IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 11, Issue 6, November-December 2024

Impact Factor: 7.394

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

crossref

निस्केयर NISCAIR

International Journal of Advanced Research in Education and TechnologY(IJARETY)

IJARETY

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 6, November-December 2024 ||

DOI:10.15680/IJARETY.2024.1106059

Object Detection in Augmented Reality: Harnessing Camera Vision

G.Lakpathi [1], Sardar Sathi Amar Singh[2], Talari Thrinath[3], Vakala Diran[4]

Assistant Professor, Department of Computer Science & Engineering, Guru Nanak Institute of Technology, India[1]

Student, Department of Computer Science & Engineering, Guru Nanak Institute of Technology, India[2]

Student, Department of Computer Science & Engineering, Guru Nanak Institute of Technology, India[3]

Student, Department of Computer Science & Engineering, Guru Nanak Institute of Technology, India[4]

ABSTRACT: Augmented Reality (AR) has made a breakthrough as a transformative technology with applications ranging from gaming and education to industry and healthcare. This project aims to develop a versatile AR framework using Python inside OpenCV, designed to work with a wide variety of markers and support the placement of graphical elements. The framework's primary functionality will depend on a user-provided markers configuration file, which enables the system to adapt different types of markers. The system will be designed to identify an extensive range of QR codes and custom markers using the OpenCV ArUco model. This ensures compatibility with a diverse set of real-world scenarios. Users can define the type, size, position, and orientation of graphical elements, allowing interactive AR experiences. This framework relies on a user-provided markers configuration file. This file will contain marker specifications and graphic element details. It can provide the output of the distance of markers and marker's Data.

KEYWORDS: Augmented Reality (AR), Object Detection, OpenCV ArUco Model, Marker Configuration, Interactive AR Framework

I. INTRODUCTION

Augmented Reality is a computer technology that adds digital content onto a live camera feed, making that digital content look as if it is part of the physical world around you.AR can provide a view of the real-time data flowing from products and allow users to control it. A marker can be an object or a visual such as logos or QR codes.AR application triggered by a specific physical marker captured by the camera to position the digital content on top of it, is called marker-based AR.OpenCV is the open-source library for computer vision and image processing. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in commercial products. By using it, one can process images and videos to identify objects, faces, or even the handwriting of a human. When it is integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis.

II. LITERATURE SURVEY

D. G. Lowe (1999) An object recognition system has been developed that uses a new class of local image features. The features are invariant to image scaling, translation, and rotation, and partially invariant to illumination changes and affine or 3D projection. These features share properties similar to those of neurons in the inferior temporal cortex, which are used for object recognition in primate vision. Features are efficiently detected through a staged filtering approach that identifies stable points in scale space. Image keys are created that allow for local geometric deformations by representing blurred image gradients in multiple orientation planes and at multiple scales. The keys are used as input to a nearest neighbor indexing method that identifies candidate object matches. Final verification of each match is achieved by finding a low residual least squares solution for the unknown model parameters. Experimental results show that robust object recognition can be achieved in cluttered partially occluded images with a computation time of under 2 seconds.

Takumi, Puerto Rico(2020) We present a new approach for resolving occlusions in augmented reality. The main interest is that it does not require a 3D reconstruction of the considered scene. Our idea is to use a contour-based approach and to label each contour point as being "behind" or "in front of", depending on whether it is in front of or behind the virtual object. This labeling step only requires that the contours can be tracked from frame to frame. A

IJARETY | ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 6, November-December 2024 ||

DOI:10.15680/IJARETY.2024.1106059

proximity graph is then built to group the contours that belong to the same occluding object. Finally, we use some kind of active contours to accurately recover the mask of the occluding object

A. P. Dempster (2020) A broadly applicable algorithm for computing maximum likelihood estimates from incomplete data is presented at various levels of generality. Theory showing the monotone behavior of the likelihood and convergence of the algorithm is derived. Many examples are sketched, including missing value situations.

W. F¨orstner, E. G¨ulch(1987) Due to the different imaging characteristics of the sensor, there are big differences between multisource images in gray and a trend of gray gradient. And the existing algorithms of image registration were time-consuming or low matching. In view of the status quo, a brief review of the SIFT algorithm is first given, and the shortcoming of SIFT, in which the matching rate is vulnerable to influence by the gray feature, is pointed out. Then a fast algorithm for multisource image registration based on geometric features of corners was presented. It adopts the geometric feature of corners rather than the gray feature. So the shortcomings of SIFT can be overcome. The novel algorithm can be used to register multisource images with large differences in gray or with different wavebands and can increase the speed and raise the matching rate of registration. This section focused on how to select the corners, how to calculate feature vectors, and the feature matching algorithm. Finally, experiments have been done to prove that this algorithm can register images quickly and efficiently.

S. M. Smith (1995) This paper presents a corner-based image alignment algorithm based on the procedures of corner-based template matching and geometric parameter estimation. This algorithm consists of two stages:1) the training phase, and 2) the matching phase. In the training phase, a corner detection algorithm is used to extract the corners. These corners are then used to build the pyramid images. In the matching phase, the corners are obtained using the same corner detection algorithm. The similarity measure is then determined by the differences in gradient vector between the corners obtained in the template image and the inspection image, respectively. A parabolic function is further applied to evaluate the geometric relationship between the template and the inspection images. Results show that the corner-based template matching outperforms the original edge-based template matching in efficiency, and both of them are robust against non-linear light changes. The accuracy and precision of the corner-based image alignment are competitive to that of edge-based image alignment under the same environment. In practice, the proposed algorithm demonstrates its precision, efficiency, and robustness in image.

III. METHODOLOGY

MODULES NAME

Modules Name:

- Data collection
- Data Preparation
- Data Distance Detection
- Marker Data Show

MODULES EXPLANATION

1) Data collection& preparation:

To create a marker-based AR experience using OpenCV we need to use ArUco marker.

2) Data Distance Detection

The video frame is analyzed to find square shapes that are candidates to be markers.After the candidates are detected

3) Marker Data Show:

Here we are exposing the marker inside data and what it can be stored by the marker

EXISTING SYSTEM DISADVANTAGES

- **Data Dependency**: CNNs require large volumes of labeled training data.
- **Computational Intensity**: Training and running CNNs demand significant computational resources
- **Overfitting**: CNNs are prone to overfitting when dealing with small datasets, causing reduced generalization to unseen data.
- **Interpretability**: CNNs are often considered "black-box" models, making understanding and interpreting their decision-making processes challenging.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 6, November-December 2024 ||

DOI:10.15680/IJARETY.2024.1106059

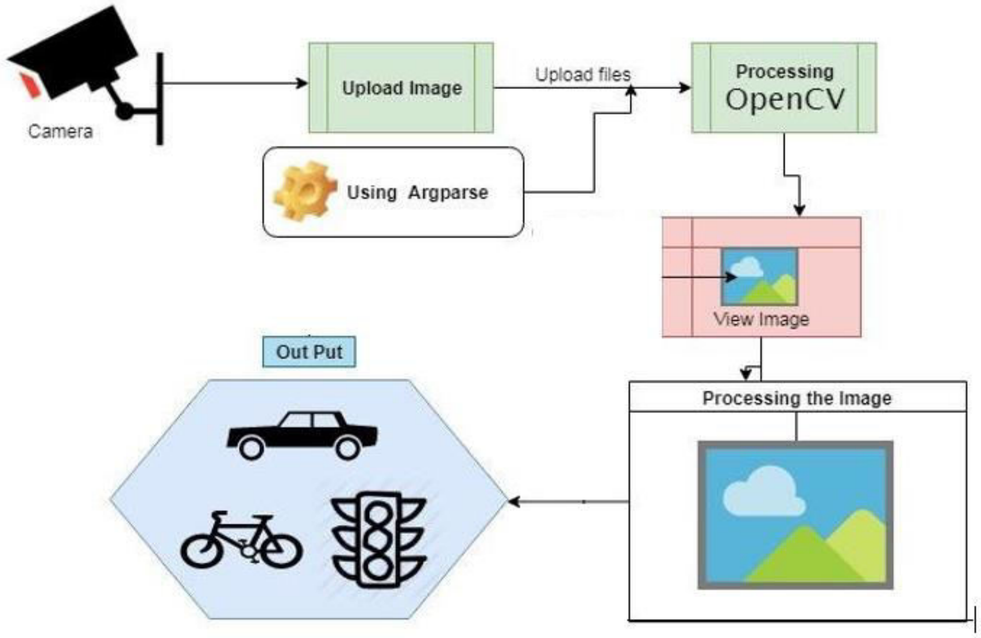
IV. PROPOSED SYSTEM

The ArUco marker is a fiducial square marker used for camera pose estimation. When an ArUco marker is detected in the video, you can augment the digital content with an image on top of the detected markers. ArUco marker is a synthetic square marker with an internal binary matrix enclosed within a wide black color border with a Unique Identifier. In the ArUco marker, black denotes one, and white represents zero. The marker size determines the size of the internal binary matrix. The odd blocks in the ArUco marker represent parity bits, and even squares in the marker represent data bits. The black border facilitates fast detection within the image.

PROPOSED SYSTEM ADVANTAGES

- Highly Accurate: ArUco markers provide high-precision pose estimation,
- Marker Pose Estimation: ArUco markers are used to estimate the pose (position and orientation) of objects in the camera view.
- Versatile Applications: ArUco markers find applications in robotics, augmented reality, virtual reality, and computer vision.
- Wide Support: ArUco markers are supported by various libraries and frameworks, making it easier to integrate them into software.

V. SYSTEM ARCHITECTURE



EXPLANATION

The system architecture consists of key components that work together to deliver augmented reality functionality. The User Interface (UI) enables users to upload marker configuration files and visualize content. The Application Layer implements marker detection and overlays using OpenCV and ArUco, while the Data Layer manages the storage of marker configurations and output data. The Integration Layer connects with external devices like cameras, and Security Measures protect user data from unauthorized access.

The Infrastructure Layer provides essential hardware and software resources. Additionally, Monitoring and Tolerance and High Availability strategies minimize downtime, supported by Deployment Tools for efficient updates.

IMPLEMENTATION

Input:

- **Source**: Live video feed from the camera (cv.VideoCapture(0)) or pre-stored images from a directory.
- **Input Types**: ArUco markers, checkerboard patterns, or plain frames.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 6, November-December 2024 ||

DOI:10.15680/IJARETY.2024.1106059

Processing:
- **Image Preprocessing**: Converts frames to grayscale for marker/pattern detection.
- **Feature Detection**:
- Detects ArUco markers using OpenCV's aruco.detectMarkers.
- Identifies checkerboard patterns for calibration.
- **Homography**: Applies a homography transformation to overlay source images onto marker-detected regions.
- **Calibration**: Computes intrinsic/extrinsic parameters using checkerboard images.

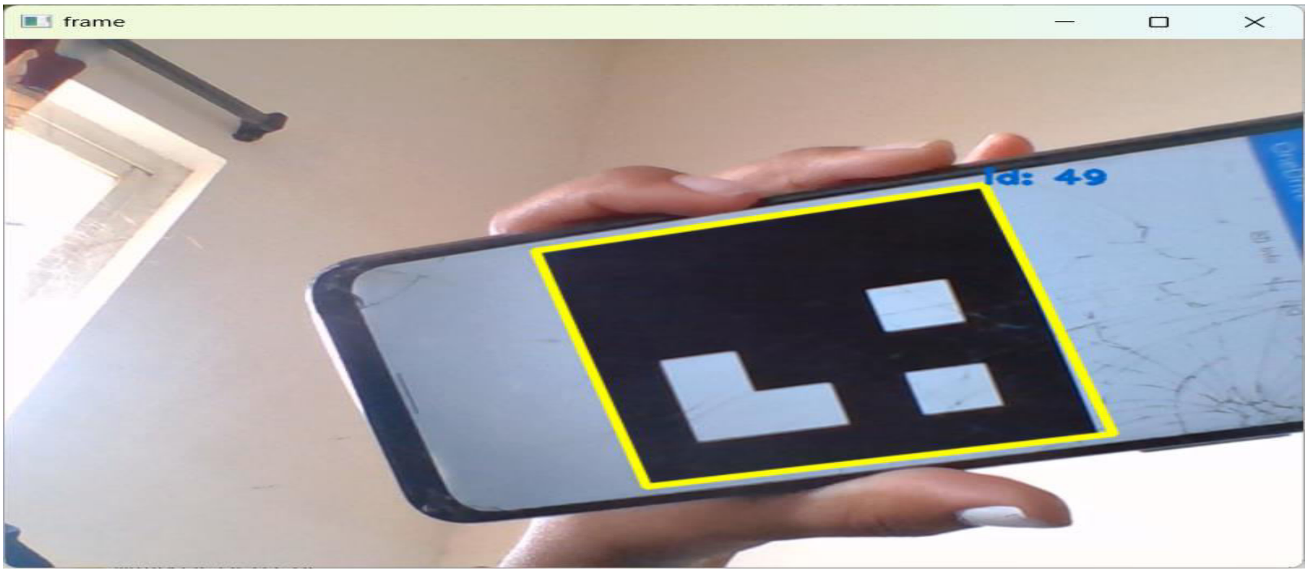
Augmentation:
- Maps an input image onto detected ArUco marker positions using homography.
- Highlights detected markers or patterns with borders and displays IDs or points.

Output:
- **Real-Time Visualization**: Displays augmented frames or detected features on the screen.
- **Data Storage**
- Saves calibration data (camera matrix, distortion coefficients) in .npz format.
- Stores generated marker images in the specified directory.
- Captures and saves frames with detected checkerboard patterns.

Termination:
- Ends the program when the user presses a specific key (e.g., 'q' for quit, 's' for saving images)

VI. EXPERIMENTAL RESULTS

RESULTS



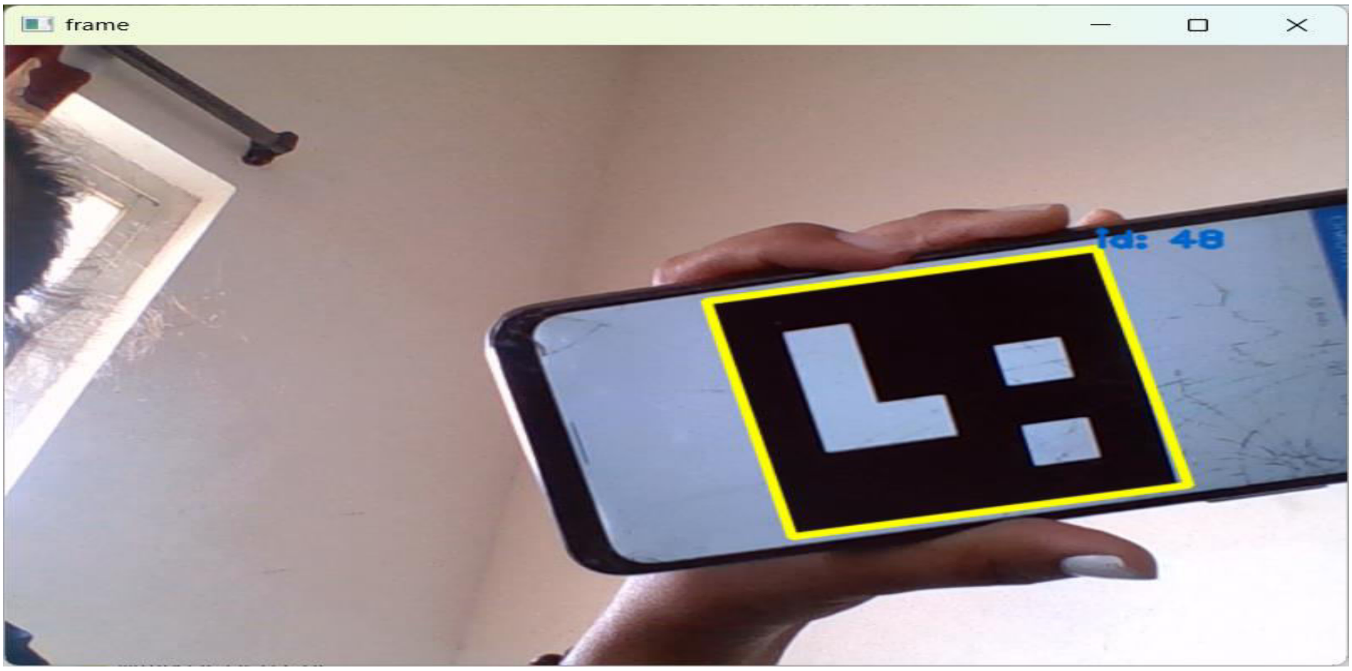
Detecting IDs of the respective markers

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 6, November-December 2024 ||

DOI:10.15680/IJARETY.2024.1106059

Detecting Image from the markers

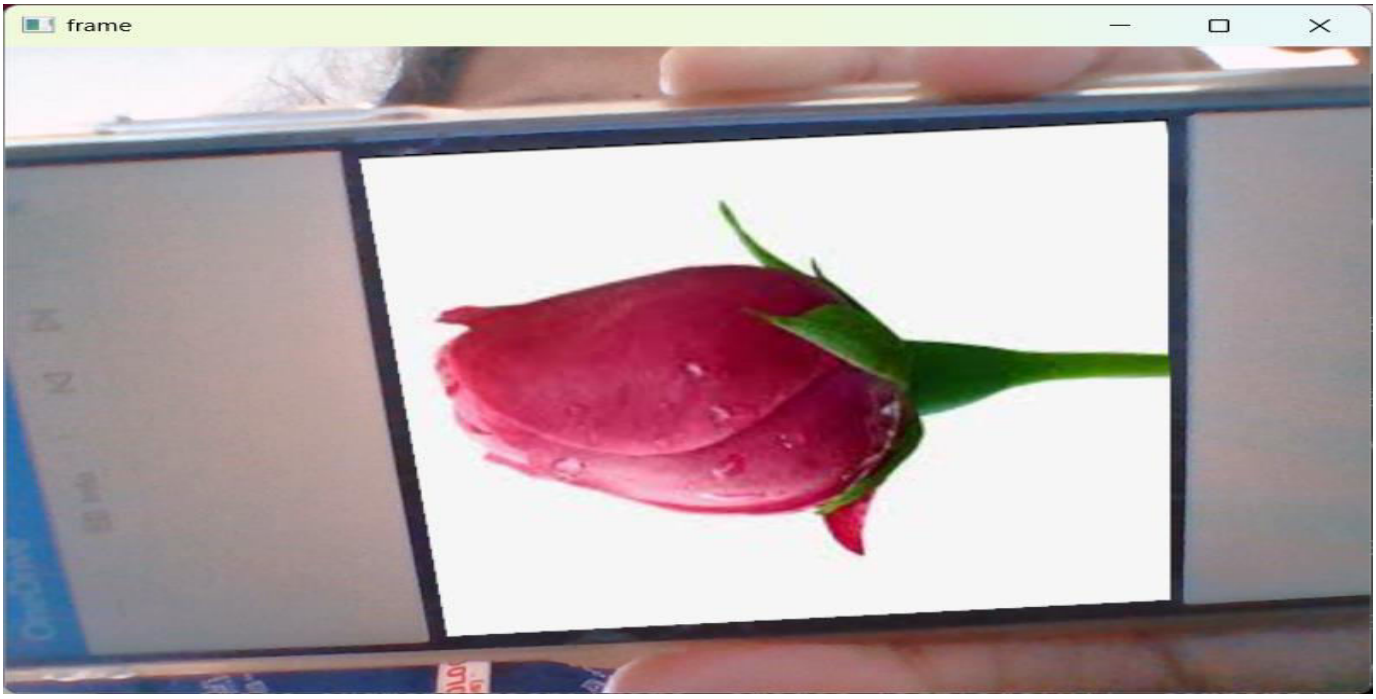


Detecting ID's of the respective markers

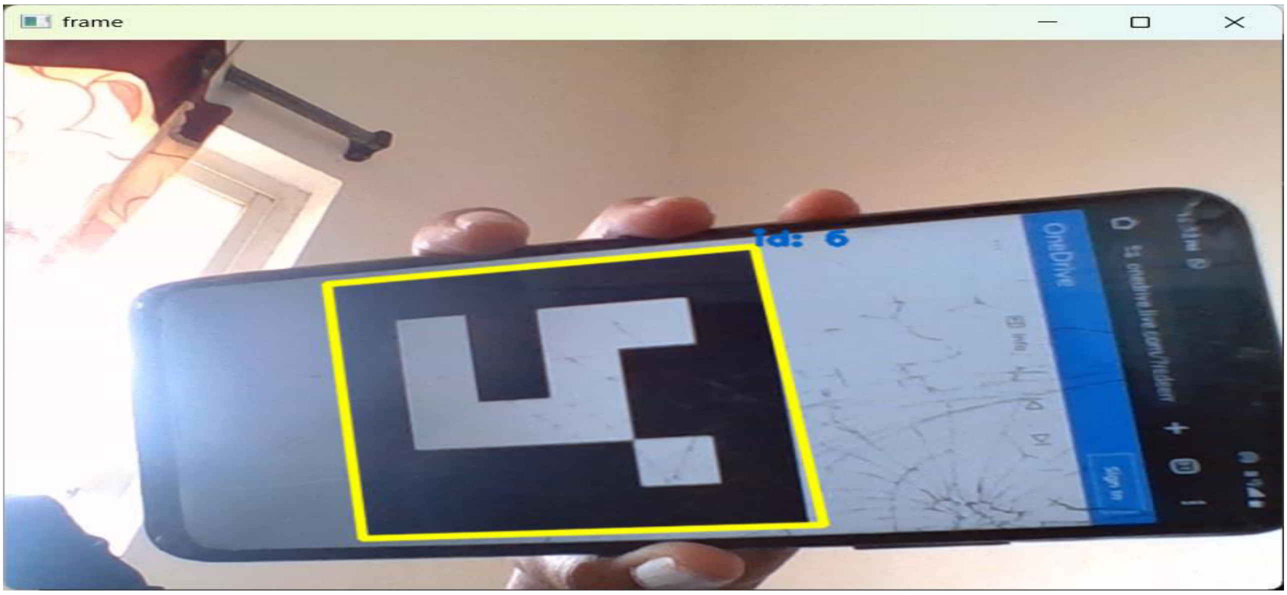
| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 6, November-December 2024 ||

DOI:10.15680/IJARETY.2024.1106059



Detecting Image from the markers

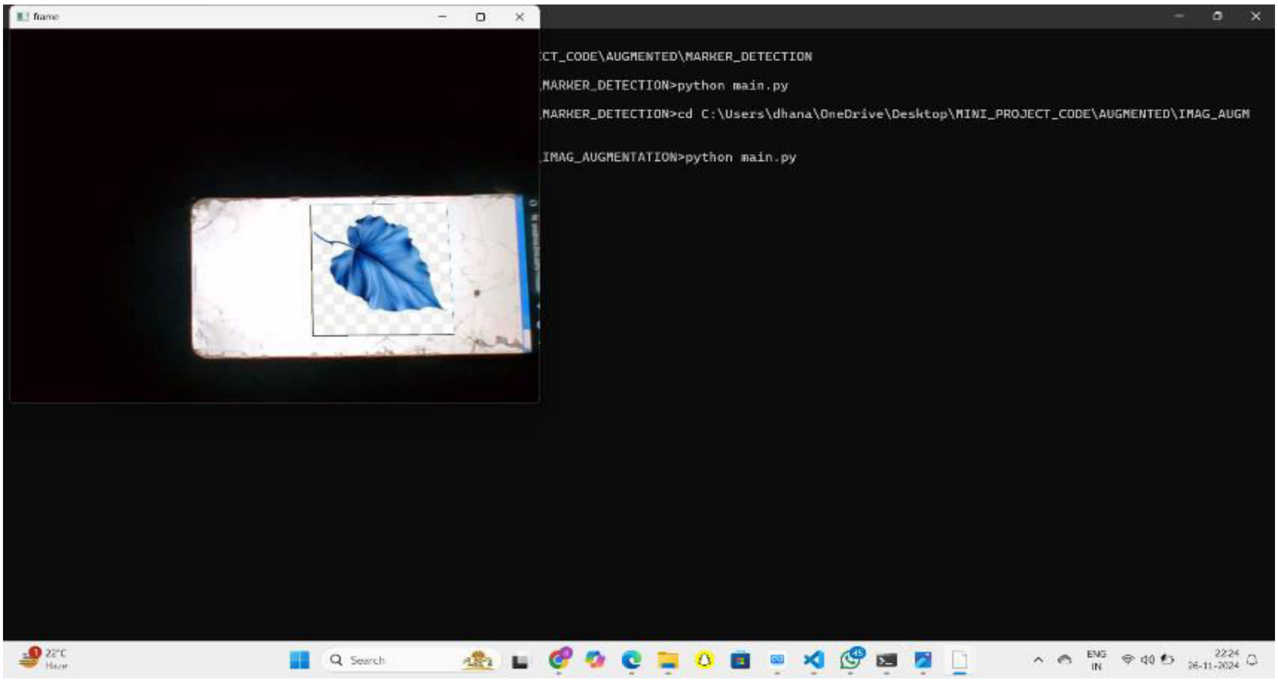


Detecting ID's of the respective markers

IJARETY | ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 6, November-December 2024 ||

DOI:10.15680/IJARETY.2024.1106059



Detecting Image from the markers

VII. CONCLUSION

The title "MARK-GP (Marker-Based Augmented Reality Framework with Graphical Placement)" signifies an innovative and promising development in the realm of augmented reality. This framework likely represents a notable advancement in the field, incorporating marker-based technology for precise object anchoring and sophisticated tools for graphical placement within the AR environment. Such a project suggests a keen focus on delivering a seamless and immersive user experience. The framework's potential applications span diverse domains, from gaming and education to marketing and industrial uses, holding the promise of impacting the AR industry significantly.

VIII. FUTURE ENHANCEMENTS

MARK-GP hints at interdisciplinary collaboration, reflecting the integration of cutting-edge technologies and the opening of new avenues for AR research and development. This work may address usability and accessibility concerns, aiming to ensure that AR technology becomes more user-friendly and widely accessible. In sum, MARK-GP appears poised to leave a substantial imprint on the augmented reality landscape, offering an exciting glimpse into the future of AR innovation.

REFERENCES

[1]R. T. Azuma,A survey of augmented reality, Presence 6 (1997) 355–385.
[2]H. Kato, M. Billinghurst, Marker tracking and HMDcalibration for a Video-Based augmented reality conferencing system, Augmented Reality, InternationalWorkshop on 0 (1999) 85–94.
[3]V. Lepetit, P. Fua, Monocular model-based 3d tracking of rigid objects:A survey, in:FoundationsandTrends in Computer Graphics and Vision, 2005.
[4] B. Williams, M. Cummins, J. Neira,P.Newman,I. Reid, J. Tard´os, A comparison of loop closing techniques in monocular slam, Robotics and Autonomous Systems.
[5] W. Daniel, R. Gerhard,M.Alessandro,T. Drummond, S. Dieter, Real-time detection and tracking for augmented reality on mobile phones, IEEE Transactions on Visualization and Computer Graphics (2010) 355–368.
[6]G.Klein,D. Murray, Parallel tracking and mapping for small ar workspaces, in:Proceedings of the2007 6th IEEE and ACM International Symposium12on Mixed and Augmented Reality, ISMAR ’07, IEEEComputer Society, Washington, DC, USA, 2007, pp.1–10.
[7]K. Mikolajczyk, C. Schmid, Indexing based on scale invariant interest points., in:ICCV, 2001, pp. 525–531.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 7.394| A Bi-Monthly, Double-Blind Peer Reviewed & Referred Journal |

|| Volume 11, Issue 6, November-December 2024 ||

DOI:10.15680/IJARETY.2024.1106059

[8]D. G. Lowe, Object recognition from local scale invariant features, in:Proceedings of the International Conference on Computer Vision-Volume 2Volume 2, ICCV '99, IEEE Computer Society, Washington, DC, USA, 1999, pp. 1150.
[9]P. Bhattacharya, M. Gavrilova, A survey of landmark recognition using the bag-of-words framework, in:InTelligent Computer Graphics, Vol. 441 of Studies Computational Intelligence, Springer Berlin Heidelberg, 2013, pp. 243–263.
[10]H. Kato, M. Billinghurst, Marker tracking and hmdcalibration for a video-based augmented reality conferencing system, in:Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, IWAR '99,IEEE Computer Society, Washington, DC, USA, 1999, pp. 85.
[11]M. Fiala, Designing highly reliable fiducial markers,IEEE Trans. Pattern Anal. Mach. Intell. 32 (7) (2010)1317–1324.
[12]D.Schmalstieg,A. Fuhrmann,G. Hesina,Z. Szalav´ari,L. M. Encarna¸c¨ao,M. Gervautz,W. Purgathofer, The studierstube augmented realityproject, Presence:Teleoper. Virtual Environ. 11 (1)(2002) 33–54.
[13]R. Munoz-Salinas,S. Garrido-Jurado,Aruco library, [Online; accessed 01-December-2013] (2013).
[14] K. Dorfmller, H. Wirth,Real-time hand and headtracking for virtual environments using infraredbeacons,in:in Proceedings CAPTECH98. 1998, Springer, 1998, pp. 113–127.
[15] M. Ribo, A. Pinz, A. L. Fuhrmann, A new opticaltracking system for virtual and augmented reality applications, in:In Proceedings of the IEEE Instrumentation and Measurement Technical Conference, 2001,pp.

IJARETY

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 7.394

www.ijarety.in editor.ijarety@gmail.com