Proceedings of the Second International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2020)
IEEE Xplore Part Number: CFP20K58-ART; ISBN: 978-1-7281-4167-1

Augmented Reality for Assistive Maintenance and Real-Time Failure Analysis in Industries

Jaya Sai Kiran Patibandla *
saikiran@dimensionnxg.com

Sunit Kumar Adhikary *
sunitk@dimensionnxg.com

Janmesh Ukey *
janmeshukey@dimensionnxg.com

* Dimension NXG, Thane, INDIA

***Abstract*— We present a methodology to solve the problem of maintenance for any machinery using augmented reality (AR) guided assistive systems. The same system can be used to implement real time fault analytics system. Given a visual feed of object, our methodology can accurately determines, tracks and maps the object in the world. The objects are categorized as stationary and nonstationary by the user. The first method presented will address the maintenance solution for object which are stationary using spatial mapping and localization technique. The second method presented will address the problem using a deep learning model that predicts accurate pose of the object in world space. The instruction sets and the necessary animations are overlaid on the objects. This methodology of ours to find the object pose is robust to occlusion, lighting conditions and works in real time on computationally inexpensive hardware.**

Keywords - AR Maintenance, Spatial Mapping, AR, 6 DOF pose, SLAM

I. INTRODUCTION

Over the last few years with the increase in mobile computing technology, the use case of augmented reality (AR), virtual reality (VR) and mixed reality (MR) has increased drastically. AR allows the augmentation of real world environment with virtual objects. VR on the other hand allows the user to interact with simulated virtual world. MR is a mix of AR and VR via immersive technologies [1].

Even though the first use case of Immersive technologies was proposed by Sutherland [2], hardware and optics constraints has restricted this technology to research. Recent innovations in wearable computing and the wide distribution of smartphones have revived the topic and sparked the development of numerous new applications [3].

Now with abundance in availability, AR/VR technologies offers intuitive and natural interactions that can be used to enhance awareness during different phases in the product life cycle, such as virtual prototyping, maintenance, simulation visualization[4]. Taking advantage of AR technology, applications can be created and used on the field to facilitate the maintaining processes, aiming to reduce the operating time and the necessity of an expert present during the operation

AR/VR technologies are being employed across wide areas such as education, medicine, industry, defense, construction and in many other applications . In industries, digital twin solution integrated with data from multiple sensors helps in maintenance and knowing the working status of machine [6]. In Education, to upskill industrial labour combined learning is used. Combined learning methodology is the process of linking the theoretical knowledge that we obtain in the classroom/workshop to the field work, as the resources for practical's are very expensive and are not easily accessible, technology has brought immersive tools in the form of AR/VR that can substantiate the physical resource [7]. Combined learning enhances the worker's performance and offers inexpensive training solution [8]

In order to provide the user a truly immersive experience, requires a robust inside-out tracking and tracking of external objects becomes pivotal. In this paper, we propose a generalized solutions to address the issues that would benefit the industries. Maintenance of a machine is one of the major time consuming process during a product life cycle. This maintenance involves checking each and every step that is present in the user manual. Usually the process is lengthy and difficult to skim to manual book, to address this maintenance issue, an efficient and low computing solution is introduced. The limitations will be addressed in preceding section along with a robust solution that can work across a wide variety of objects that comprises both mobile and immobile, by using the state-of-the-art deep learning technology [9]. Further, the same pipeline can be used to implement the fault detection.

II. RELATED WORK

Augmented Reality complements the real world by superimposing virtual objects in the user's environment and allows for the manipulation of the environment in real time. Milgram et al. Described augmented reality as overlaying virtual content on top of real environment.

In medicine [10] implemented a tool to visualize the medical data obtained from CT and MRI scans in real time using a video see through Head Mounted Device(HMD) [11] has detailed the usage of AR, VR technology in the various fields of medicine in depth.

In education and training [12] used it to create a simulator to train for welding. [13] used the AR, VR to simulate the manipulation operation of centrifugal pumps and also simulated industrial environment to provide immersive user experience.

In manufacturing industries, [14] proposed a virtual assembly system to evaluate the assemblability to replace the physical evaluation as much as possible. [15] used an AR system for visualization of maintenance, [16] employed AR for the maintenance of a Flapper valve. [17] Implemented an AR system for fault detection and troubleshoot HVAC systems. [18] Adopted AR guided product disassembly and assembly during maintenance.

In relation to the work we will be presenting, [19] proposed a method that can track objects in real time, eventually this leads to overlaying of virtual content over real environment.

[20] Introduced the notion of indoor environment reconstruction for augmented reality. [8], [21], [9] used deep neural networks to estimate the position of object in world frame. We imbibed the methods of [19], [20], [9] and built a generalized system that can be used for maintenance, training and for fault analysis in industries. In this work we focus on how AR can be used for robust maintenance of industrial appliances and fault analysis.

III. METHODOLOGY

Maintenance operations are tedious tasks as they involves checking each and every step of a manual guide for the corresponding machinery. Usually these applications are lengthy and involves carrying large volumes of manuals and in some cases an extra resource just to help him figure out the process from the manuals. Sometimes the user has to access places where he might not have enough room to get extra material. Moreover traditional techniques can't guide the worker on the go. We present a simple yet powerful solution to address this issue using a low cost indigenous AR headset.

We will explain the two methods that we incorporated in our system to deal with dynamic and static objects. Method 1 explains the process that we used for Maintenance architecture, Method 2 explains the architecture that we employed for dynamic and fast moving objects.

a) **Method 1**

A typical AR system can blend virtual content over real world, Interact with augmented world, localize in the real environment and map the real world. The figure below describe how a typical AR system works

In the figure 1, overview of AR pipeline has been described. AR systems uses generally inside out tracking [22]. In our pipeline we are using a stereo camera for Tracking. We are incrementally reconstructing the environment by fusing the depth map and pose from the stereo camera. The process of SLAM and real time reconstruction [20] are beyond the scope of this paper. The pose is sent to UNITY3D and virtual objects are rendered on the standalone headset. We are using the same pipeline. First we reconstruct the environment with the help of pose obtained from SLAM and depth images. Along with dense reconstruction, our SLAM stores the feature based 3d points and the corresponding features. These 3d points corresponding to the features are stored and are further used to relocalize during the track loss of SLAM [23].

We save the sparse map of the environment along with the feature map. We convert the dense map into formats that are compatible with UNITY3D. We are converting the map to obj file and are performing post processing to decrease the level of detail of mesh and to fill occlusions. The map is loaded into Unity3d and the virtual markers and instruction sets are placed as required for creating the AR experiences.

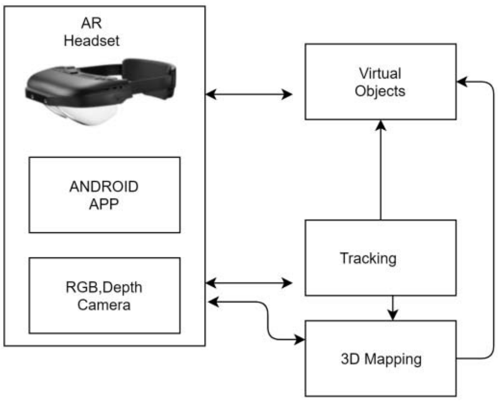


Figure1. System overview showing the overall steps in end to end AR pipeline.

The application is loaded onto AR device and the sparse feature map of the corresponding room is loaded, since the dense map is also stored previously, we load the dense map of the environment corresponding to the sparse feature map. Now with the help of sparse feature SLAM, the position of camera gets relocated. The virtual content appears when the user approaches the object of interest, he will be able to interact with the instruction placed on the object. Since the slam is robust to any outliers and drifts the virtual content appears always on the objects which they are placed. This procedure is computationally inexpensive unlike the model based tracking and other solutions. Here we take advantage of dense reconstruction to overlay virtual content unlike the model based on surface registration based techniques.

Even though this method is computationally inexpensive this is only useful when the object of interest is static. To resolve this we also developed another solution that can track objects in dynamic environments

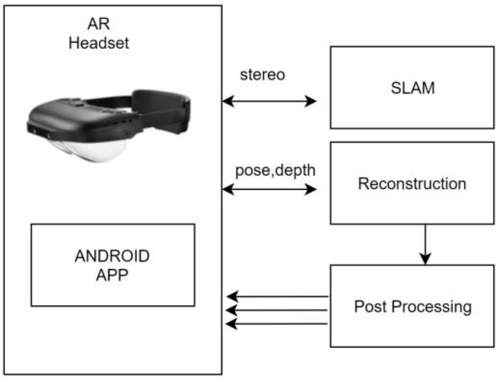


Figure2. System overview showing the overall steps in end to end AR pipeline.

b) **Method 2**

The alternative solution deals with real environment situations where the objects rapidly change position is a scene, in such an environment we obtain the pose of any object of interest using a novel deep learning technique. Algorithm is particularly designed keeping the hardware in mind. This method is robust, works in real time irrespective of the lighting conditions.

Pose estimation from RGB images as in [8][21] is done by estimating the projections of 3D bounding box in 2D images and doing a PnP with the real 3D coordinates , which is not feasible in a real-life scenario as this. In [9] Nigam et. al explores predicting the complete 6Dof pose from RGB images. However due to the absence of depth data this poses may not be very accurate. We exploit both RGB and depth data to obtain pose of an object without any manual intervention.

In this technique we first identify and segment each instance of the object in the scene. We extract the depth image corresponding to the segmented object in the RGB frame and obtain the corresponding point cloud. We process this point cloud to segment the object of interest in the 3D space. We do a segmentation in the 3D space, as segmentation performed in the 2D space need not correspond to a segmentation in the 3D space. The point cloud corresponding to the segmented 3D object is then processed through a deep learning technique to obtain the pose. It is summarized in Algorithm 1. Once the objects are segmented we use predicted pose in 3D space to place anchors/instructions/animations on objects like we did in method 1. The next sections presents the results that we obtained using our generalized framework.

Algorithm 1 Obtain Pose And Understand Scene

```
while !terminate do
  RGB ← xnxgSensor.getRGB()
  Depth ← xnxgSensor.getDepth()
  instanceSeg ← xnxgIdentifyAndSegmentObject(RGB)
  foreach object ∈ instanceSeg do
    pcdSegmentedObject ← segmentObject3D(RGB, Depth, Object)
    poseObject ← poseSegmentedObject(pcdSegmentedObject)
end
```

IV. RESULTS

This paper concentrates on discussing the methods that would help anyone in accelerating the maintenance process. As discussed in the previous sections, we were using a novel SLAM solution for tracking and have developed our own dense reconstruction systems and developed deep leaning based network for specific tracking of dynamic objects. In this study, we have taken a modular drawer cabinet, usually in industries, parts of different items are placed in different cabinets and during maintenance, assembly or any general purpose operation, the user has to open these cabinets and fetch the parts. Here we present an example for the proposed AR assist system that can be used to enhance maintenance of machinery and for monitoring the health of machines. In the above figure 3.a we see a drawer cabinet, and fig 3.b shows the corresponding dense 3d reconstruction. In fig 3.c we overlay virtual markers on the drawer from the instruction set. These markers are placed on the map by loading the map into unity3d. Fig 3.d shows the virtual tray placed inside the reconstructed to simulate real functionality of drawer. When the user approaches this cabinet drawer, the arrows spawns on the object that we need to interact with and a voice solution dictates the instructions. The figure 4.a shows the overlaying of markers on a real world

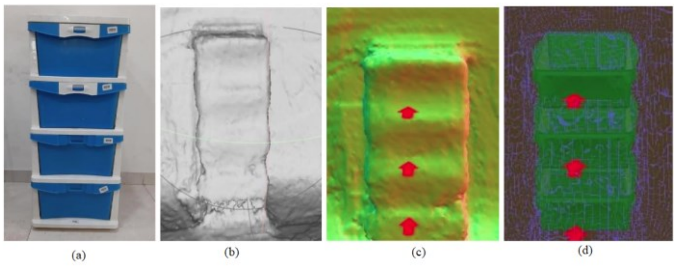
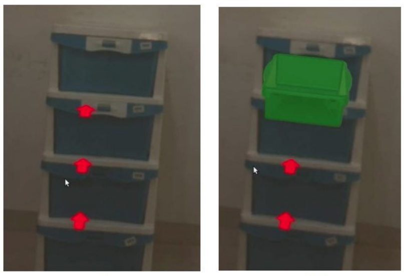


Figure 3. Results from the reconstruction and overlaid virtual objects. a) Tray, object under study, b) Reconstructed object (phong shaded), c) Virtual markers placed over Tray, d) Wireframe of the Tray

object during the operation process and the Fig.4b shows the open state of drawer. This would help the user access the tasks without spending any time in referring to complex manuals. As the instructions are overlaid directly on the object.



(a) (b)

Figure 4. Virtual instructions overlaid on real world objects. a) AR scene with visual markers, b) AR scene with virtual objects

In the Method 2. That we described above, a natural scene with various objects like drill machine, funnel and cans are placed on the surface. Then the RGB and depth images are used to generate a 3D pointcloud, this pointcloud is fed to the deep neural architecture and the results are process to identify the objects. Once objects are identified their poses estimated using the second methodology. Now suppose the user need to drill a hole in a machine which is placed nearby, or needs to place the funnel in an appropriate position, animation will pop up over

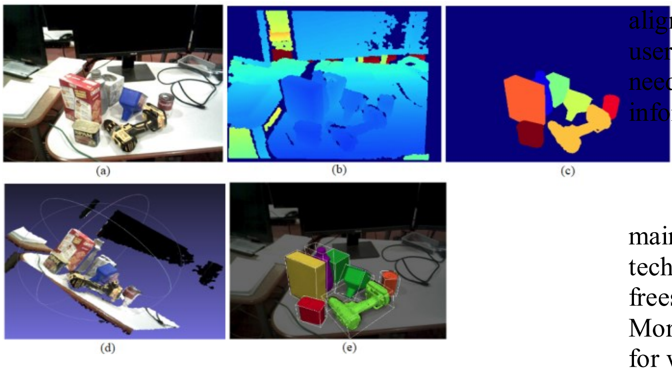


Figure 5. Results obtained using our deep neural architecture. a)RGB Scene, b) Depth scene, c) Objects Segmented from RGB Image, d)Point cloud of the entire scene, e) Projected object models after finding pose

the drill machine guiding the user exactly how to complete the process, the pose models will be running in real time and will aid the user, so that no mistakes whatsoever can occur. Now the drill machine can be put in any scene and a novice can also be aided to do their job properly. This generalized framework can be extended to any other operation. For example, to use this for monitoring device health, the user decides that nature of environment. For example, if it is a room full of machines, that are almost static, the user first scans the complete room where machinery is present. The reconstructed environment map is post processed and loaded into to Unity3d. Then the virtual tags are placed to show various devices status such as temperature, voltage, current, RPM etc. If the environment is a place where objects are mobile and frequently moves from a place to place, we use process described in methodology 2, the virtual markers are then placed over detected objects, the data for these virtual markers is obtained via pre implemented IOT frameworks [24]. The application is loaded into AR headset, then a

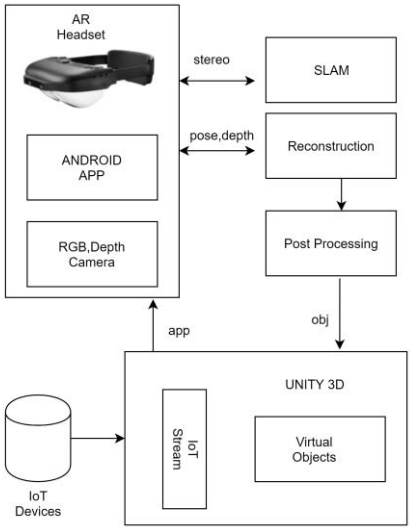


Figure 6 Pipeline to extend the system for Fault analysis

relocalization of coordinate system is performed. This aligns virtual content with real world content and the user will be able to all the data in real time without needing any extra hardware to gather the device information.

V. CONCLUSION

We presented a novel methodology for maintenance to aid the workers of the future. The technique thus presented is ideal for a worker as it frees his hands to do other labour intensive jobs. Moreover, the headset based solutions remains ideal for workers in the assembly lines, where new workers may come in and join, as it helps drastically reduce the time required for up-skilling them. In addition the AI assisted headset is ideal for reducing the chances of failure during assembly. The hybrid methodology for finding pose is ideal for the industrial use case as it is robust to occlusion, change in lighting conditions, and works in real-time on a computationally inexpensive device.

VI. REFERENCES

[1] P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays", In IEICE Transactions on Information Systems, Vol. E77-D, no. 12, Dec. 1994

[2] I. E. Sutherland “The Ultimate Display,” Proceedings of the IFIP Congress, pp. 506–508, 1965

[3] M. Quandt, B. Knoke, C. Gorldt, M. Freitag, and K.-D. Thoben, “General Requirements for Industrial Augmented Reality Applications,” Procedia CIRP, vol. 72, pp. 1130–1135, 2018

[4] Manikandan, R., and S. Smys. "Controller based performance measures of speed control of electrical motor for industrial applications." In *2017 International Conference on Inventive Systems and Control (ICISC)*, pp. 1-5. IEEE, 2017.

[5] J. S. Ortiz, J. S. Sánchez, P. M. Velasco, C. R. Sánchez, W. X. Quevedo, V. D. Zambrano, O. Arteaga, and V. H. Andaluz, “Teaching-Learning Process through VR Applied to Automotive Engineering,” Proceedings of the 2017 9th

International Conference on Education Technology and Computers - ICETC 2017, 2017

[6] S. Rabah, A. Assila, E. Khouri, F. Maier, F. Ababsa, V. Bourny, P. Maier, and F. Mérienne, “Towards improving the future of manufacturing through digital twin and augmented reality technologies,” Procedia Manufacturing, vol. 17, pp. 460–467, 2018

[7] J. Martín-Gutiérrez, “Virtual Technologies Trends in Education,” EURASIA Journal of Mathematics, Science and Technology Education, vol. 13, no. 1, 2016

[8] B. Tekin, S. N. Sinha, and P. Fua, “Real-Time Seamless Single Shot 6D Object Pose Prediction,” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018

[9] A. Nigam, A. Penate-Sanchez, and L. Agapito, “Detect Globally, Label Locally: Learning Accurate 6-DOF Object Pose Estimation by Joint Segmentation and Coordinate Regression,” IEEE Robotics and Automation Letters, vol. 3, no. 4, pp. 3960–3967, 2018

[10] C. Bichlmeier, T. Sielhorst, S. M. Heining, and N. Navab, “Improving Depth Perception in Medical AR,” Bildverarbeitung für die Medizin 2007 Informatik aktuell, pp. 217–221

[11] P. Pantelidis, A. Chorti, I. Papagiouvanni, G. Paparoidamis, C. Drosos, T. Panagiotakopoulos, G. Lales, and M. Sideris, “Virtual and Augmented Reality in Medical Education,” Medical and Surgical Education - Past, Present and Future, 2018

[12] K. Fast, T. Gifford, and R. Yancey, “Virtual Training for Welding,” Third IEEE and ACM International Symposium on Mixed and Augmented Reality

[13] E. P. Yugcha, J. I. Ubilluz, and V. H. Andaluz, “Virtual Training for Industrial Process: Pumping System,” Lecture Notes in Computer Science Augmented Reality, Virtual Reality, and Computer Graphics, pp. 393–409, 2019

[14] X. Fan, R. Yang, D. Wu, and D. Ma, “Virtual Assembly Environment for Product Design Evaluation and Workplace Planning,” Virtual Reality & Augmented Reality in Industry, pp. 147–161, 2011

[15] D. B. Espíndola, C. E. Pereira, R. V. Henriques, and S. S. Botelho, “Using Mixed Reality in the Visualization of Maintenance Processes,” IFAC Proceedings Volumes, vol. 43, no. 3, pp. 30–35, 2010

[16] L. E. Garza, G. Pantoja, P. Ramírez, H. Ramírez, N. Rodríguez, E. González, R. Quintal, and J. A. Pérez, “Augmented Reality Application for the Maintenance of a Flapper Valve of a Fuller-kynion Type M Pump,” Procedia Computer Science, vol. 25, pp. 154–160, 2013

[17] R. Ghimire, K. R. Pattipati, and P. B. Luh, “Fault diagnosis and augmented reality-based troubleshooting of HVAC systems,” 2016 Ieee Autotestcon, 2016

[18] M. Chang, S. Ong, and A. Nee, “AR-guided Product Disassembly for Maintenance and Remanufacturing,” Procedia CIRP, vol. 61, pp. 299–304, 2017

[19] B. K. Seo , H. Wuest, "A Direct Method for Robust Model-Based 3D Object Tracking from a Monocular RGB Image", Computer Vision -- ECCV 2016 Workshops, pp. 551–562

[20] S. Izadi, A. Davison, A. Fitzgibbon, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, and D. Freeman, “KinectFusion,” Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST 11, 2011

[21] S. Peng, Y. Liu, Q. Huang, X. Zhou, and H. Bao. Pvnet: Pixel-wise voting network for 6dof pose estimation. In CVPR, 2019

[22] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, “ORB-SLAM: A Versatile and Accurate Monocular SLAM System,” IEEE Transactions on Robotics, vol. 31, no. 5, pp. 1147–1163, 2015

[23] Y. Feng, Y. Wu, and L. Fan, “Real-time SLAM relocalization with online learning of binary feature indexing,” Machine Vision and Applications, vol. 28, no. 8, pp. 953–963, 2017

[24] A. Abdelgawad and K. Yelamarthi, “Internet of Things (IoT) Platform for Structure Health Monitoring,” Wireless Communications and Mobile Computing, vol. 2017, pp. 1–10, 2017