	0.11505	80	0.09664	75	0.08626
29	85	0.11595	82	0.09925	80	0.08655
30	87	0.11505	84	0.09975	79	0.08875
31	87	0.10625	84	0.09805	82	0.08766
32	87	0.11025	86	0.09854	85	0.08485

Table 4.3: Scale of the object in the images taken at all 32 locations at three different levels and their respective matching accuracy.

Below table (Table 4.4) depicts a set of images taken in order to obtain the linear regression function at location one, together with the template (reference) image at location one.

	a) Geo tagged image taken using a mobile device which is 5m away from the object near location one.
	b) Geo tagged image taken using a mobile device which is 6m away from the object near location one.
	c) Geo tagged image taken using a mobile device which is 7m away from the object near location one.
	Template image which is taken using a Canon EOS 1300D DSLR Camera, Geo tagged using a 10cm accurate RTK device.

Table 4.4: Set of user images and reference images taken at location 1.

Below table (Table 4.4) depicts the functions obtained using linear regression for each of the selected locations.

Location	**Function**
1	0.15959 - 0.01105 x
2	0.168943 - 0.01218 x
3	0.182493 - 0.014195 x
4	0.173597 - 0.013165 x
5	0.164583 - 0.01235 x
6	0.17015 - 0.01335 x
7	0.151683 - 0.01085 x
8	0.16015 - 0.010845 x
9	0.150717 - 0.00965 x
10	0.16397 - 0.01126 x
11	0.175823 - 0.011595 x
12	0.18005 - 0.0121 x
13	0.15887 - 0.00815 x
14	0.167683 - 0.01148 x
15	0.151323 - 0.009545 x
16	0.15541 - 0.010455 x
17	0.162883 - 0.01225 x
18	0.171117 - 0.01238 x
19	0.199457 - 0.01587 x
20	0.205983 - 0.0173 x
21	0.19659 - 0.015645 x
22	0.180777 - 0.013705 x
23	0.153903 - 0.009835 x
24	0.159703 - 0.011235 x
25	0.165247 - 0.01215 x
26	0.18085 - 0.0142 x
27	0.188813 - 0.01478 x
28	0.185687 - 0.014395 x
29	0.188783 - 0.0147 x
30	0.180083 - 0.01315 x
31	0.15309 - 0.009295 x
32	0.17408 - 0.0127 x

Table 4.5: Functions obtained using linear regression for the data in figure 4.1.

4.4 Filtering the reference images to find the most optimal set of locations

Since the accuracy of the A-GPS chips used in mobile devices are approximately 5 meter, to filter the reference images the following approach was used.

- Obtain the user’s approximate GPS location value given by the A-GPS chip in the user’s mobile device. This value can be extracted using the geo tag of the user image.
- Query the reference images by using that received GPS value and the geo tags of the reference images and find the closest 8 locations to the users approximate GPS location, out of the 32 selected.
- Select those eight locations as the optimal set of locations for that particular user location.
- Number of optimal locations is highly based on the scale of the site and the distance between two selected locations.
- Eight optimal locations were selected by assuming that a mobile device has a minimum of 500cm position localization accuracy. This number of optimal locations were selected in order to make this proposed methodology work even for the worst position localization accuracy of 500cm that can be obtained using a mobile device.

To obtain the distance between two GPS locations, this research have used the great circle distance finding method with Vincenty equation.

4.5 Using template matching to find the best matching location and the scale of the object in the user image

When performing the template matching to find the best matching template, the algorithm mentioned in the section 3.4 was used. All the optimal reference images were scaled down.

- As mentioned in section 3.4 the user image was scaled down 20% vertically and horizontally so that the number of pixels in the user image is reduced. Hence, the template matching algorithm with the scaled down user image will take less amount of time compared to the 100% scaled user image.
- When scaling down the template (reference) image for the object used in the evaluation process, the template image was scaled down to 6.5 % vertically and horizontally of its original scale.
- Then the scale of the scaled down template image was increased gradually. Evaluations were done under two categories. One increasing the scale by 0.01% and the other increasing the scale by 0.1% at each iteration.
- The scale of the template image was increased gradually up until it reached the level 14 % of its original scale.
- For all these iterations, a correlation score was computed for each of the different scale values.
- Highest 10 correlation scores of each optimal template images were saved and the template with the largest sum over its highest 10 correlation scores was taken as the best matching template and its geo tag was taken as the best matching location.
- The average of the 10 scale values of that template image which had the highest 10 correlation scores was taken as the scale of the object in the user image.
- Evaluations were done under two categories. One using gray scaled images and the other by using original images that were loaded using alpha channel.
- All of the above mentioned percentage values are highly depended on this specific use case. And they are subject to change with the change of the objects that are reconstructed.

- For template matching, this research has used the Opencv template matching algorithms (OpenCV 2.4.13.7 documentation).
- Out of all the other template matching algorithms TM_CCOEFF_NORMED algorithm out performs the others according to previous carried out studies. Hence, TM_CCOEFFF_NORMED was the algorithm used in this proposed approach.

Below figure (Figure 4.3) depicts a pseudo code of the algorithm which was used to find the best matching location and the scale of the object in the user image.

```
Load the user image as source image;
Extract the geo tag of the user image;
Load all the reference images at optimal locations: optimal_image_array[];
Reduce source image size to 0.2 (20%);

Best_match_rate = 0
Best_match_location = 0
Best_match_scale = 0

For T_image in optimal_image_array;
    Template = T_mage;
    Accuracy_arr=[]
    #reducing the template image size to 6.5%
    #increasing template image size by 0.01% or 0.1%
    #increase it up until it reach 14%
    For j in range(650,1401,1):
        Reduce template size to j;
        Perform template matching (source image, template image) with TM_CCOEFF_NORMED
        Accuracy=matching_accuracy
        Accuracy_arr.append(j,Accuracy)

    Sorted_accuracy_arr = Sort Accuracy_arr on Accuracy

    Match_rate = average of first 10 Accuracy in sorted_accuracy_arr
    Match_scale = average of first 10 j values in sorted_accuracy_arr

    If(best_match_rate<Match_rate):
        Best_match_rate=Match_rate
        Best_match_location = location of Image
        Best_match_scale = Match_scale
```

Figure 4.3: Pseudocode to find best matching location and the scale of the object in the user image.

4.6 Calculating the distance from the user to the best matching location.

Obtain the function at the best matching location. Use the scale of the object in the user image as the ‘y’ value in the function and solve that function to obtain the ‘x’ value which will be the distance from the user to the object.

All the reference images were taken 5m away from the object. Hence, the distance from the object to the best matching location would be 5m. Distance from the user to the best matching location can be obtained by subtracting 5m from the ‘x’ value obtained by solving the equation.

4.7 Obtaining user’s accurate location value

As mentioned in the literature review this task was done using geo tag of the best matching template image (which was measured using a 10cm accurate RTK device), bearing associated with that best matching location and the distance from the user to the best matching location. These 3 variables were used as inputs to the Great Circle Distance method with Vincenty equation in order to obtain user’s accurate GPS location.

4.8 Augmentation

To implement this proposed methodology and to perform the necessary augmentation, the following tools and software were used during this research.

- In order deploy this system to a mobile device Unity 2017.4.12f1 (64-bit) version was used to build the relevant application.
- In order to do the precise augmentation using the accurate GPS values of the user, Kudan-Unity3D-Plugin was used in this research.
- After the initial augmentation to keep track of the augmented object this research has used extended tracking technologies which are equipped in Kudan-Unity3D-Plugin.

4.9 Summary

This chapter provided a thorough explanation on how the conducted experiments were implemented. Moreover it provided information about the data set used for that conducted experiments. Following chapter will describe on how the conducted experiments were evaluated.

Chapter 5 - Results and Evaluation

5.1 Introduction

This chapter describes the results of the conducted experiments. As mentioned in chapter 1 the main problem that arises when considering augmented reality based reconstruction using mobile devices, is the accuracy of the position localization achieved by using the A-GPS chips in the mobile device. Purpose of this research paper is to bridge that gap and provide an accurate position localization method consuming minimal time which can be used with mobile devices for augmented reality based reconstruction.

In order to evaluate the proposed methodology the same data set was used to conduct three different experiments. Below are the brief explanations of the three experiments.

1. Experiment 1 - Template matching using original images with image size increasing by 0.01% in each iteration (step size 0.01%).
2. Experiment 2 - Template matching using original images with image size increasing by 0.1% in each iteration (step size 0.1%).
3. Experiment 3 - Template matching using gray scaled images with image size increasing by 0.01% in each iteration (step size 0.01%).

Sole purpose of conducting three experiments were to analyze the processing time of each of these experiments together with their accuracy levels. In order to evaluate the accuracy of the position localization obtained from the proposed methodology, position localization results obtained in each of the above three experiments were compared against position localization results obtained using a RTK devices. The error of the above results were compared against the error between the position localization results obtained using the A-GPS chip in mobile devices and the position localization results obtained using a RTK device. When obtaining position localization results of a RTK device, a near 10cm accurate RTK device was used. Compared to the position localization accuracy of mobile devices which is 500cm this accuracy of the RTK device is extremely high. Hence, position localization results of the RTK device were considered as the ground truth and it was assumed that they represented the most accurate location of the user.

When calculating the average time taken in each experiment to perform position localization, the time this proposed methodology takes starting from receiving the image to producing the accurate GPS localization value was taken and then their average was calculated for each experiment.

All of the above mentioned three different experiments were conducted using the same data set. For all those experiments below table (Table 5.1) depicts the accurate locations of the user obtained using a 10cm accurate RTK device. Where first 11 of those locations are 550cm away from the object, second 9 locations are 650 cm away from the object and third 6 locations are 750cm away from the object.

Number	**Latitude**	**Longitude**
Locations which are 550cm away from the object		
1	06.90086082089624	79.86086002130470
2	06.90086762089624	79.86083202130470
3	06.90087342089624	79.86080382130480
4	06.90084162025025	79.86074337635690
5	06.90080452025025	79.86073017635730
6	06.90075606106556	79.86072926699860
7	06.90073777910373	79.86076647869570
8	06.90073157910373	79.86079677869570
9	06.90074487974971	79.86086012364220
10	06.90077787974971	79.86087292364250
11	06.90081107974971	79.86088482364280
Locations which are 650cm away from the object		
1	06.90087846268870	79.86082686391440
2	06.90087803794288	79.86077220960620
3	06.90083216075064	79.86072952907110
4	06.90079416075065	79.86071722907210
5	06.90073363731118	79.86074383608710
6	06.90072763731117	79.86077293608720
7	06.90072153731117	79.86080383608720
8	06.90078553924901	79.86088477092770
9	06.90084128980309	79.86088492900650
Locations which are 750cm away from the object		
1	06.90088900448114	79.86082010652410
2	06.90085940125088	79.86073108178390
3	06.90072249551860	79.86075029347870
4	06.90071679551859	79.86077979347880
5	06.90071674842826	79.86084540065320
6	06.90079369874817	79.86089701821320

Table 5.1: Position localization values of the user measured using 10cm accurate RTK Device.

For all those experiments below table (Table 5.2) depicts the position localization values obtained using the A-GPS in mobile devices for all of the 26 images.

Number	**Latitude**	**Longitude**
Locations which are 550cm away from the object		
1	06.90087069444444	79.86083244444440
2	06.90088433333333	79.86078952777780
3	06.90087997222222	79.86079119444440
4	06.90082675000000	79.86074205555550
5	06.90080394444445	79.86074747222220
6	06.90078213888889	79.86075619444440
7	06.90075958333333	79.86081450000000
8	06.90071116666667	79.86081369444440
9	06.90073380555556	79.86092377777780
10	06.90078386111111	79.86089244444440
11	06.90081227777778	79.86090952777780
Locations which are 650cm away from the object		
1	06.90086955555555	79.86080544444440
2	06.90090633333333	79.86074650000000
3	06.90081633333333	79.86075188888890
4	06.90080302777778	79.86073936111110
5	06.90075269444444	79.86075438888890
6	06.90075122222222	79.86079663888890
7	06.90070655555556	79.86084369444440
8	06.90079163888889	79.86087252777780
9	06.90085583333333	79.86088100000000
Locations which are 750cm away from the object		
1	06.90087255555556	79.86079863888890
2	06.90085177777778	79.86072941666670
3	06.90075800000000	79.86078483333330
4	06.90071325000000	79.86081930555560
5	06.90070502777778	79.86088963888890
6	06.90078516666667	79.86089191666670

Table 5.2: Position localization values of the user measured using A-GPS chip in mobile device.

Below table (Table 5.3) depicts the distance between the position localization values obtained using a 10cm accurate RTK device and the position localization values obtained using A-GPS chips in mobile devices. This distance is calculated using the great circle distance method with Vincenty equation.

Number	**Distance (m)**
550cm away from the object	
1	3.237416141
2	5.046866992
3	1.572354081
4	1.650971252
5	1.912550736
6	4.144085772
7	5.829297056
8	2.931022838
9	7.140760338
10	2.256515747
11	2.733444274
650cm away from the object	
1	2.563988258
2	4.226718958
3	3.028251879
4	2.635216011
5	2.408708599
6	3.696642603
7	4.706318782
8	1.511904142
9	1.665957562
750cm away from the object	
1	2.989657447
2	0.862932501
3	5.476166938
4	4.384332771
5	5.057991172
6	1.099179694

Table 5.3: Distance between position localization values obtained using RTK device and mobile device.

Following are google map images of the above recorded position localization points obtained using 10cm accurate RTK device and position localization points obtained using A-GPS chips in mobile devices. First figure (Figure 5.1) depicts the mapping for the first 11 locations taken 550cm away from the object. Second figure (Figure 5.2) depicts the mapping for the following 9 locations taken 650cm away from the object. Third figure (Figure 5.3) depicts the mapping for the following 6 locations taken 750cm away from the object.

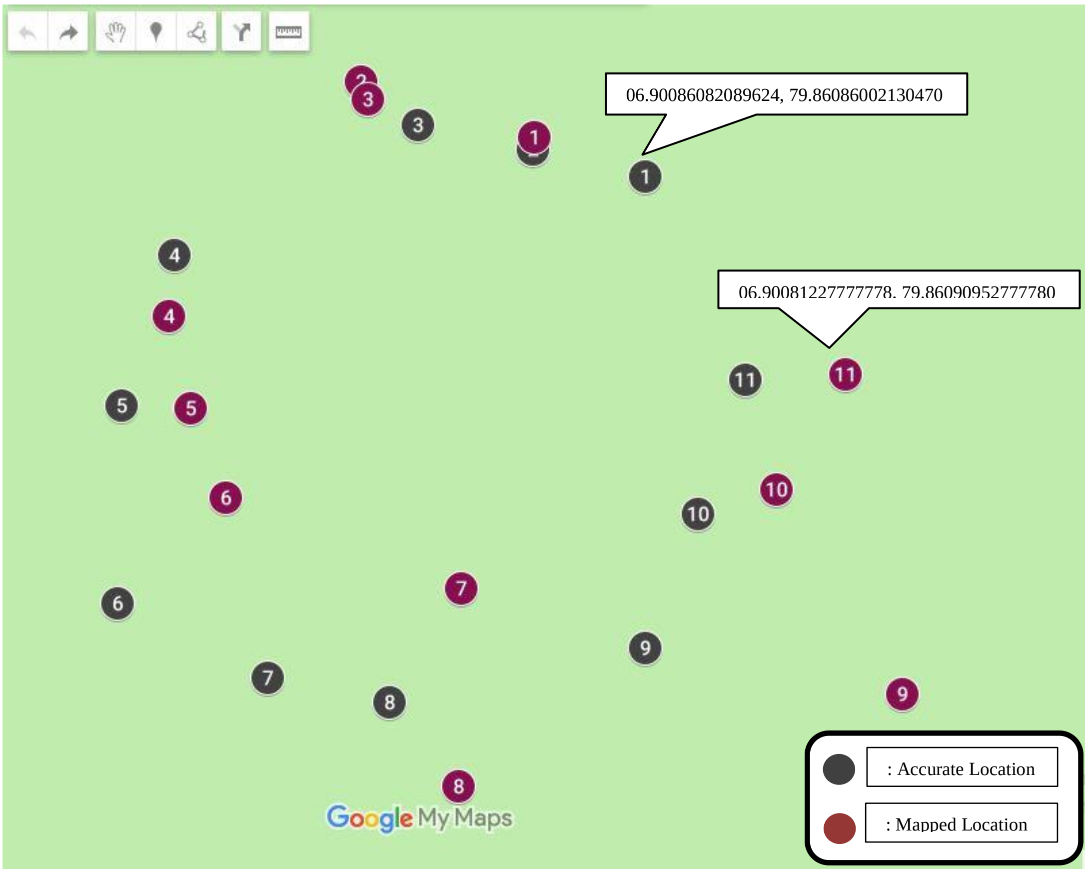


Figure 5.1: Mapping of position localization data of RTK device against data obtained using mobile device for locations that are 550cm away from the object.

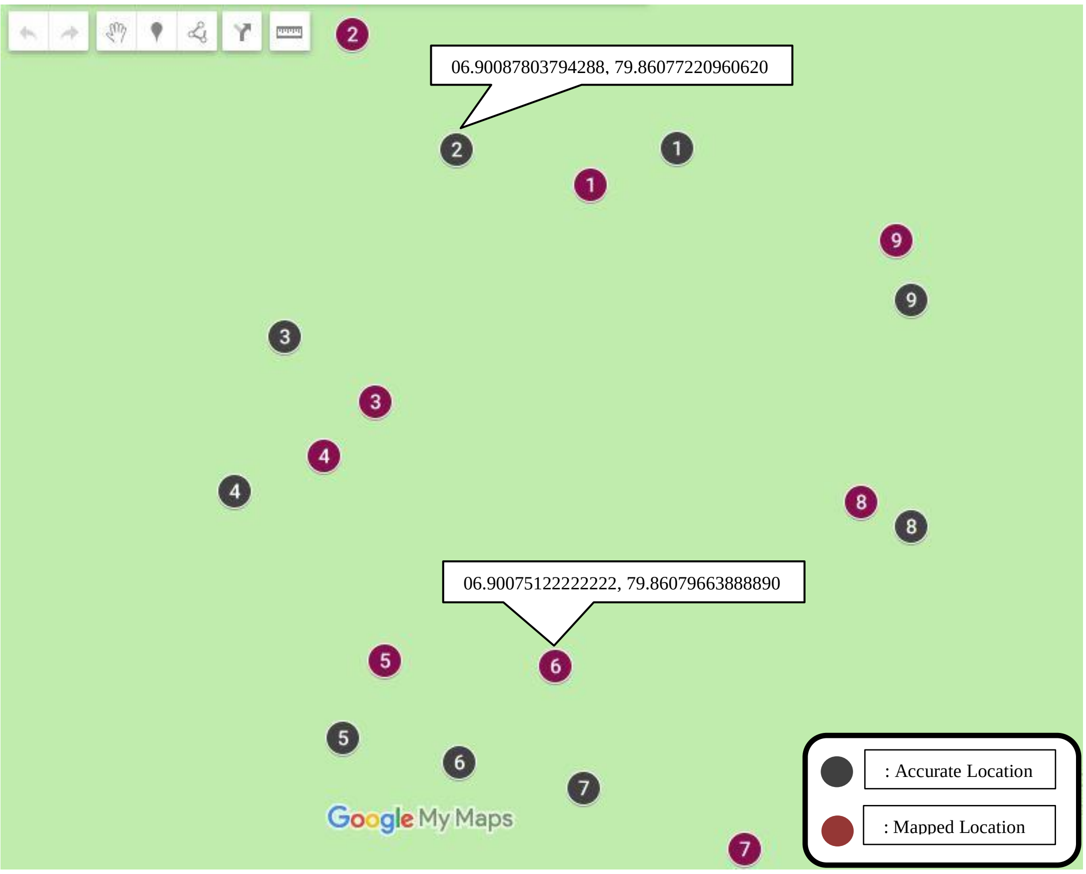


Figure 5.2: Mapping of position localization data of RTK device against data obtained using mobile device for locations that are 650cm away from the object.

.

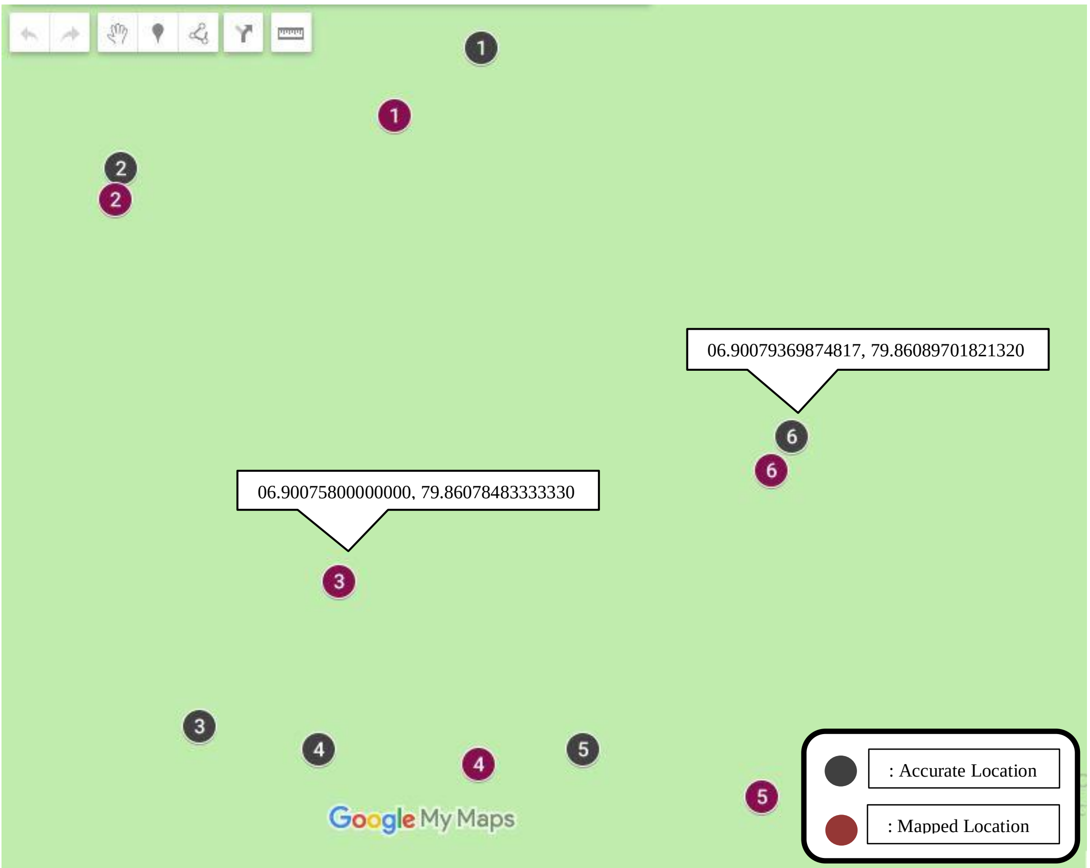


Figure 5.3: Mapping of position localization data of RTK device against data obtained using mobile device for locations that are 750cm away from the object.

Below table (Table 5.4) depicts a summary of results obtained by comparing position localization values obtained by only using the A-GPS chips in mobile devices against the position localization values obtained using a RTK device for the evaluated 26 locations. In this table 'x' represents the distance between actual location obtained through a RTK device and the location given by the A-GPS chips in mobile devices.

	x between 0cm and 10cm	**x between 10cm and 40cm**	**x between 40cm and 150cm**	**x between 150cm and 500cm**	**x between 500cm and 800cm**
Percentage of locations classified into each class	00.0 %	00.0 %	7.69 %	73.08 %	19.23 %

Table 5.4: Summary of results obtained by comparing data obtained using RTK device against data obtained using mobile devices.

Below figure (Figure 5.4) depicts a graphically represents a summary of results obtained by comparing data obtained using RTK device against data obtained using mobile devices.

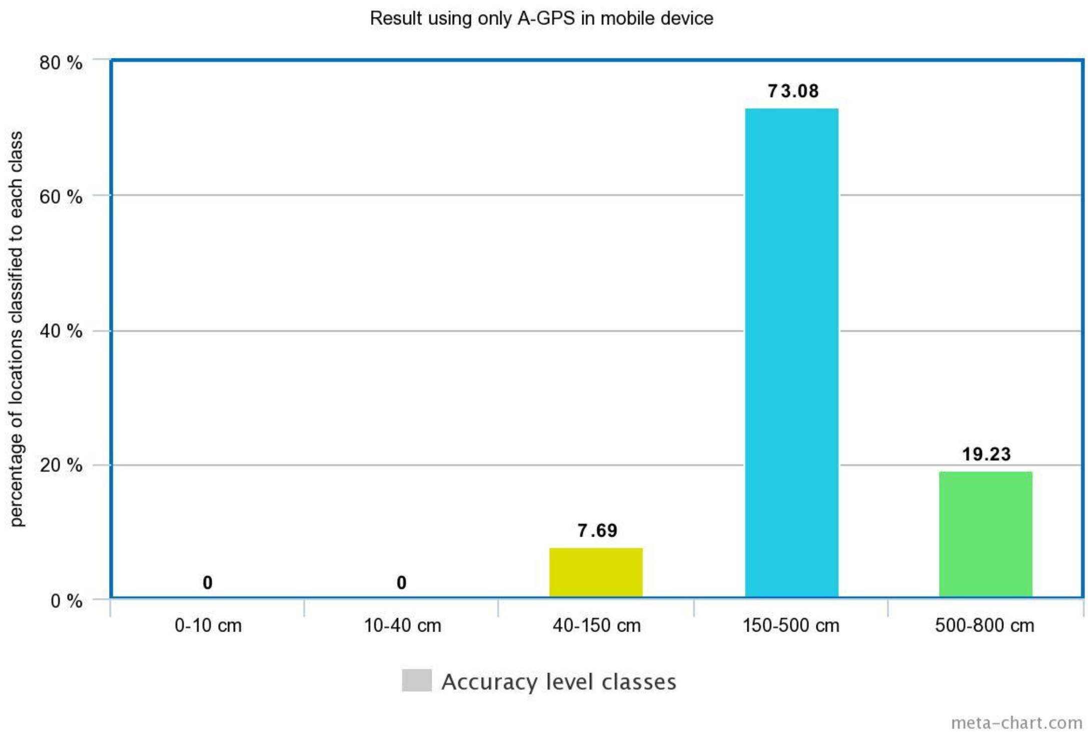
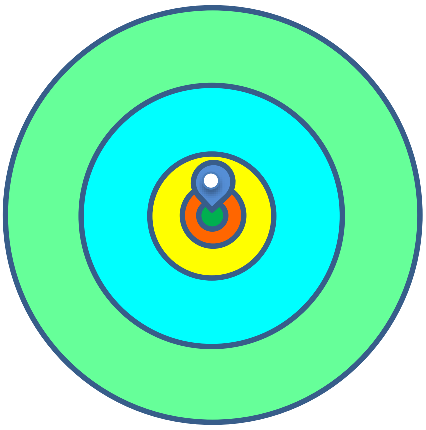


Figure 5.4: Graphical summary of the results obtained by comparing data obtained using RTK device against data obtained using mobile devices.

Below figure (Figure 5.5) represents the probabilities of an evaluated location being mapped to its correct location with a particular range of accuracy using only the A-GPS chips in mobile device.



Color	Distance Range	Probability
	0cm – 10 cm	0.0
	10cm – 40 cm	0.0
	40cm – 150cm	0.0769
	150cm – 500cm	0.7308
	500cm – 800cm	0.1923

Figure 5.5: Probabilities of matching accuracy levels using mobile devices.

5.2 Template matching using original images with 0.01% step size

Below table (Table 5.5) depicts position localization values obtained for the 26 user locations when using the proposed methodology with original images while increasing the template image size by 0.01% after each iteration.

Number	Latitude	Longitude
Locations which are 550cm away from the object		
1	06.90086186370590	79.86086042186240
2	06.90087020473238	79.86083301379200
3	06.90087385504557	79.86080398806750
4	06.90084167493010	79.86074323381840
5	06.90080535323779	79.86072800493850
6	06.90075561870113	79.86072827278540
7	06.90073374028803	79.86077528032440
8	06.90073117903483	79.86079662502380
9	06.90074499390067	79.86085982607530
10	06.90077748253794	79.86087395908770
11	06.90081062043184	79.86088602098530
Locations which are 650cm away from the object		
1	06.90088027955383	79.86082756179750
2	06.90087949088424	79.86077156229400
3	06.90083272403036	79.86072806072220
4	06.90079489833142	79.86071530635810
5	06.90072643888585	79.86075180818020
6	06.90072650668267	79.86078352367020
7	06.90071712888609	79.86079122816870
8	06.90078534459852	79.86088527833940
9	06.90084088436875	79.86088401779340
Locations which are 750cm away from the object		
1	06.90088766075013	79.86081959037840
2	06.90084849552132	79.86072545400620
3	06.90072296124916	79.86075047237220
4	06.90071744167050	79.86078004167460
5	06.90071683629165	79.86084536150840
6	06.90080521669200	79.86090010736770

Table 5.5: Position localization data obtained through the proposed methodology in Experiment 1.

Below table (Table 5.6) depicts the distance between the position localization values obtained using a 10cm accurate RTK device and the position localization value obtain through the proposed methodology for all the 26 user locations in Experiment 1.

Number	Distance (m)
550cm away from the object	
1	0.123529237
2	0.306076289
3	0.051406776
4	0.0168253
5	0.257052768
6	0.120276282
7	1.070376589
8	0.047371378
9	0.035226001
10	0.122576116
11	0.141741524
650cm away from the object	
1	0.21522237
2	0.175887188
3	0.17382327
4	0.227611073
5	1.187430955
6	1.176771758
7	1.476220081
8	0.060067456
9	0.110235216
750cm away from the object	
1	0.159175807
2	1.356997346
3	0.055169552
4	0.076541918
5	0.010630597
6	1.318732217

Table 5.6: Distance between position localization values obtained using RTK device and Experiment 1.

Following are google map images of the above recorded position localization points obtained using 10cm accurate RTK device and position localization points obtained through the proposed methodology using original images with 0.01% step size. First figure (Figure 5.6) depicts the mapping for the first 11 locations taken 550cm away from the object. Second figure (Figure 5.7) depicts the mapping for the following 9 locations taken 650cm away from the object. Third figure (Figure 5.8) depicts the mapping for the following 6 locations taken 750cm away from the object.

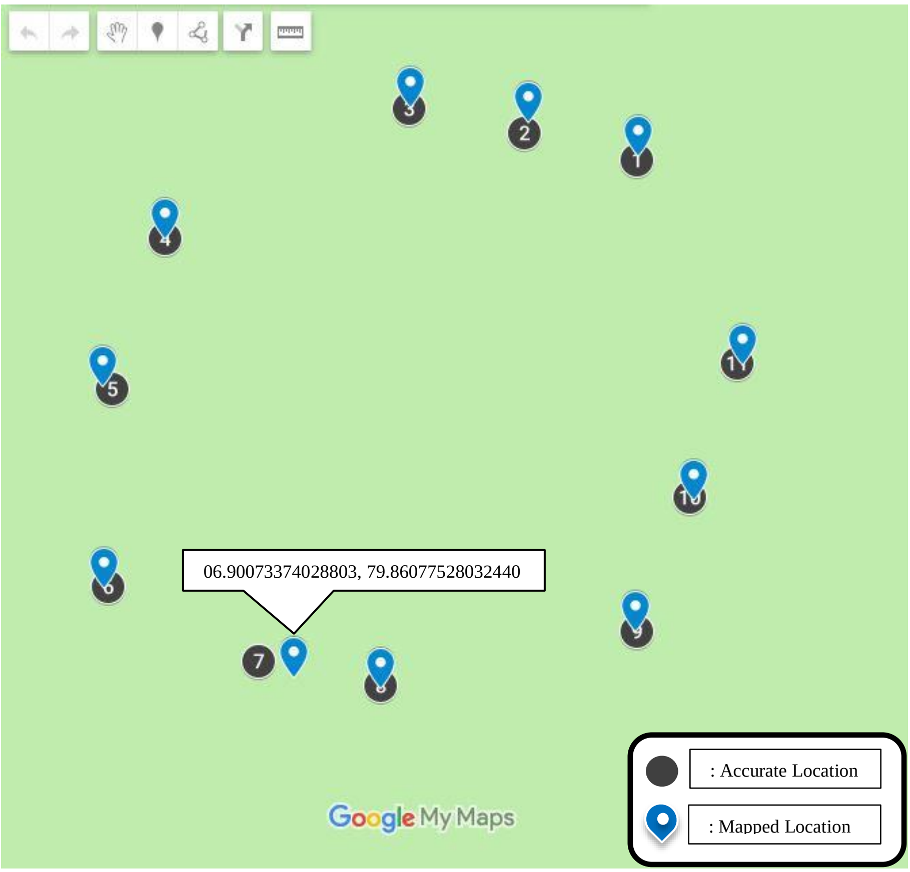


Figure 5.6: Mapping of position localization data of RTK device against data obtained in Experiment 1 for locations that are 550cm away from the object.

Figure 5.7: Mapping of position localization data of RTK device against data obtained in Experiment 1 for locations that are 650cm away from the object.

Figure 5.8: Mapping of position localization data of RTK device against data obtained in Experiment 1 for locations that are 750cm away from the object.

Below table (Table 5.7) depicts a summary of results obtained by comparing position localization values obtained in Experiment 1 against the position localization values obtained using a RTK device for the evaluated 26 locations. In this table 'x' represents the distance between actual location obtained through a RTK device and the location given in Experiment 1.

	x between 0cm and 10cm	**x between 10cm and 40cm**	**x between 40cm and 150cm**	**x between 150cm and 500cm**	**x between 500cm and 800cm**
Percentage of locations classified into each class	30.77 %	46.15 %	23.08 %	00.0 %	0.00 %
Average processing time for a single location			828.75 Seconds		

Table 5.7: Summary of results obtained by comparing data obtained using RTK device against data obtained in Experiment 1.

Below figure (Figure 5.9) depicts a graphically represents a summary of results obtained by comparing data obtained using RTK device against data obtained using the proposed methodology in Experiment 1.

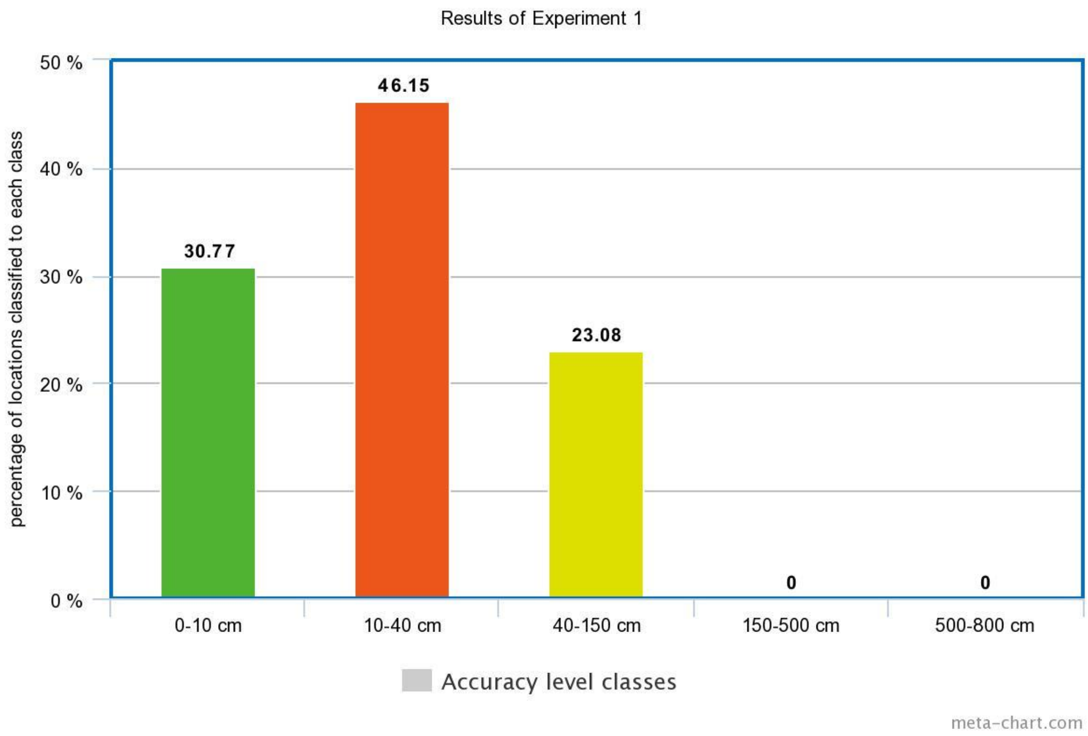
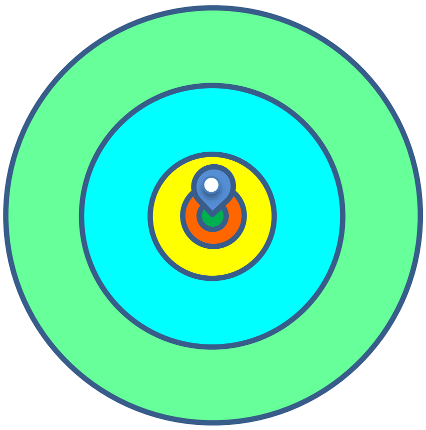


Figure 5.9: Graphical representation of the summary of results obtained by comparing data obtained using RTK device against data obtained in Experiment 1.

Below figure (Figure 5.10) represents the probabilities of an evaluated location being mapped to its correct location with a particular range of accuracy using the proposed methodology in Experiment 1.



Color	Distance Range	Probability
	0cm – 10 cm	0.3077
	10cm – 40 cm	0.4615
	40cm – 150cm	0.2308
	150cm – 500cm	0.0
	500cm – 800cm	0.0

Figure 5.10: Probabilities of matching accuracy levels using proposed method in Experiment 1.

5.3 Template matching using original images with 0.1% step size

Below table (Table 5.8) depicts position localization values obtained for the 26 user locations when using the proposed methodology with original images while increasing the template image size by 0.1% after each iteration.

Number	**Latitude**	**Longitude**
Locations which are 550cm away from the object		
1	06.90083434818781	79.86086932773690
2	06.90087062794415	79.86083317635360
3	06.90087303809791	79.86080367426660
4	06.90084148786746	79.86074372145030
5	06.90080513455371	79.86072857500070
6	06.90075626949572	79.86072973544490
7	06.90073396475187	79.86077536654410
8	06.90073102504372	79.86079656587370
9	06.90074516726079	79.86085937416300
10	06.90078956606713	79.86087427388290
11	06.90081091676595	79.86088524850630
Locations which are 650cm away from the object		
1	06.90088106563167	79.86082786374090
2	06.90087648217213	79.86077290273080
3	06.90083209241927	79.86072970719630
4	06.90080871302037	79.86071924670990
5	06.90072742762569	79.86075218796870
6	06.90072610199462	79.86078336822400
7	06.90071808979061	79.86079159726530
8	06.90081021950922	79.86088706610400
9	06.90084071929905	79.86088364679950
Locations which are 750cm away from the object		
1	06.90088743942973	79.86081950536610
2	06.90084810700665	79.86072646678040
3	06.90072265439886	79.86075035450680
4	06.90071800823377	79.86078025929930
5	06.90071752743459	79.86084505359160
6	06.90080568734032	79.86089888048920

Table 5.8: Position localization data obtained through the proposed methodology in Experiment 2.

Below table (Table 5.9) depicts the distance between the position localization value obtained using a 10cm accurate RTK device and the position localization value obtained through the proposed methodology for all the 26 user locations in Experiment 2.

Number	Distance (m)
550cm away from the object	
1	3.103037631
2	0.356209147
3	0.045326405
4	0.040852186
5	0.189568748
6	0.056670935
7	1.069006939
8	0.065632888
9	0.088723436
10	1.300979942
11	0.050295379
650cm away from the object	
1	0.308339648
2	0.188335298
3	0.021086473
4	1.624714394
5	1.150468437
6	1.165366044
7	1.405307398
8	2.741158983
9	0.155116696
750cm away from the object	
1	0.185392998
2	1.349156655
3	0.018812702
4	0.143655918
5	0.094303345
6	1.341701682

Table 5.9: Distance between position localization values obtained using RTK device and Experiment 2.

Following are google map images of the above recorded position localization points obtained using 10cm accurate RTK device and position localization points obtain through the proposed methodology using original images with 0.1% step size. First figure (Figure 5.11) depicts the mapping for the first 11 locations taken 550cm away from the object. Second figure (Figure 5.12) depicts the mapping for the following 9 locations taken 650cm away from the object. Third figure (Figure 5.13) depicts the mapping for the following 6 locations taken 750cm away from the object.

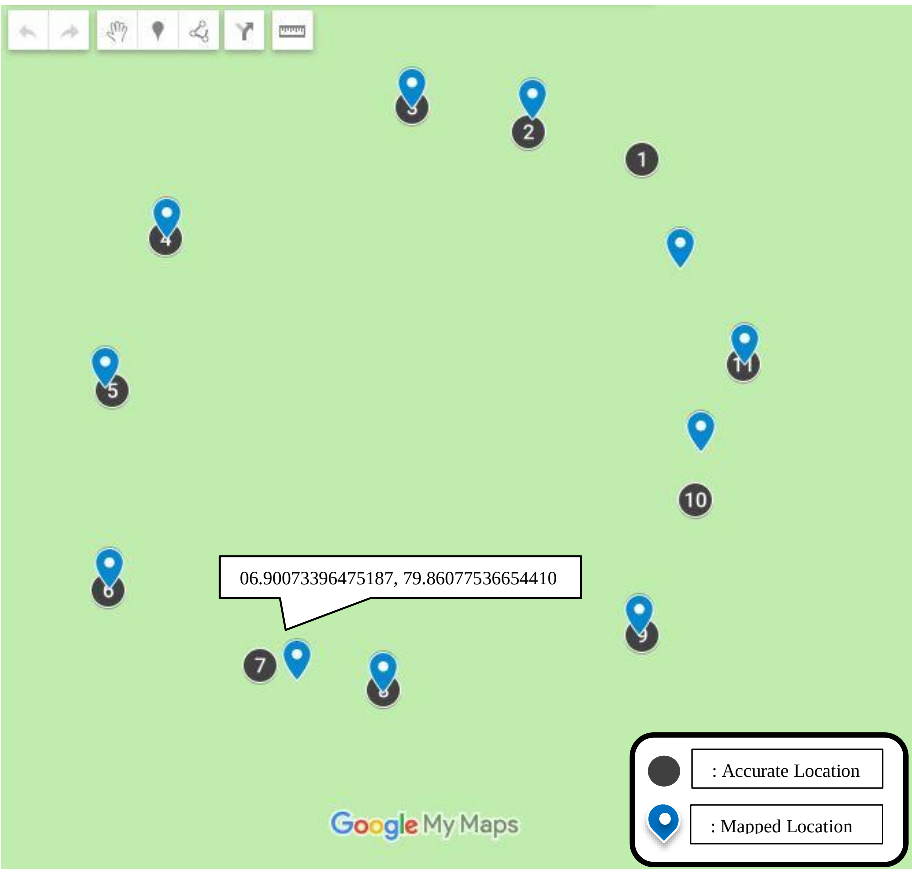


Figure 5.11: Mapping of position localization data of RTK device against data obtained in Experiment 2 for locations that are 550cm away from the object.

Figure 5.12: Mapping of position localization data of RTK device against data obtained in Experiment 2 for locations that are 650cm away from the object.

Figure 5.13: Mapping of position localization data of RTK device against data obtained in Experiment 2 for locations that are 750cm away from the object.

Below table (Table 5.10) depicts a summary of results obtained by comparing position localization values obtained in Experiment 2 against the position localization values obtained using a RTK device for the evaluated 26 locations. In this table ‘x’ represents the distance between actual location obtained through a RTK device and the location given in Experiment 2.

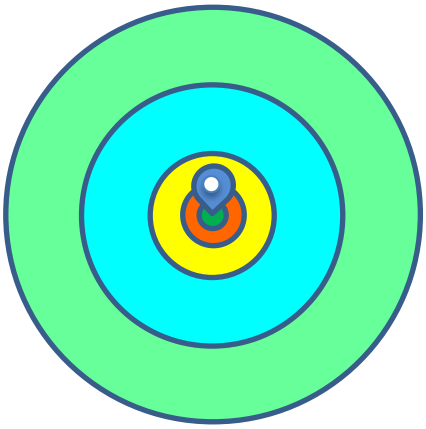
	x between 0cm and 10cm	x between 10cm and 40cm	x between 40cm and 150cm	x between 150cm and 500cm	x between 500cm and 800cm
Percentage of locations classified into each class	34.62 %	26.92 %	26.92 %	11.54 %	00.0 %
Average processing time for a single location			101.55 Seconds		

Table 5.10: Summary of results obtained by comparing data obtained using RTK device against data obtained in Experiment 2.

Below figure (Figure 5.14) depicts a graphically represents a summary of results obtained by comparing data obtained using RTK device against data obtained using the proposed methodology in Experiment 2.

Figure 5.14: Graphical representation of the summary of results obtained by comparing data obtained using RTK device against data obtained in Experiment 2.

Below figure (Figure 5.15) represents the probabilities of an evaluated location being mapped to its correct location with a particular range of accuracy using the proposed methodology in Experiment 2.



Color	Distance Range	Probability
	0cm – 10 cm	0.3462
	10cm – 40 cm	0.2692
	40cm – 150cm	0.2692
	150cm – 500cm	0.1154
	500cm – 800cm	0.0

Figure 5.15: Probabilities of matching accuracy levels using proposed method in Experiment 2.

5.4 Template matching using gray scaled images with 0.01% step size

Below table (Table 5.11) depicts position localization values obtained for the 26 user locations when using the proposed methodology with gray scaled images while increasing the template image size by 0.01% after each iteration.

Number	Latitude	Longitude
Locations which are 550cm away from the object		
1	06.90083450746560	79.86086968571350
2	06.90087053175966	79.86083313940770
3	06.90087385504557	79.86080398806750
4	06.90084164615123	79.86074330883870
5	06.90080539299854	79.86072790129080
6	06.90075561870113	79.86072827278540
7	06.90073379396416	79.86077530094220
8	06.90073117903483	79.86079662502380
9	06.90074499390067	79.86085982607530
10	06.90077748253794	79.86087395908770
11	06.90081062043184	79.86088602098530
Locations which are 650cm away from the object		
1	06.90088034790842	79.86082758805340
2	06.90087933854438	79.86077163016420
3	06.90083272403036	79.86072806072220
4	06.90079489833142	79.86071530635810
5	06.90072643888585	79.86075180818020
6	06.90072650668267	79.86078352367020
7	06.90071707344929	79.86079120687460
8	06.90078532034988	79.86088534155040
9	06.90084088147279	79.86088401128480
Locations which are 750cm away from the object		
1	06.90088766075013	79.86081959037840
2	06.90084849552132	79.86072545400620
3	06.90072296124916	79.86075047237220
4	06.90071744167050	79.86078004167460
5	06.90071629220039	79.86084560391100
6	06.90080521669200	79.86090010736770

Table 5.11: Position localization data obtained through the proposed methodology in Experiment 3.

Below table (Table 5.12) depicts the distance between the position localization values obtained using a 10cm accurate RTK device and the position localization value obtained through the proposed methodology for all the 26 user locations in Experiment 3.

Number	**Distance (m)**
550cm away from the object	
1	3.099832586
2	0.344815315
3	0.051406776
4	0.007969878
5	0.269322589
6	0.120276282
7	1.069989112
8	0.047371378
9	0.035226001
10	0.122576116
11	0.141741524
650cm away from the object	
1	0.223319524
2	0.157445543
3	0.17382327
4	0.227611073
5	1.187430955
6	1.176771758
7	1.48047459
8	0.067550377
9	0.11102261
750cm away from the object	
1	0.159175807
2	1.356997346
3	0.055169552
4	0.076541918
5	0.055229094
6	1.318732217

Table 5.12: Distance between position localization values obtained using RTK device and Experiment 3.

Following are google map images of the above recorded position localization points obtained using 10cm accurate RTK device and position localization points obtain through the proposed methodology using gray scaled images with 0.01% step size. First figure (Figure 5.16) depicts the mapping for the first 11 locations taken 550cm away from the object. Second figure (Figure 5.17) depicts the mapping for the following 9 locations taken 650cm away from the object. Third figure (Figure 5.18) depicts the mapping for the following 6 locations taken 750cm away from the object.

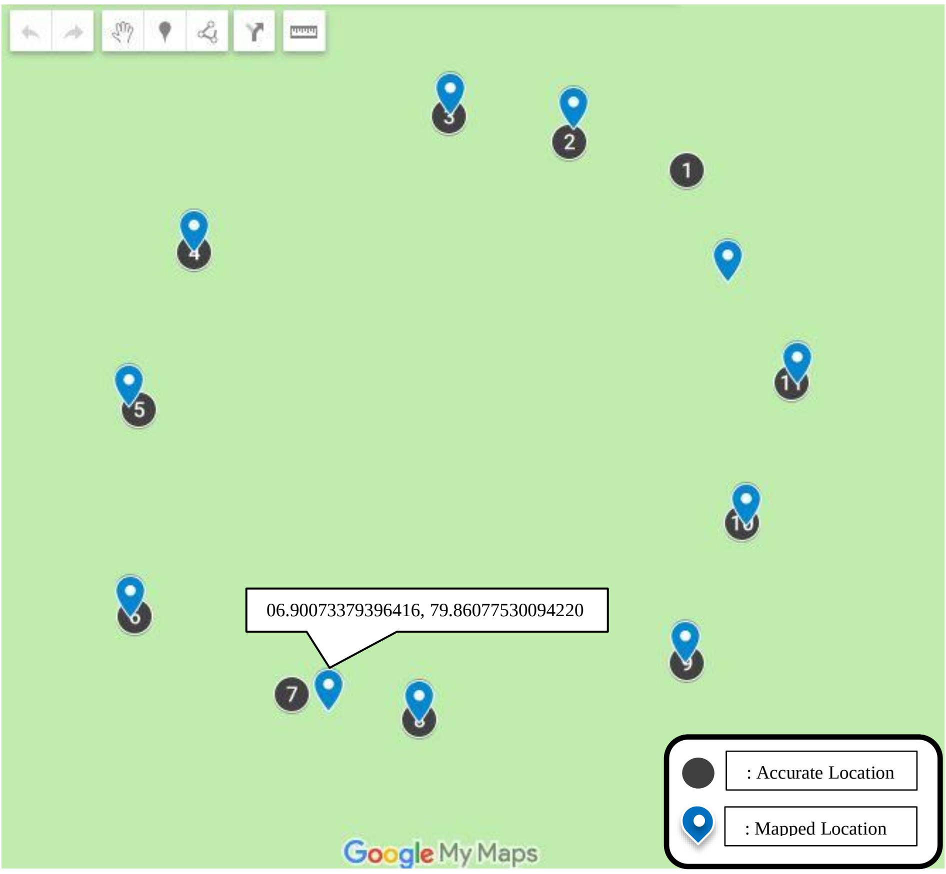


Figure 5.16: Mapping of position localization data of RTK device against data obtained in Experiment 3 for locations that are 550cm away from the object.

Figure 5.17: Mapping of position localization data of RTK device against data obtained in Experiment 3 for locations that are 650cm away from the object.

Figure 5.18: Mapping of position localization data of RTK device against data obtained in Experiment 3 for locations that are 750cm away from the object.

Below table (Table 5.13) depicts a summary of results obtained by comparing position localization values obtained in Experiment 3 against the position localization values obtained using a RTK device for the evaluated 26 locations. In this table 'x' represents the distance between actual location obtained through a RTK device and the location given in Experiment 3.

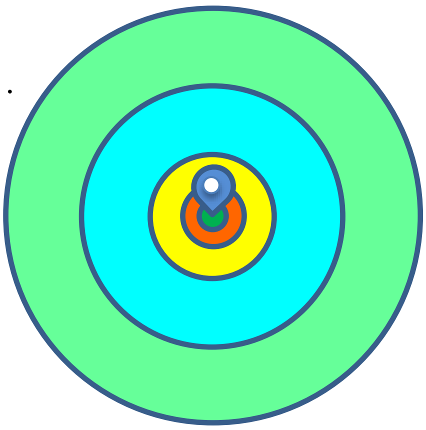
	x between 0cm and 10cm	**x between 10cm and 40cm**	**x between 40cm and 150cm**	**x between 150cm and 500cm**	**x between 500cm and 800cm**
Percentage of locations classified into each class	30.77 %	42.31 %	23.08 %	3.85 %	00.0 %
Average processing time for a single location			601.84 Seconds		

Table 5.13: Summary of results obtained by comparing data obtained using RTK device against data obtained in Experiment 3.

Below figure (Figure 5.19) depicts a graphically represents a summary of results obtained by comparing data obtained using RTK device against data obtained using the proposed methodology in Experiment 3.

Figure 5.19: Graphical representation of the summary of results obtained by comparing data obtained using RTK device against data obtained in Experiment 3.

Below figure (Figure 5.20) represents the probabilities of an evaluated location being mapped to its correct location with a particular range of accuracy using the proposed methodology in Experiment 3.



Color	Distance Range	Probability
	0cm – 10 cm	0.3077
	10cm – 40 cm	0.4231
	40cm – 150cm	0.2308
	150cm – 500cm	0.0385
	500cm – 800cm	0.0

Figure 5.20: Probabilities of matching accuracy levels using proposed method in Experiment 3.

5.5 Devices used during the experiments

- Obtaining geo tags for the reference images
    - RTK device - navcom SF340
- Obtaining geo tagged user images
    - Device -Nokia 6.1 (Nokia 6 2018)
    - Rear camera – 16MP
- Position localization with the proposed method
    - System Model: HP Pavilion 15 Notebook PC
    - Processor: Intel(R) Core(TM) i7-5500U CPU @ 2.40GHz (4 CPUs), ~2.4GHz
    - Memory: 8192MB RAM
    - Available OS Memory: 8114MB RAM

5.6 Augmentation results obtained using the proposed methodology.

Below figure (Figure 5.21) depicts a screenshot of the performed augmented reality based reconstruction using the proposed methodology for position localization using a mobile device. When reconstructing the object this study focused on precise and accurate placement of the object. Hence, physical attributes of the actual object and the reconstructed object are not the same.

Figure 5.21: Screenshot of the performed AR reconstruction.

5.7 Summary

Below table (Table 5.14) depicts a summary of all the position localization results obtained in all the three experiments together with the position localization results obtained for by only using A-GPS chips in mobile devices for all the evaluated 26 locations.

Position localization method	**x between 0cm and 10cm**	**x between 10cm and 40cm**	**x between 40cm and 150cm**	**x between 150cm and 500cm**	**x between 500cm and 800cm**	**Average processing time (Seconds)**
A-GPS in mobile devices	00.0 %	00.0 %	7.69 %	73.08 %	19.23 %	-
Experiment 1	30.77 %	46.15 %	23.08 %	00.0 %	0.00 %	828.75
Experiment 2	34.62 %	26.92 %	26.92 %	11.54 %	00.0 %	101.55
Experiment 3	30.77 %	42.31 %	23.08 %	3.85 %	00.0 %	601.84
Average of Experiment 1,2 & 3	32.05%	38.46%	24.36%	5.13%	00.0%	-

Table 5.14: Summary of all the experiments.

Below figure (Figure 5.22) graphically represents a summarization of all the position localization results obtained using the proposed methodology in all three experiments. This chart represents the average accuracy levels obtained for all the accuracy classes in all three experiments.

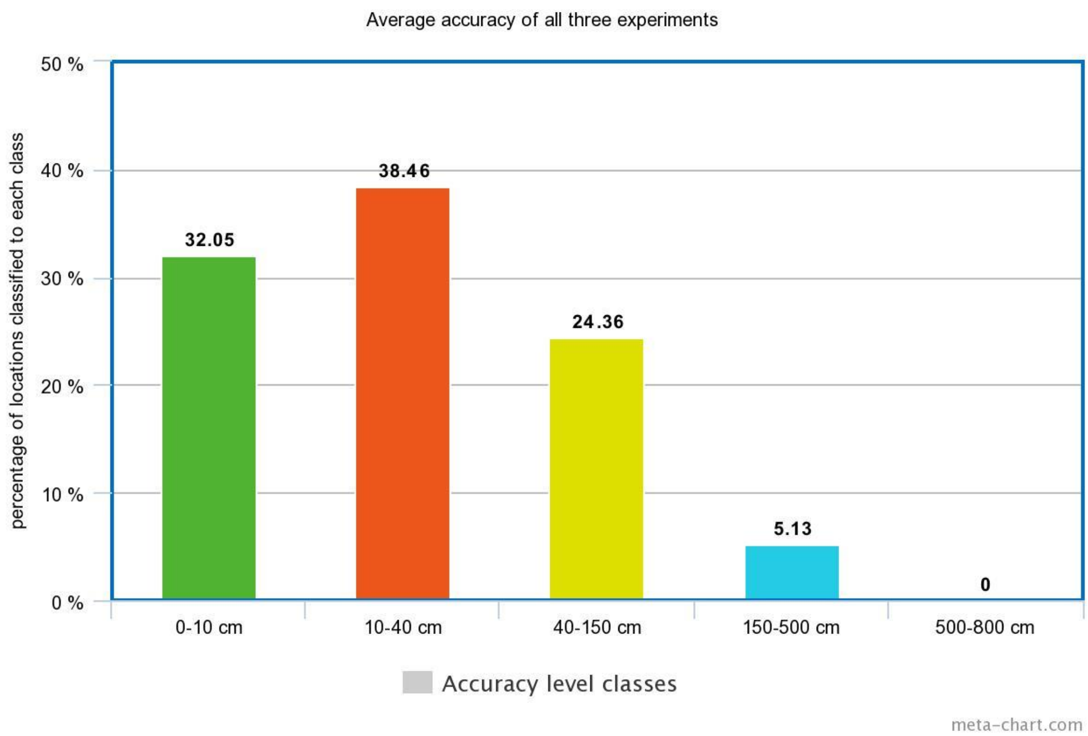
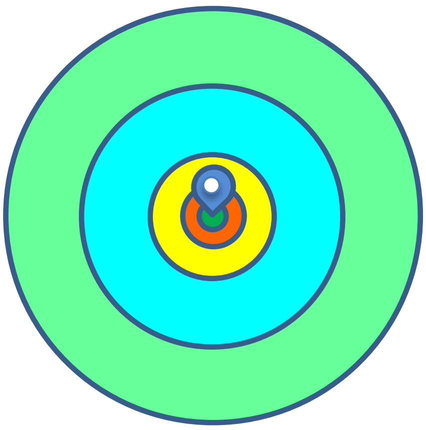


Figure 5.22: Graphical representation of average accuracy obtained using the proposed method.

Below figure (Figure 5.23) graphically represents the average probability of an evaluated location being mapped to its correct location with a particular range of accuracy using the proposed methodology



Color	Distance Range	Probability
	0cm – 10 cm	0.3205
	10cm – 40 cm	0.3846
	40cm – 150cm	0.2436
	150cm – 500cm	0.0513
	500cm – 800cm	0.0

Figure 5.23: Average probabilities of matching accuracy levels using the proposed methodology of this thesis.

Below table (Table 5.15) depicts distance based average accuracy for each accuracy level for all the evaluated 26 locations, categorized based on the distance from the location to the object.

Distance from location to the object	Number of locations	x between 0cm and 10cm	x between 10cm and 40cm	x between 40cm and 150cm	x between 150cm and 500cm
550cm	11	42.42 %	39.39 %	12.12 %	6.06 %
650cm	9	11.11%	48.15%	33.33%	7.41%
750cm	6	44.44%	22.22%	33.33%	00.0 %

Table 5.15: Distance based average accuracy for each accuracy level.

According to the above table (Table 5.15) relatively large amount of locations have been classified with an accuracy level of 40cm when the locations are 550cm away from the object. When the distance from the location to the object increases the accuracy level of the position localization achieved using this methodology decreases. This is largely due to the fact that the implemented template matching algorithm does not have the capability to perform with a considerable accuracy level when the images are taken substantial distance (more than 800cm for this particular object) away from the object. When the image is taken substantial distance away from the object then the user image will contain more background information rather than the object itself. Hence, it is necessary that at least 40% of the user image should contain the object in order to this proposed methodology to work with considerable accuracy.

This chapter presents the results and the evaluation of each experiment together with a screenshot of the performed augmentation using this proposed methodology. All the evaluations were described graphically and numerically. Furthermore each and every result of all the carried out experiments have been compared against existing position localization technology for mobile devices which uses A-GPS chips.

Chapter 6 - Conclusions

6.1 Introduction

This thesis presents how well image processing techniques and mobile sensor information can be combined to perform accurate position localization for augmented reality based reconstruction of ancient monuments. Experimental and exploratory approaches have been taken to compare the results from different techniques and methods.

Augmented reality has become one of the rapid growing technologies in the industry of information technology. This technology is used to overlay digital information on top of a real world environment. Augmented reality is often combined with many other industries to build various different application. Some of those industries include navigation, interior designing, apparel industry, gaming and archeology. This rapid growth of augmented reality is largely due to the rapid growth in the industry of smartphones. Since mobile devices possess the hardware requirements which can handle augmented reality based applications, development of these two industries have gone hand in hand. Even though different capabilities of mobile devices have hugely influenced augmented reality industry to grow in heaps and bounds, low accuracy level of position localization obtained using mobile devices has become a constant issue when considering augmented reality applications using mobile devices.

One such application is archeological reconstruction using augmented reality for mobile devices. Often the accuracy level of position localization obtained using mobile devices have not been accurate enough to perform augmented reality based reconstruction. And some of the existing solutions to this issue often requires additional hardware equipment. This paper presents a novel approach which can be used to bridge this gap and perform highly accurate position localization using mobile devices.

Three different experiments were conducted in order to evaluate the proposed methodology against existing position localization methodology of mobile devices. Results obtained from those experiments justify that the proposed methodology out performs the existing position localization methodology of mobile devices by a considerable margin. The existing position localization methodology of mobile devices have not been able to calculate any of the evaluated location with a 40cm accuracy. But in Experiment 1, the proposed methodology have been able to calculate the accurate location with 40cm accuracy for 76.92% of the evaluated locations .which used the proposed methodology. In Experiment 2, the proposed methodology have been able to calculate the accurate location with 40cm accuracy for 61.54% of the evaluated locations. In Experiment 3, the proposed methodology have been able to calculate the accurate location with 40cm accuracy for 61.54% of the evaluated locations. Out of the three different experiments conducted results justify that the Experiment 1 which used the original images with 0.01% step size out performs Experiment 2 which used original images with 0.1% step size and Experiment 3 which used gray scaled images with 0.01 step size. When considering most correctly matched locations Experiment 1, Experiment 2 and Experiment 3 has 76.92 %, 61.54 % and 61.54 % of the evaluated locations match with an accuracy level which of 40cm respectively. Hence, Experiment 1 out performs the other two experiments on this aspect. When considering worst matched locations Experiment 2 and Experiment 3 has 11.54 % and 3.85 % of the evaluated locations match with an accuracy level which of 150cm to 500cm respectively. Where as in Experiment 1 none of the locations were match with such low accuracy. Hence, Experiment one again out performs the other two experiments on this aspect. But when considering the aspect of obtaining a timely feedback and obtaining the location value using as less time possible then Experiment 2 out performs the other two experiments as the average processing time of Experiment 2, Experiment 1, Experiment 3 are 101.55s, 828.75s, and 601.84s respectively.

Considering the accuracy level of the localization value this proposed methodology got some promising results. Therefore the proposed methodology will be useful for position localization for augmented reality based reconstructions which use mobile devices.

6.2 Conclusions

Even though the methods used in Experiment 1 and Experiment 3 did not provide a timely feedback, method used in Experiment 2 provided a feedback consuming a minimal amount of time compared to the other two experiments.

Considering the accuracy level of the localization value, this proposed methodology got some promising results which were substantially better than the existing position localization methodology of mobile devices. Out of the 26 evaluated locations, position localization values using mobile devices didn't had any location matched with at least 40cm accuracy level whereas when using the proposed methodology, an average of 66.67% of the evaluated locations were matched with this accuracy level. Out of the 26 evaluated locations, 92.31% were not matched with at least 150cm accuracy level when using A-GPS chips in mobile devices to perform the position localization. But when considering the proposed methodology only an average of 5.13% of the evaluated locations were not matched with at least 150cm accuracy.

The definitive goal of this research project was to combine image processing techniques and mobile sensor information in order to achieve less time consuming and accurate position localization using mobile devices for augmented reality based reconstructions of large objects. With the results of Experiment 1, Experiment 2 and Experiment 3 it can be said that the primary goal of this research was achieved.

Considering the secondary goals and objectives, the implemented algorithm in Chapter 4 was the algorithm which was used to identify the scale of an object in the user image and it was further developed and trained using linear regression to accurately calculate user's position from a given new user image and the proposed methodology was successfully evaluated against the assumed ground truth data which was obtained using a RTK device. Hence, all the secondary goals have been achieved under certain conditions.

Therefore the proposed methodology will be useful for position localization for augmented reality based reconstructions which use mobile devices.

6.3 Limitations

- In the delimitations it is mentioned that different this proposed solution is only for day time use as it won't consider different lighting condition. But when considering different lighting conditions, drastic changes of lighting in the day time will also be not considered.

6.4 Implications for further research

- Even though this research have been able to address the main objectives of this research and provide much more accurate position localization results than the existing position localization methodology used in mobile devices the proposed methodology is requires much more time than the existing methodology. There for another study can be carried out on trying to improve the processing time of this methodology as further enhancement.
- As also mentioned in the limitations this proposed methodology does not perform with the expected accuracy when there are drastic changes in the lighting conditions during the day time. Another future enhancement would be to overcome this issue.
- When the distance from the user's location to the object is relatively large (800cm for the evaluated object), more than 60% of the user image is filled with background information rather than the object. This has an effect on the implemented algorithm as the object size in the user image is too small track with the implemented algorithm. This can be addressed as a future enhancement.
- The average processing times of Experiment 1, 2 and 3 are 828.75s, 101.55s and 601.84s respectively. These values are dependent on the device which was used to process. Another future enhancement would be to try out this proposed methodology on a much more efficient device.

References

A. F. Seber, G., & J. Lee, A. (2003). *Linear Regression Analysis.* Auckland, New Zealand: WILEY.

Brunelli, R. (2009). *Template Matching Techniques in Computer Vision: Theory and Practice,.* Wiley.

Didier Stricker, & Thomas Kettenbach. (2001). Real-time and Markerless Vision-Based Tracking for Outdoor Augmented Reality Applications. *IEEE and ACM International Symposium on Augmented Reality (ISAR'01).*

Ebrahim Karami, Siva Prasad, & Mohamed Shehata. (2015). Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images. *Newfoundland Electrical and Computer Engineering Conference.* St. johns, Canada.

Gerhard Reitmayr, Tobias Langlotz, Daniel Wagner , Alessandro Mulloni, Gerhard Schall, Dieter Schmals, & Qi Pan. (2010, July 7-10). Simultaneous Localization and Mapping for Augmented Reality. *International Symposium on Ubiquitous Virtual Reality*.

Gerhard Schall, Daniel Wagner, Gerhard Reitmayr, Elise Taichmann, Manfred Wieser, Dieter Schmalstieg, & Bernhard Hofmann-Wellenhof. (2009). Global Pose Estimation using Multi-Sensor Fusion for Outdoor Augmented Reality. *IEEE International Symposium on Mixed and Augmented Reality 2009*, (pp. 153-162). Orlando, Florida, USA.

Jack C.P. Cheng, Keyu Chen, & Weiwei Chen. (2014). Comparison of marker-based AR and markerless AR: A case study on indoor decoration system. *Lean & Computing in Construction Congress (LC), 2 (CONVR).* Heraklion, Greece.

Jeffreys. H, & Jeffreys B. S. (1988). Lagrange's Interpolation Formula. In *Methods of Mathematical Physics, 3rd ed* (p. 260). Cambridge, England: Cambridge University Press.

JongBae Kim , & HeeSung Jun. (2008, August). Vision-based location positioning using augmented reality for indoor navigation. *IEEE Transactions on Consumer Electronics, 54* (3), 954 - 962.

Kifana, B. D., & Abdurohman , M. (2012, APRIL 04). Great Circle Distance Methode for Improving Operational Control System Based on GPS Tracking System. *International Journal on Computer Science and Engineering (IJCSE), 04*.

Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision, 50*(2), 91-110.

Ludovico Carozza, David Tingdahl, Fr´ed´eric Bosch´e, & Luc van Gool. (2012). Markerless Vision-Based Augmented Reality for Urban Planning. *Computer-Aided Civil and Infrastructure Engineering*, 1–16.

Mark Billinghurst, Adrian Clark, & Gun Lee. (2014). *A Survey of Augmented Reality* (Vol. 8). Foundations and Trends in Human-Computer Interaction. doi:10.1561/1100000049

Milgram, P., & Kishino, F. (1994, DECEMBER). A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information and Systems, Vol E77-D.*

Number of smartphone users worldwide from 2014 to 2020 (in billions). (2016). Retrieved OCTOBER 03, 2018, from Statista:

https://www.statista.com/statistics/330695/number-of-smartphone-usersworldwide/.

OpenCV 2.4.13.7 documentation. (2018, May 15). Retrieved from OpenCV: https://docs.opencv.org/2.4/modules/imgproc/doc/object_detection.html?highlight=matchtemplate#matchtemplate

Paul A. Zandbergen, & Sean J. Barbeau. (2011, JULY). Positional Accuracy of Assisted GPS Data from High-Sensitivity GPS-enabled Mobile Phones. *THE JOURNAL OF NAVIGATION, 64*, 381–399.

Qing Hong Gao, Tao Ruan Wan, Wen Tang, & Long Chen. (2017). A Stable and Accurate Marker-less Augmented Reality Registration Method. *International Conference on Cyberworlds*, (pp. 41-47).

Sidnei Alves de Araújo, & Hae Yong Kim. (2010). Color-Ciratefi: A color-based RST-invariant template matching algorithm. *17th International Conference on Systems, Signals and Image Processing*, (pp. 101-104).

Soyoung Hwang, & Donghui Yu. (2012, JULY). GPS Localization Improvement of Smartphones Using Built-in Sensors. *International Journal of Smart Home, Vol. 6*(No. 3).

Vassilios Vlahakis, John Karigiannis, Manolis Tsotros, Michael Gounaris, Luis Almeida, Didier Stricker, . . . Nikos Ioannidis. (2001). ARCHEOGUIDE: First results of an Augmented Reality, Mobile Computing System in Cultural Heritage Sites. doi:10.1145/584993.585015

Waring, E. (1779). Philos. Trans.

Welch, G., & E. Foxlin. (2002, Nov.-Dec). Motion tracking: no silver bullet, but a respectable arsenal. *IEEE Computer Graphics and Applications, 22*(6), 24 - 38. doi: 10.1109/MCG.2002.1046626

Yan Ke, & Rahul Sukthankar. (2004). PCA-SIFT: A more distinctive representation for local image descriptors. *CVPR*, (pp. 506–513).

Zandbergen, P. A. (2009). Accuracy of iPhone Locations: A Comparison of Assisted GPS, WiFi and Cellular Positioning. *Transactions in GIS*, 5–26.

Appendix A: Diagrams

Below diagram depicts several images which were used during the evaluations of this proposed methodology.

Appendix B: Code Listings

Below is the python code of the implemented approach which was used to calculate user's accurate location.

```
from sympy.solvers import solve
from sympy import Symbol
import cv2
import numpy as np
from geopy.distance import vincenty
from geopy import Point
from geopy.distance import distance, VincentyDistance
from operator import itemgetter
from GPSPhoto import gpsphoto
import time

#Tag of the images (image name without extension)
user_image_tag_arr=['user_image']

#Template matching methods
methodl = 'cv2.TM_CCOEFF_NORMED'
methodx=methodl

#Souce image reduce size
source_size = 0.2

#Range of template image sizes
lower_range = 0.0650
upper_range = 0.1400

#Step size in each iteration
step_size = 0.001

#Path for template images
path_templates = 'G:\\research\\Site 07\\templates\\JPG\\'

#Template images
template_arr = ['REF (1).jpg','REF (2).jpg','REF (3).jpg','REF (4).jpg','REF (5).jpg','REF (6).jpg','REF (7).jpg','REF (8).jpg','REF (9).jpg','REF (10).jpg',
                'REF (11).jpg','REF (12).jpg','REF (13).jpg','REF (14).jpg','REF (15).jpg','REF (16).jpg','REF (17).jpg','REF (18).jpg','REF (19).jpg',
                'REF (20).jpg','REF (21).jpg','REF (22).jpg','REF (23).jpg','REF (24).jpg','REF (25).jpg','REF (26).jpg','REF (27).jpg','REF (28).jpg',
                'REF (29).jpg','REF (30).jpg','REF (31).jpg','REF (32).jpg'
                ]
#Llatitudes of the 32 reference images
ref_latitudes = [6.9008566,6.9008588,6.9008615,6.9008634,
                 6.9008658,6.9008679,6.9008692,6.900865647,
                 6.9008513,6.90084,6.9008273,6.9008165,
                 6.9008029,6.9007893,6.900778155,6.9007579,
                 6.9007463,6.9007436,6.900742,6.9007403,
                 6.9007379,6.9007358,6.9007342,6.9007374,
                 6.9007465,6.9007569,6.9007683,6.9007795,
                 6.9007904,6.9008018,6.9008127,6.900835773]

#Longitudes of the 32 reference images
ref_longitudes = [79.8608584,79.8608493,79.8608396,79.8608304,
                  79.860822,79.860812,79.8608022,79.86077773,
                  79.8607522,79.8607476,79.8607422,79.8607383,
                  79.8607344,79.8607299,79.860726,79.8607334,
                  79.8607487,79.8607584,79.8607681,79.8607778,
                  79.8607879,79.8607984,79.8608087,79.8608362,
                  79.8608559,79.8608611,79.8608649,79.8608687,
                  79.8608721,79.8608759,79.8608806,79.86087253]

#Size of the template varies between lower range and upper range by step size
lower_range = int(lower_range*10000)
upper_range = int(upper_range*10000)
step_size   = int(step_size*10000)


#Start the clock
t0 = time.clock()

for user_image_tag in user_image_tag_arr:

    #User locatioin Given by the mobile device(Geo tag of the user image)
    data = gpsphoto.getGPSData(user_image_tag+'.jpg')
    user_location = (data['Latitude'],data['Longitude'])

    #Array for most optimal locations
    optimal_location_list = []
    sorted_optimal_list=[]
```

```
for i in range (0,32,1):
    ref_location = (ref_latitudes[i],ref_longitudes[i])

    distance = vincenty(user_location, ref_location).kilometers*1000

    optimal_location_list.append (distance)

#Array for the sorted distances from object to the user location given by the geo tag of the user image
sorted_optimal_list=sorted(optimal_location_list)


#Array to hold reference images of optimal locations
optimal_match_temp_arr=[]

for k in range (0,8,1):

    #Path of template images
    path_template =path_templates + template_arr[optimal_location_list.index(sorted_optimal_list[k])]

    #Read templates
    temp = cv2.imread(path_template,-1)

    #Append template to the array
    optimal_match_temp_arr.append(temp)



#Path of user image
path_user_image =  user_image_tag +'.jpg'

#Read user image
user_img = cv2.imread(path_user_image,-1)

#Resize the user image
image = cv2.resize(user_img,None,fx=source_size, fy=source_size, interpolation = cv2.INTER_AREA)
w, h, x = image.shape[::-1]

#Variables for best matching location
best_match_rate=0
best_match_size = 0
best_match_location = 0

#Variable to keep track of the location that is checked
count = 0

for temp in optimal_match_temp_arr:

    print("checking template : " + str(template_arr[optimal_location_list.index(sorted_optimal_list[count])]))
    accuracy_array=[]
    accuracy_array_sorted=[]
    max_accuracy = 0
    object_scale = 0

    print("------------------")

    for i in range(lower_range,upper_range,step_size):
        #Resizing the template
        size=i/10000
        template = cv2.resize(temp,None,fx=size, fy=size, interpolation = cv2.INTER_AREA)

        #Evaluating method
        method = eval(methodx)

        # If method is CCOEFF
        res = cv2.matchTemplate(template,image,method)
        min_val, max_val, min_loc, max_loc = cv2.minMaxLoc(res)
        print(str(i)+")"+str(max_val))

        accuracy_array.append([i,max_val])

    accuracy_array_sorted = sorted(accuracy_array, key=itemgetter(1),reverse=True)

    #Get the best 10 scales and accuracies for this template
    for j in range (0,10):
        max_accuracy = max_accuracy +  accuracy_array_sorted[j][1]
        object_scale = object_scale +  accuracy_array_sorted[j][0]

    #Getting the average of best matching scales for this template
    max_accuracy =max_accuracy /10
    object_scale=object_scale /10

    #Selecting the best matching location and scale
    if(best_match_rate<max_accuracy):
        best_match_rate = max_accuracy
        best_match_size = object_scale/10000
        best_match_location = optimal_location_list.index(sorted_optimal_list[count]) + 1

    count = count + 1


#Set of funcctioins at each location
x = Symbol('x')

function1=0.15959 - 0.01105 *x
function2=0.168943 - 0.01218 *x
function3=0.182493 - 0.014195 *x
function4=0.173597 - 0.013165 *x
function5=0.164583 - 0.01235 *x
function6=0.17015 - 0.01335 *x
function7=0.151683 - 0.01085 *x
function8=0.16015 - 0.010845 *x
function9=0.150717 - 0.00965 *x
function10=0.16397 - 0.01126 *x
function11=0.175823 - 0.011595 *x
function12=0.18005 - 0.0121 *x
function13=0.15887 - 0.00815 *x
function14=0.167683 - 0.01148 *x
function15=0.151323 - 0.009545 *x
function16=0.15541 - 0.010455 *x
function17=0.162883 - 0.01225 *x
function18=0.171117 - 0.01238 *x
function19=0.199457 - 0.01587 *x
function20=0.205983 - 0.0173 *x
function21=0.19659 - 0.015645 *x
function22=0.180777 - 0.013705 *x
function23=0.153903 - 0.009835 *x
function24=0.159703 - 0.011235 *x
function25=0.165247 - 0.01215 *x
function26=0.18085 - 0.0142 *x
function27=0.188813 - 0.01478 *x
function28=0.185687 - 0.014395 *x
function29=0.188783 - 0.0147 *x
function30=0.180083 - 0.01315 *x
function31=0.15309 - 0.009295 *x
function32=0.17408 - 0.0127 *x
```

```
function_list = [function1,function2,function3,function4,function5,function6,function7,function8,function9,function10,
                 function11,function12,function13,function14,function15,function16,function17,function18,function19,function20,
                 function21,function22,function23,function24,function25,function26,function27,function28,function29,function30,
                 function31,function32
                 ]


#Selecting the best function
functionx = function_list[best_match_location-1]

#Solving the best function
answer_set = solve(functionx-best_match_size,x)

#distance to the object from the user's location
distkm = (answer_set[0] - 5)/1000


#best matching location gps
lat1= ref_latitudes[best_match_location-1]
lon1= ref_longitudes[best_match_location-1]


if((best_match_location >=1) and (best_match_location<=7)):
    #Bearing of the best matching location
    bearing = 21
if((best_match_location >=9) and (best_match_location<=15)):
    #Bearing of the best matching location
    bearing = 291
if((best_match_location >=17) and (best_match_location<=23)):
    #Bearing of the best matching location
    bearing = 201
if((best_match_location >=25) and (best_match_location<=31)):
    #Bearing of the best matching location
    bearing = 111

if (best_match_location == 8):
    #Bearing of the best matching location
    bearing = 336
if (best_match_location == 16):
    #Bearing of the best matching location
    bearing = 246
if (best_match_location == 16):
    #Bearing of the best matching location
    bearing = 246
if (best_match_location == 24):
    #Bearing of the best matching location
    bearing = 156
if (best_match_location == 32):
    #Bearing of the best matching location
    bearing = 66


# Given: lat1, lon1, bearing, distMiles calculate new location data
new = VincentyDistance(kilometers=distkm).destination(Point(lat1, lon1), bearing)


lat2 = new[0]
lon2 = new[1]

#Print locatioin data
print(lat2)
print(lon2)
#Printing  the processing time
print (time.clock())
```