# Literature Review

The chapter serves as a review of the available research that was done in the area of Workplace Assistant Augmented Reality. I shall be discussing the two components involved in the Augmented Reality application as well as the research that has inspired my approach. The chapter is divided into three parts; the first part involves Image and Object Recognition techniques involved in Augmented Reality, the second is about applying user profiling methods with Augmented Reality, and the third is about different image and object recognition techniques involved in augmented reality technologies.

## Workplace Augmented reality

``Augmented Reality (AR) technology has rarely been discussed outside of the computer science world. It has taken years for this technology to become closer to a stable existence and will most likely take several more years before it will be used by average citizens" [9]. Augmented Reality has a wide variety of applications. One of its main applications in the 4.0 Industry is the use of AR in assisted learning. Every workplace needs adjustment and some form of training for employees to be adjusted to the process of the work they might be doing. Augmented Reality may help in assisting the employees by providing them with additional overlaid instructions to guide them through the whole process of adjustment as well as to provide them with training.

Workplace training normally comes in two forms: on the job training and off the job training. ``OJT may be viewed as an apprenticeship where a novice AMT is mentored by an AMT who is an expert" [12]. It is a traditional form of training, especially for teaching maintenance. However, ``OJT may not be the best method for training because the feedback to learners may be infrequent and unmethodical" [12]. Off the job training maybe provided through face to face conversations or through use of multimedia. Augmented Reality can combine the two aspects of training into one. Where the user is given on the job training using multimedia, which is overlaid on top of the real-world environment.

There are several useful applications for Augmented Reality at a workplace. However, not every workplace might necessitate for AR. ``there are situations where an AR system may be used to enhance the task completion process or display and/or communication of information in conjunction with traditional technologies" [9]. As discussed in [9], the following are workplace conditions where AR is applicable; distance communication with 2D or 3D objects provided for visualization, training and education when making use of real life tools, recording of information obtained while training, and a collaborative design and interaction of 3D models is required.

An advantage which Augmented Reality provides is, for the workers and managers ``the ability to author their own environment by embedding the relevant information needed for task completion" [9]. The common problem which is face during work training is for the expert individual to provide the respective information to the trainee in the most understandable way possible. Through AR technology the trainee can tailor how that information is presented. Therefore, Augmented Reality may be capable of understanding its user's such that it may adapt to future possible users.

## Recommendation System for Augmented Reality

Information during a job training is crucial for an employee to learn and adjust to the new environment. However, an overwhelming amount of information directed towards a new employee may demotivate them. Augmented Reality is a tool for providing interactive information towards the user as well as from the user. Despite that, ``The fact that the typical scene of these applications mix real and virtual elements can be a motivating factor for users. However, this feature may also make the interaction more complicated, which can affect the user experience in performing tasks within the application" [25].

``Recommender systems (RS) have proven to be a valuable tool for online users to cope with the information overload" [7]. Recommender systems provide tailor made information to different users based on the users' preference. ``Thus, it is important to oﬀer the user a personal response, but also a context-dependent and constrained by the limited computing capacities of the mobile devices" [7,2,29,22].

Collaborative filtering techniques have been widely adapted in recommender systems. However, traditional recommender system in Augmented Reality cannot be easily adapted and deployed since they differentiate in the following areas: location, timing, first time use of the application, and immediate response from the AR application as discussed in [30]. In augmented reality distance-based filtering and visibility-based filtering are commonly used. In [30] a random walk algorithm was incorporated, which recommendations are based on user preferences, behaviour patterns, history records and information from social media. ``The stationary distribution of the random walk represents the ranking score, inspired by the PageRank algorithm"[30]. Location information along with the user's personal preference for recommendation in the random walk algorithm.

An alternative to using location or distance-based recommendation, Augmented Reality applications can make use of time-based recommendation systems. By time-based meaning, the amount of time one would generally spend on completing a task using the AR application. A task may take a lot of time to be completed by the user due to several factors, such as the task itself is complicated or simply the AR app is incapable of providing the user with the right instructions and guidance into solving the task, simply because it may lack different forms of interactive techniques. In [25], ``a set of procedures to conduct experiments with users to identify how a set of aspects related to the user proﬁle can be considered to improve mobile AR technology usage" [25] were defined.

The target audience is crucial when providing a final end-user Augmented Reality product. One must keep in mind that the users' age might greatly vary especially in a workplace. Some people might suffer from eyesight problems therefore small pieces of text will not be helpful for them. The application cannot assume that the user will easily interpret what is being overlaid on the screen. They might need to be guided along, as to understand what different symbols, colour and size of the symbols being displayed might mean.

## Computer Vision Approaches in Augmented Reality

Augmented Reality applications make use of several computer vision approaches to recognize images, objects and text. As previously discussed, Vuforia makes use of both traditional and deep learning approaches. Using deep neural networks will ensure highly accurate and efficient results. However, ``it is well-known that training high capacity models such as deep neural networks require huge amounts of labelled training data. This is particularly problematic for tasks where annotating even a single image requires significant human effort" [4].

As discussed in [18] marker-based applications have been the main driving force of applying augmented reality in real life. ``Most of the current approaches to 3D tracking are based on what can be called recursive tracking" [18]. Therefore, the system must be initialized by hand and with a little bit of occlusion between the camera and the object being recognized the system fails to perform. However, a new computer vision approach has improved augmented reality, which can register the camera without camera pose introduction. This approach is called Tracking-by-Detection and in [18] it is tested to see its benefits. The approach works by extracting feature points from inputted frames during run-time. The features are then ``matched against a database of feature points for which the 3D locations are known" [18]. Traditional approaches are then used to improve the estimation of the calculated pose.

According to the research conducted in [21], in order to solve the occlusion problem within current Augmented Reality technology a S-G Hybrid Recognition method was implemented. The approach takes ``advantage of robustness of the SURF feature-based object identification and combine it with high reliability and effectiveness of the Golay error correction code detection" [21]. SURF along with SIFT are two traditional vision approaches commonly used for feature-based detection. The advantage of SURF is scale and rotation in-variance. Golay error correction code on the other hand is a marker identification approach. ``A marker based on the Golay error correction code (ECC) can be composed of a large white square in the top left corner and e.g. 24 black or white squares that encode a number. The large square provides information about the marker orientation" [21]. They tested the three main aspects which may hinder an AR application: distance variance, angle variance, and occlusion. As a result, the S-G approach it was found that an object can be placed 2m away from the camera, the angles comparison was completely influenced by the SURF algorithm where it was able to detect under 55 degrees angle to the camera's axis, and that it could not be affected by up to 55\% obstruction.

Another approach into solving the occlusion problem in augmented reality is by applying deep learning techniques as described in [10]. In this research, they ``present a temporal 6-DOF tracking method which leverages deep learning to achieve state-of-the-art performance on challenging datasets of real-world capture"[10]. Deep learning architectures can be trained on large amounts of data, and as a result this solves the occlusion, angle variance, and distance variance problem. Their approach involved getting a 3D model of the object and training the tracker for that specific object. Training involved two steps; first one was using a frame to capture the object in its predicted position, and secondly the frame of the object's actual position. ``To encourage the network to be robust to a variety of situations, we synthesize both these frames by rendering a 3D model of the object and simulating realistic capture conditions including object positions, backgrounds, noise, and lighting" [10].

Deep learning architectures work well when making use of GPUs. The GPU is commonly used to run deep learning neural networks; hence the network takes less processing time to train and test. In [20], is presented ``YOLO-LITE, a real-time object detection model developed to run on portable devices such as a laptop or cell phone lacking a Graphics Processing Unit (GPU)" [20]. YOLO-LITE primarily is designed to obtain a smaller, faster and more efficient model. `You Only Look Once (YOLO) was developed to create a one step process involving detection and classiﬁcation. Bounding box and class predictions are made after one evaluation of the input image"[20]. The architecture developed runs at 10 frames per second, and its goal is to prove that shallow networks can run on non-gpu devices, and that shallow networks do not require batch normalization. The model had 18 trials achieving results of 33.77\% mAP and 21 FPS, and 12.26\% and 21 on PASCAL VOC and COCO dataset respectively.

## Conclusion

Different approaches were defined and revised in this chapter. Traditional computer vision techniques, deep learning techniques, recommender systems, and augmented reality solutions were studied to obtain relevant information to acquire a state-of-the-art Workplace Assistant Augmented Reality application. In the following chapters, the design and implementation of the proposed method shall be presented.

# References

[1] Advances in computer vision. Advances in Intelligent Systems and Computing, 2020.

[2] Gediminas Adomavicius and Alexander Tuzhilin. Context-Aware Recommender Systems, pages 217–253. Springer US, Boston, MA, 2011.

[3] Omer Akgul, H. Ibrahim Penekli, and Yakup Genc. Applying deep learning in augmented reality tracking. 2016 12th International Conference on Signal-Image Technology Internet-Based Systems (SITIS), pages 47–54, 2016.

[4] Hassan Alhaija, Siva Mustikovela, Lars Mescheder, Andreas Geiger, and Carsten Rother. Augmented reality meets computer vision : Eﬃcient data generation for urban driving scenes. International Journal of Computer Vision, 08 2017.

[5] Nader Barzegar and Shahroz Farjad. A study on the impact of on the job training courses on the staﬀ performance (a case study). Procedia - Social and Behavioral Sciences, 29:1942 – 1949, 2011. The 2nd International Conference on Education and Educational Psychology 2011.

[6] Gaurav Bhorkar. A survey of augmented reality navigation. ArXiv, abs/1708.05006, 2017.

[7] Carlos Plaza de Miguel. Arlodge : Context-aware recommender system based on augmented reality to assist on the accommodation search process. 2014.

[8] Nevin Deniz, Aral Noyan, and ¨Oznur Gu¨len Ertosun. Linking person-job ﬁt to job stress: The mediating eﬀect of perceived person-organization ﬁt. Procedia - Social and Behavioral Sciences, 207:369 – 376, 2015. 11th International Strategic Management Conference.

[9] James Ford and Tobias Ho¨llerer. Augmented reality: Information for workplace decision-makers, managers, workers and researchers. 02 2020.

[10] Mathieu Garon and Jean-Fran¸cois Lalonde. Deep 6-DOF tracking. IEEE Transactions on Visualization and Computer Graphics, 23(11), 2017.

[11] Ivar Grahn. The vuforia sdk and unity3d game engine : Evaluating performance on android devices. 2017.

[12] T. Haritos and N. D. Macchiarella. A mobile application of augmented reality for aerospace maintenance training. In 24th Digital Avionics Systems Conference, volume 1, pages 5.B.3–5.1, Oct 2005.

[13] Sumitkumar Kanoje, Sheetal Girase, and Debajyoti Mukhopadhyay. User proﬁling trends, techniques and applications. ArXiv, abs/1503.07474, 2015.

[14] Sumitkumar Kanoje, Debajyoti Mukhopadhyay, and Sheetal Girase. User proﬁling for university recommender system using automatic information retrieval. Procedia Computer Science, 78:5 – 12, 2016. 1st International Conference on Information Security Privacy 2015.

[15] Sneha kasetty sudarshan. AUGMENTED REALITY IN MOBILE DEVICES. PhD thesis, 05 2017.

[16] Stan Kurkovsky, Ranjana Koshy, Vivian Novak, and Peter Szul. Current issues in handheld augmented reality. pages 68–72, 06 2012.

[17] Andy Lee. Comparing deep neural networks and traditional vision algorithms in mobile robotics. 2016.

[18] Vincent Lepetit. On computer vision for augmented reality. pages 13 – 16, 08 2008.

[19] Paramjinang Moita. Adjustment to the work place by new recruits in libraries. Vol.5(2):71–85, 04 2015.

[20] Jonathan Pedoeem and Rachel Huang. YOLO-LITE: A real-time object detection algorithm optimized for non-gpu computers. CoRR, abs/1811.05588, 2018.

[21] David Proch´azka, Ondˇrej Popelka, Tomas Koubek, Jaromir Landa, and Jan Kolomaznik. Hybrid surf-golay marker detection method for augmented reality applications. Journal of WSCG, 20:197–204, 01 2012.

[22] Francesco Ricci, Lior Rokach, and Bracha Shapira. Recommender Systems Handbook, volume 1-35, pages 1–35. 10 2010.

[23] M. Romilly. 12 Best Augmented Reality SDKs. (2019, Jan 25).

[24] Edward Rosten and Tom Drummond. Machine learning for high-speed corner detection. In Aleˇs Leonardis, Horst Bischof, and Axel Pinz, editors, Computer Vision – ECCV 2006, pages 430–443, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[25] S. R. R. Sanches, M. Oizumi, C. Oliveira, E. F. Damasceno, and A. C. Sementille. Aspects of user proﬁles that can improve mobile augmented reality usage. In 2017 19th Symposium on Virtual and Augmented Reality (SVR), pages 236–242, Nov 2017.

[26] Rodrigo Silva, Jauvane Oliveira, and G. Giraldi. Introduction to augmented reality. 01 2003.

[27] Simon Taylor and Tom Drummond. Binary histogrammed intensity patches for eﬃcient and robust matching. International Journal of Computer Vision, 94(2):241–265, Sep 2011.

[28] Mathangi Vijayan. Impact of job stress on employees’ job performance in aavin, coimbatore. 06 2018.

[29] Hongzhi Yin, Yizhou Sun, Bin Cui, Zhiting Hu, and Ling Chen. Lcars: a locationcontent-aware recommender system. pages 221–229, 08 2013.

[30] Zhuo Zhang, Shang Shang, Sanjeev R. Kulkarni, and Pan Hui. Improving augmented reality using recommender systems. In Proceedings of the 7th ACM Conference on Recommender Systems, RecSys ’13, page 173–176, New York, NY, USA, 2013. Association for Computing Machinery.