# Abstract

Starting a new job in an office, can be very stressful for an intern or a new employee especially if it is their ﬁrst day in office. It takes time, to learn what other employees’ job is and how they can be beneﬁcial for you and your adjustment. It might take some time for new members to learn the rooms and their purpose within the office building, as well as to understand and learn on how to use certain equipment, for example an automatic key lock or simply a coﬀee machine. The Workplace Assistant Augmented Reality will try to understand who the user is, through user proﬁling and provide the necessary process for the user to learn and understand the information relevant to the user.

The application will guide the interns through a process adjusted just for them to get to know the people around them, the building and any relevant equipment which they might make use of daily. It will have user proﬁling implemented along with object and image recognition techniques using Vuforia to overlay new information on the tablet or their mobile phones. The application will provide navigation using only image recognition techniques provided by the Vuforia Library to guide them through the building as well as information on every oﬃce, the people who might be working in that ofﬁce as well as some additional information such as how to use the coffee machine. The main expected outcome is that the users will easily get adjusted to the workplace through a user-friendly immersive experience, provided by the augmented reality application.

The research and experimentation carried out, will ultimately compare whether using Vuforia's augmentation techniques, can suffice into completing certain Augmented reality tasks and if not, better augmentation techniques that of which will be compared with Vuforia's, will be finally recommended.

# Introduction

## Problem Definition

“Person- job fit is a substantial factor for decreasing job stress and the adjustment of employees to an organization is an important issue for eliminating stress.” [Job-fitToJob-stress]

“New employees all bring expectations to their new jobs that are based on factors like their previous job experiences, their understandings of the profession, beliefs and experiences held by peers or family, promises made during recruitment, and their evaluation of the work situation during their interview.”[ Adjustment To The Work Place By New Recruits In Libraries] The ﬁrst month at the workplace might seem overwhelming. During their first few months of settling and adjustments, “a period of learning how to “fit in” and adjusting to how things work in the new setting.” may be carried out by the company for the employee’s benefit.

Providing an assistant augmented reality application, to help speed up the process for the employee to adjust to their new workplace environment may offer several challenges. There are several Augmented Reality libraries, which provide all the necessary techniques for one to build such applications, without having the need to be highly skilled in any form of programming, especially where it involves Artificial Intelligence. When it comes to feature extraction especially, if one is making use of traditional computer vision techniques such as SIFT and SURF alone can be challenging, for example, “The SIFT algorithm deals with the problem that certain image features like edges and corners are not scale-invariant. In other words, there are times when a corner looks like a corner but looks like a completely different item when the image is blown up by a few factors.” [Comparing Deep Neural Networks and Traditional Vision Algorithms in Mobile Robotics]

## Motivation

“Whilst employees can be reasonably expected to adjust to changes in jobs over time, poor job or employee job ﬁt can result in increased stress and ineﬃciency in organizations.”[Job-fit-To-Job-stress]

A workplace is defined as the environment in which people work in. Adjusting to a new environment, especially if that place is your work, can come with several challenges, such as; adjusting to new people, finding certain offices within the environment and using certain job equipment. “When humans feel a loss of control this causes physiological changes which can exacerbate feelings of stress.”[ Job-fit-To-Job-stress] Job stress has become a common term in industry, as it interests several companies in sustaining a healthy working environment for their employees. “Workload is one of the major factors which affect the employees’ productivity and efficiency. Job stress caused by high workload has become common in today’s scenario.” [Impact of Job stress on employee’s job]

Such level of stress can increase from certain adjustment work needed, for the employee to settle within a company such as, filling in papers and handing them to the right offices, and learning to use certain equipment around them. Therefore, proper training should always be provided, whether it is detailed or not. “Application of training in the workplace and proper implementation of training can directly lead to improving the employees' performance.” [A Study on the impact of on the job training]

There are two types of training, on the job training and oﬀ the job training. On the job training is a method of imparting knowledge and training directly while on the job. Oﬀ the job training is a method of imparting knowledge and training while not on the place of work, for example through a site. The idea behind it, is to minimise stress levels and allow the employee to improve without any pressure. ”Training, which aims at empowerment, development, and qualifying employees through knowledge and skills, referrers to end-oriented, organized, logical, on-going planned attempts to bring about the desired, change in the knowledge, skills, capability and attitude employees.”[A Study on the impact of on the job training]

### Why the Problem is Non-Trivial

There have been previous attempts at making indoor augmented reality applications to guide users around a place. However, most attempts are normally done using ArCore and by acquiring a 3d model of the building. ArCore is useful for catching movements and current positioning, light detection and has the anchoring feature where a virtual object is given a marker to monitor the object’s displacement. However, ArCore is incompatible with several devices which proves to be useless when applying the application in real life scenarios, as not everyone will have the latest phone with the latest specs. Vuforia on the other hand is user friendlier and can be used on several operating systems.

The second problem is that the Augmented reality application can be fed a 3d model directly to anchor positions within the map and display the respective augmented information. This can serve useful when applying indoor augmented reality navigation. However, creating a 3d model of the workplace can come with several problems. Firstly, the company would not want to freely hand out a plot of their indoor workplace, as it goes against company’s policy. Secondly, one would not be testing and experimenting with anything if a 3d model of the workplace had to be used. In this project several features will be tested from the Vuforia’s library such as feature detection, and the library will be used to its full potential.

## Approach

The proposed solution is to develop a workplace assistant augmented reality (WAAR) application. The application will assist the users by providing them with augmented reality information to; guide them to offices, provide them with information of offices and rooms while walking in the corridor, and instructions on how to use the office’s coffee machine. The application will make use of user profiling techniques, to understand the user’s requirements, and display relevant information which is related to what they intend to use the application. A form will be needed to be filled before using the application. The form will be quite short, and data shall not be stored anywhere. The data collected will be used only to display relevant markers on the augmented reality application. Once the application is closed all data shall be forgotten about the previous user, at least for our testing purposes only.

Augmented Reality development will be handled by Vuforia’s libraries as it has some features which the application can well beneﬁt from. It” enables businesses and app developers to quickly spin-up high ﬁdelity, mobile-centric, immersive AR experiences.” [ M. Romilly. 12 Best Augmented Reality SDKs. (2019, Jan 25).] For our research, use of its image and object segmentation shall be used to identify ofﬁce workplace markers, along with Unity, the proper content shall be overlaid using game objects. There shall be cases where model target along with Vuforia’s deep learning techniques shall be used to scan some objects in 3D. Vuforia is ideal because it can develop augmented reality application for Android and IOS devices.

Indoor navigation can be done in several ways. One can use GPS signals, beacons, RSS or WIFI signals, or simple Augmented Reality itself. Now Augmented reality can be location based or marker based. So, the proposed solution for our problem would be using Augmented Reality marker-based navigation. Several markers around the oﬃce building will be used to segment images or objects and the proper directions will be displayed by recognizing the markers in view. This will allow the company to keep the application useful for oﬄine use. For scenarios when WIFI or any other signals are down, user can still make good use of the application, for example in case one might need it during an emergency to ﬁnd the ﬁre exit.

## Aims and Objectives

### Aim

The aim of this project is to research and develop a workplace assistant augmented reality application using image and object detection provided by Vuforia, which are filtered through user profiling.

### Objectives

* Collect images and perform image and object detection techniques using the Vuforia Library.
* Use Augmented Reality techniques from the detected images and objects to overlay and augment information as well as navigation information.
* Provide user profiling to filter out unnecessary information for augmentation.
* Compare and contrast other image and object recognition techniques apart from the ones provided by Vuforia.

### Report Layout

The layout of the report is as follows. Chapter 2, an overview of the background behind the technologies used. Chapter 3 contains the literature review done in attempting to solve the problem at hand. Chapter 4, a brief overview of the system as well with its design. Chapter 5 will be discussing the implementation process. Chapter 6 discusses the evaluation methods and approaches for the application, it will include both user and AI evaluation, as well as results obtained. Chapter 7 will outline the limitations, challenges, and future possibilities into further development of the application along with the technologies used. Finally, the project will end with a conclusion.

* Workplace
* Augmented Reality
* User Profiling
* Vuforia
* Human computer Interaction Techniques and Methods
* Image and object Detection Techniques
* ----------------------------------------------------------------------------------

### User Profiling

“User Profiling is the process of Extracting, Integrating and Identifying the keyword-based information to generate a structured Profile and then visualizing the knowledge out of these findings.” \cite{KANOJE20165} Through user profiling the system is capable to tailor the required information for the user to see and make use of. It is annoying for users to have to go through irrelevant documents or data to find what is specific to what they require.

"User profile generation is done when we get users’ complete information while he registers into our system. We have identified different user attributes for profiling him into our system".\cite{KANOJE20165} User profiling has taken the form of recommender systems, where the system provides user specific recommendations in a personalized form. There are two forms of User profiling. Explicit User profiling, “In this approach users’ behaviour is predicted by analysing the user’s available data” \cite{Kanoje2015UserPT}. This is known also as static profiling, in which analysis of static and predictable user data is made. The second type is implicit User profiling, it “relies more on what we have known about user in future i.e. systems tries to learn more about the user.” \cite{Kanoje2015UserPT}. It is also referred to as Adaptive Profiling. After performing extraction , one might end up with redundant information.

To clean the information and see unique pieces of information one must perform filtering. There are three filtering techniques for user profiling, rule-based, collaborative and content based filtering techniques. Content based filtering, "recommends items based on a comparison between the content of the items with a user profile and selects those items whose content best matches with the content of another item." \cite{Kanoje2015UserPT} Collaborative Filtering is the process of grouping users with a similar search criteria. The filtering is based on previous sought items as well and items which they are more likely to search for next.

### Augmented Reality

“Augmented Reality (AR) is a new technology that involves the overlay of computer graphics on the real world”[ Introduction to Augmented Reality]. It is a term which refers to mixed reality, where the digital world and reality are combined and interwoven together. Augmented Reality is a new form of technology that focuses on displaying realistic overlays on reality to provide extra information and content to what we see with our naked eye. It is an enhanced version of a real-world environment, through the form of media devices such pictures, videos, 3D models and sounds. “Therefore, AR can be differentiated from Virtual Reality, given that the former overlays digital information in a real environment, rather than completely replacing it.”\cite{CROFTON2019102178}.

There are different categories of Augmented Reality. The first category is, marker-based AR, where the augmented overlay is only displayed once a marker is detected through a camera. It is also known as image recognition. The second category is, markerless augmented, which makes use of an accelerometer, a GPS and velocity tracker to detect the location of the phone and display the AR overlay in that specific location, given its location is predefined. The third category is projection based which basically projects data in the form of light rays on objects, for example an augmented/ projected keyboard. The last category is superimposed AR, where the AR partially replaces the real view with an augmented one of that object, IKEA makes use of this application through their digital catalogues.

There are several Augmented Reality Devices. The first category are Optical See-Through HMD. “Optical See-Through AR uses a transparent Head Mounted Display to show the virtual environment directly over the real world” [ Introduction to Augmented Reality]. HMD performs best when its perfectly fits to the users eyes and sit comfortable on their face, making it easy for the users to move around with them worn. The second category are Virtual Retinal Devices. “The VRD projects a modulated beam of light (from an electronic source) directly onto the retina of the eye producing a rasterized image” [ Introduction to Augmented Reality]. The third category Video See-Through HMDs, the monitor based Augmented Reality which “uses merged video streams but the display is a more conventional desktop monitor or a hand held display” [ Introduction to Augmented Reality]. Finally the projection displays which projects on surfaces and is useful for multiple user interaction. One such example of projection based AR is Tilt Five.

### Vuforia

Vuforia is the ``most popular SDK for developing AR-applications on a wide selection of devices."\cite{Grahn2017TheVS}. Similar to ARCore and ARKit, Vuforia can be used on multiple devices to recognise images, objects and text. One can use third party engines such as Unreal or Unity to display overlays to create Augmented Reality apps.

Image recognition works by detecting natural features such as edges and corners in an image. ``the feature tracking algorithm can determine what is a feature and map the positions of these features in the image." \cite{Grahn2017TheVS}. By shifting the positions of the image features like edges are intensified, even more corners as their position changes after shifting. Vuforia thus, makes use of Pose feature detection techniques where it takes into consideration the position and orientation of the natural features. It can make use of extended tracking, where the engine detects surrounding features as well. A proper image with high quality feature detection is an image that contains uniquely distinct features, which are not repetitive. For example, a dark circle is difficult to recognize and establish uniquely features.

Object Recognition in Vuforia works in several ways. The first method is by using a scanning plane which acts as an image target. Once it is recognized, ``provides the scanner with the XYZ-coordinates of the detected features on the 3D object."\cite{Grahn2017TheVS}. The latest form of object detection is by using their latest release, modelling targets. They have added a new feature for target modelling using their own database on the cloud, where it contains multiple views and/or multiple objects with automatic recognition. It makes use of a deep learning library which is added as an SDK. One only needs to train a model target database where your 3D model is uploaded to the Cloud for training purposes. Text recognition in Vuforia makes use of Optical character recognition tools. The image cause through segmentation and the characters are extracted from the image background.

### Human Computer Interaction Techniques and Methods

There are several approaches to Human computer interaction designing. Firstly, the Anthropomorphic Approach, which is the designing of a human like interface. Secondly is the Cognitive Approach, which takes into consideration cognitive abilities of the human for interfacing. Thirdly is the Empirical approach, which compares usability of multiple conceptual designs. This involves performing Human task performance measures and A/B Testing. Finally, the last approach is Predictive Modelling Approach. It is the examination of the amount of time the user took to complete a task using the interface provided.

There are eight research methods with respect to mobile HCI research. The first one is case studies. Case studies are intensive empirical studies using groups, organizations, individuals, systems or tools. Data is collected to qualitative and quantitative approaches. The second is Field studies, "Ethnographic field studies are characterized by researchers spending significant amounts of time in the field and, to some extent, immersing themselves into the environment they study."\cite{KJJEGRCO}. Third one is Action research, fourth one is Laboratory Experiments where a controlled environment is used for research purpose, Fifth is survey research, sixth one is applied research where it "builds on trial and error on the basis of the researchers capabilities of reasoning through intuition, experience, deduction and induction."\cite{KJJEGRCO}.The final two research method are basic research and normative writings.

### Augmented reality Navigation

Outdoor navigation normally makes use of GPS localisation. However, for indoor navigation this can be a problem. There are several ways how to provide indoor localisation. One can make use of beams either by Bluetooth signals or WIFI signals. This will provide the user with continuous mapping, but it has also resulted to be buggy at times. The alternative to that is by using offline waypoints. The user simply scans a marker to get a location or augment pre-programmed information within that location. "the user needs to update his/her location by scanning another way-point on the way."\cite{Bhorkar2017ASO}.

One main challenge in augmented reality navigation is the process of registration. “Registration is the process of correctly aligning the virtual information with the real world in order to preserve the illusion of coexistence.”\cite{Bhorkar2017ASO}. Although proper visual registrations must be met for the Augmentation to be as realistic as possible, one must not forget that the user still needs to focus on what is on their path.

The improvement of AR can help provide navigation information without distracting the user from looking away to a secondary screen or view, “For example, showing navigation markers on the windshield of the car or augmenting the video camera output of a smartphone with the navigation path”\cite{Bhorkar2017ASO}. To provide an augmented reality navigation system there are several steps one need to take, “1. Acquire the real-world view from the user’s perspective. 2. Acquire the location information for tracking the user. 3. Generate the virtual world information based on the real-world view and the location information. 4. Register the virtual information generated with the real-world view.”\cite{Bhorkar2017ASO}.

### Deep Learning in Augmented Reality

The detection problem has been solved using camera-based tracking systems to apply it to Augmeted Reality, using deep learning techniques. The Vuforia Library has applied such techniques to scan 3D objects and create model targets for them, to be easily recognizable within any developed AR app. This provides new advantages such as detectability from any angle of the recognizable real-world object. “Known model of the object can be used to determine the position and orientation of the object. Rendering of the virtual object follows easily.”\cite{Akgul2016ApplyingDL}. There are two ways how the object can be recognized. As previously discussed one can use traditional artificial vision techniques or use Convolutional Neural Networks for improved detection.

Model-based AR tracking is achievable in two steps. Firstly, using video tracking which “yields the pose of the camera with respect to the known target.”\cite{Akgul2016ApplyingDL}. Secondly, the pose is sent to an algorithm for tracking. For detection, algorithms such as SIFT and SURF are commonly used. Vuforia makes use of a CNN which is based off AlexNet. It extracts a number of key points using a corner detection algorithm such as FAST\cite{CornerDetection}. In \cite{ Akgul2016ApplyingDL} a CNN implementation was trained using AlexNet to detect patches. FAST was used to detect features on a reference image, it extracts 15 by 15 patches across each feature. HIPS \cite{Taylor2011} was used for 8 by 8 for sparse sampled patches from the original set of patches. When comparing the overall performance of the CNN used in \cite{Akgul2016ApplyingDL} with an algorithm such as ORB the re-projection error shows that it was far improved in the DeepAR. “DeepAR method produces consistently more inliers than HIPS. However, as can be seen in Figure 12 the percentage of inliers vs outliers are less for DeepAR.”\cite{ Akgul2016ApplyingDL}.

In the study conducted within \cite{ Akgul2016ApplyingDL}, it is concluded that “The detector performance is very strong especially in the presence of error in feature localization”\cite{ Akgul2016ApplyingDL}. It is comparable to one of the best feature detection algorithms to date. Within their experiment the CNN was kept in its most simplistic form to detect 2D objects. However, it is expected for Vuforia’s CNN to perform better due to its 3D object tracking and detection abilities.