# Designing a Real-time Interactive Spatial Augmented Reality Platform

Aye Chan Zay Hta[1],2*, YunLi Lee[1],2 and Wai Chong Chia[1],2
[1] Department of Computing and Information Systems, School of Engineering and Technology,
[2] Research Centre for Human-Machine Collaboration (HUMAC),
Sunway University, Bandar Sunway, Selangor, Malaysia
* Corresponding author E-mail: ayechan.aczh@gmail.com

(Received 11 July 2021; Final version received 22 July 2021; Accepted 7 September 2021.)

**Abstract**

The Spatial Augmented Reality (SAR) system is an interactive platform that allows any virtual contents to be projected onto any physical structure (object). However, most of the SAR platforms use pre-rendered models in a static setting that does not allow the physical structure to be modified dynamically. This has limited interactions between the users and the system because users are unable to make rapid changes to the physical structure to express their ideas. This limitation can be a huge challenge when it comes to application such as city planning, where rapid real-time prototyping is able to provide a better visualization to the impacts that the changes could bring to the city environment. Therefore, this research project aims to design and develop a tracker-based SAR system to resolve the aforementioned limitation. The main contributions of this research project include, (i) a SAR system that supports real-time physical structure reconstruction and projection mapping, and (ii) a SAR platform constructed using Lego blocks and easily accessible hardware and software. The hardware involves the design of the physical set-up to support a real time reconstruction, and dynamic projection mapping. The software involves real time object detection, tracking, and projection mapping. Real time object detection is carried out using colour tracking, and recording Lego positions, while dynamic projection mapping is done through marker tracking and coordinate mapping. Based on preliminary evaluations conducted in the laboratory, the experimental results shown that the proposed SAR system is able to (i) successfully project virtual content onto physical structure built using Lego blocks in real-time, and (ii) detect changes made to the physical structure.

Keywords: Dynamic Reconstruction, Lego blocks, Real-time marker assignment, Spatial Augmented Reality (SAR), Tracker-based SAR

## 1. Introduction

Digital industry predicts that augmented reality (AR)/virtual reality (VR) would continuously grow and the revenue would increase more than twenty-five billion in the next five years (Makarov, 2021). The future applications of these AR/VR technologies are taken place in various sectors such as business, marketing, education, navigation, health, and others. AR constitutes the integration of virtual resources together with real world physical elements, in which computer-generated graphical components are displayed in the user's digital devices along with the elements of real environment. Milgram and Kishino (1994) explained the operational definition of AR by stating the term that describes any case in which the real environment is "augmented' virtually by computer graphics. The mix reality environment is in between the spectrum of extremes of real and virtual worlds, where the user can interact with both real and virtual objects which are presented at the same display as shown in Fig. 1.

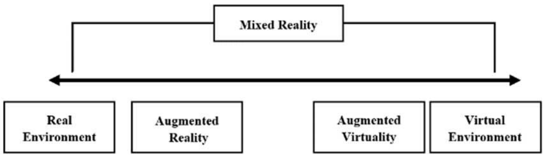



**Fig. 1.** Virtuality Continuum (Milgram and Kishino, 1994)

Furthermore, with current technologies, limitations in the user's field of view and the ergonomics of wearable AR devices are still challenging. With projection mapping techniques, we can project out these virtual contents onto the real world allowing more seamless blend between the virtual and real world. However, most of the projection mapping systems are static setting without allowing for dynamic reconstruction of physical structure, thus limitations in user interactions (Cortes et al., 2018).

There is a need of interactive tangible user interfaces/platform to allow better data visualization and more innate interactions. With Spatial Augmented Reality (SAR) system, the users can view a larger field in real-time and experience immersive interactions among multiple users. SAR used a projection technology to display the surfaces of a variety of objects with video projection (Ball, 2018; Park et al., 2014). It provides many opportunities for display of events including live shows, museums, exhibitions, conferences, trainings, and designing of products by using audio, video, projectors and software. The audience can appreciate the effect of a combination of audio, video together with 3D modality well beyond the traditional ways (Ball, 2018). SAR allows evaluation of the products by users ahead of physical development of a prototype, thus saving time and cost related to the development of the product (Ball, 2018). Park et al. (2014) mentioned that SAR is useful to show the virtual products which are similar to the real ones without limitation of space issues. From their evaluation of SAR designs, it shows that SAR provides more flexible and intuitive environment with high sense of immersion than using digital display. However, they pointed out that SAR required a more complex set-up of the hardware compared to traditional computer-aided colour design and the immersive experience depends on the projector's performance as the resolution and lumens of the projector are key factors.

### 1.1 Dynamic Reconstruction SAR Platform

Most SAR platforms were using pre-rendered models and they were used in a static setting without allowing for dynamic manipulation of physical setting. These SAR platforms rely on finger tracking and touch gestures to allow for user interaction. These techniques are viable for 2D projections and touch screen applications. However, in 3D projections the use of movable tangible objects can improve the user's interaction as it allows for a more natural user interaction. A study conducted by Al-Megren and Ruddle (2016) which compares tangible interaction with multi touch interaction showed that the time required to complete tasks were faster as well as less errors occurring in tangible interaction. Those that did allow for dynamic manipulation use markers which limit the virtual projections that are to be mapped and require the use of pre-rendered models, such as in the case of Winder and Larson (2017) which supports 16 types of different markers, all of which are pre-rendered and assigned to the specific marker.

Further literature review was conducted to find SAR platforms that support real-time dynamic reconstruction. Kim et al. (2014) and Guo et al. (2018) support a basic version of dynamic reconstruction by using depth sensors such as the Microsoft Kinect. Based on the depth information, the virtual content responded on to the physical setting. However, they still have limitation in their platforms as they only use the vertical depth information to control the change of virtual content, and therefore their applications cannot detect the actual shape of the object.

There are several issues to be considered such as most SAR platforms do not offer dynamic manipulation of physical setting to the system as it requires object detection and dynamic projection mapping. Additionally, others only used pre-rendered models in a static setting limiting the types of physical structure to pre-determined shapes.

A real-time interactive SAR platform is required to address this problem. In this project, dynamic reconstruction is introduced in SAR system that allows users to manipulate the physical shape of the tangible object that they are interacting with, in real time. Thus, the proposed solution offers an additional layer of user interaction, which overcomes the pre-rendered models and pre-determined shapes in the SAR platform.

### 1.2 Aims & Objectives

This research aims to develop a general-purpose SAR platform with an additional layer of user interaction through real-time reconstruction, by using a simple, cost effective and easily accessible hardware.

The objectives of the current research project are as follow.

(1) to identify state-of-the-art tracking and calibration techniques that should be used in developing the SAR platform.

(2) to perform real-time projection mapping and 3D reconstruction for objects constructed using Lego blocks.

(3) to track the objects constructed using Lego blocks in real-time using image processing.

(4) to evaluate the usability and generalization of the SAR platform based on feedback /ratings given by end users.

### 1.3 Project Scope

The proposed real-time interactive SAR platform includes the designing and developing of a tracker-based real-time SAR system and a physical set-up of SAR platform using Lego blocks. Lego blocks are popular and common, therefore supporting the objective of using easily accessible hardware. Furthermore, Lego block can be easily deconstructed and reconstructed by the user allowing for more possibilities in terms of shapes. In this project, only standard 2x2, 2x3 and 2x4 Lego blocks are used. This SAR platform supports the object detection, marker detection and tracking and correct projections mapped from the physical structure to offer a more natural and organic way of user interaction that reduces the barrier between the virtual and the real world.

## 2. Literature review

A literature review was performed before starting the project in order to identify the existing or similar work on the SAR platforms. This gave a variety of projects with different designs of framework and highlights of issues and challenges that are relevant to SAR platforms. SAR applications require physical set-up design and software architecture design. Most physical designs involve of top, bottom, or even side camera position as well as front projections or rear projections. Common framework design for software architecture were, some form of object detection and tracking which is either marker based (Winder and Larson, 2017; Mousavi et al., 2013; Laviole, 2012) or marker-less (Kim et al., 2014; Guo et al., 2018; Wilson, 2005; Park, 2017), interactive features either direct interaction with the physical content (Winder and Larson, 2017; Kim et al., 2014; Guo et al., 2018; Mousavi et al., 2013; Mousavi et al., 2013; Wilson, 2005) or indirectly (via mobile devices) (Mendes et al., 2019) and projections mapped onto 2D (Mousavi et al., 2013; Laviole, 2012) or 3D objects (Winder and Larson, 2017; Kim et al., 2014; Guo et al., 2018; Laviole, 2012). In addition, the framework of common SAR application designs involves of camera and projector calibration (Fleischmann and Koch, 2016) to support for dynamic projections and to allow accurate projection mapping onto 3D objects. SAR framework design involves of object detection and/or hand/finger detection (Mousavi et al., 2013; Laviole, 2012; Wilson, 2005) to support interactivity. Most SAR platforms are also found to be not portable and some are designing to make it more portable but still these wearable SAR platforms can be bulky, cumbersome and unergonomic to use for extended periods of time.

Furthermore, additional literature review was conducted with a main focus on interactive techniques used as well as the use of dynamic reconstruction in SAR platforms. It was found that most 2D projection-based SAR platforms include virtual buttons and hand gestures as their interactive features. Whereas, most 3D projection-based SAR platforms mainly used tangible objects and some were using real-time reconstruction as their interactive features. Additionally, some 3D projection-based SAR platforms also included an external display which provided more detailed information that supplemented the platform. The common techniques used in existing SAR platforms such as marker-based tracking, depth sensing, free-form tracking, and real-time reconstruction are compared and described in Table 1.

Most SAR platforms prefer to offer direct forms of interaction as it is a main advantage of SAR compared to traditional AR, where users cannot interact with the physical object/world directly and are required to do so through a secondary device such as a touch screen or mobile device. Table 2 shows the comparison of various features of the existing SAR platforms such as whether they support 3D or 2D projections, usage of external display, animation, support of video or audio. In addition, the interactive features like tangible objects, virtual buttons, hand gestures and real-time reconstruction are also compared in Table 2. This table compares and highlights the limitations of each existing work, further confirms the advantages of a system which allows the users to manipulate the physical shape of the tangible object that they are interacting with, in real-time. Furthermore, interaction with tangible objects can provide a more natural form of interaction.

These common techniques and interactive features showed in Table 1 and 2, are useful as a benchmark to consider the required features in the proposed project.

**Table 1.** Comparison of techniques used in existing SAR platforms.

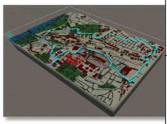
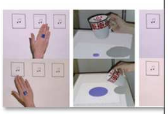
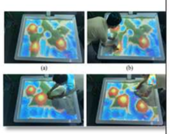
| Author | Image | Techniques used | | | |
|---|---|---|---|---|---|
| | | Marker based | Depth sensing | Free-form tracking | Real-time reconstruction |
| Winder & Larson (2017) |  | ✓ (Support for 16 types of different markers) | ✗ (No use of depth sensor, uses marker id to determine which content to project) | ✗ (No free-form tracking using slot by grid basis) | ✗ (Does not support for different shapes that are not predetermined by the marker) |
| Mendes et al (2019) |  | ✗ | ✗ | ✗ | ✗ |
| | | No moving parts, interaction is done through mobile app | | | |
| Park (2017) |  | N/A (Mostly likely static setting with dynamic contents being projected) | ✓ (Using an IR camera to gather depth information on pre-defined 3D objects) | ✓ (Using OpenCV image processing library to track object) | ✗ (No support for building of Lego blocks or changing of new shapes) |
| Kim et.al (2014) |  | ✗ | ✓ (Using Kinect to gather 3D depth information | ✓ | ✓ (Allows users to rearrange the square blocks but limited to vertical depth and not actual shape) |
| Mousavi et al. (2013) |  | ✓ (Uses colour markers on user's fingers for tracking purposes) | ✗ (Plans to use Kinect in future works but not yet implemented) | ✓ (Supports free-form tracking of hand based on colour markers) | ✗ |
| Laviole (2012) |  | ✓ (Uses markers on page sides for tracking purposes) | ✓ (Uses Kinect sensor for depth sensing but mainly used for gesture and finger tracking) | ✓ (Use of AR markers supports for free-form tracking) | ✗ |
| Gou et al. (2018) |  | ✗ | ✓ (Uses Kinect v2 to gather depth information as the sand is build up) | ✗ (No tracking needed as projections are changed through depth information) | ✓ (Supports real-time reconstruction of the sand box in any shape but limited to vertical depth and not actual shape) |

**Table 2.** Comparison of various interactive features in existing SAR platforms.

| Author | 3D projections | 2D projections | External display | Animation | Video | Audio | Interactivity | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | Tangible objects | Virtual Buttons | Hand Gestures | Real-time reconstruction |
| Winder & Larson (2017) | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ |
| Wilson (2005) | ✗ | ✓ | ✗ | ✗ | ✓ | ✓ | ✗ | ✓ | ✓ | ✗ |
| Mendes et al (2019) | ✓ | ✗ | ✓ | ✓ | ✗ | ✗ | ✓ | ✗ | ✗ | ✗ |
| Park (2017) | ✓ | ✗ | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ |
| Kim et al (2014) | ✓ | ✗ | ✗ | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ |
| Mousavi et al (2013) | ✗ | ✓ | ✗ | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ | ✗ |
| Guo et al (2018) | ✓ | ✗ | ✗ | ✓ | ✗ | ✗ | ✓ | ✗ | ✗ | ✓ |
| Laviole (2012) | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ | ✗ |
| Proposed Project | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ |

## 3. Methodology

In this section, an overview of the proposed architecture is first described. This includes the set-up of the proposed SAR platform by using low-cost and easily accessible hardware. This is followed by the description of the proposed framework, which highlights the techniques used to achieve real-time 3D reconstruction. However, there are several ways (or algorithms) to implement this module. Therefore, evaluation metrics were set-up to evaluate several potential techniques before choosing the best in slot for the proposed framework. The actual implementation of the proposed framework is discussed in more details in Section 4.

### 3.1 Architecture of Real-Time Interactive SAR Platform

The proposed real-time interactive SAR platform uses markers for tracking after object detection is completed together with depth sensor to detect the shape of an object. It supports free-form tracking through the marker without restricting the user to a grid and slot basis which was employed in Bits and Bricks (Winder and Larson, 2017). Real-time reconstruction is also supported for the detection of actual shape of an object to allow users to build up Lego blocks as they desire. This makes the project innovative in providing the additional feature of real-time reconstruction for actual shape which was not presented in earlier literatures. Those existing works mainly consisted of pre-assigned shapes and even those that supported real-time reconstruction only allowed limited interactions such as in Kim et al. (2014) and Guo et al. (2018) where the virtual content changes only based on the depth axis and thus do not support for more complex shapes.

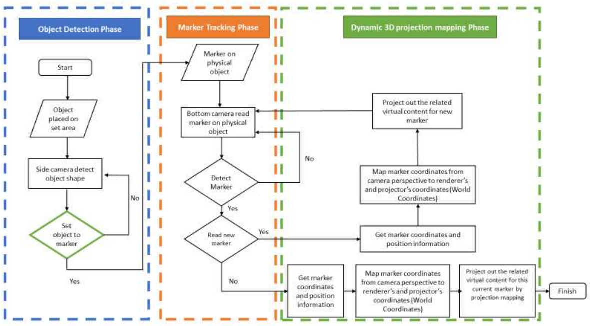



**Fig. 1.** Three phases of the proposed real-time interactive SAR platform.

The system design involves three phases as described in Fig. 2.

Phase 1 includes object detection in real-time with the 3D reconstruction that allows the user to create a new object by building Lego blocks and then assign it to a marker in real-time.

Phase 2 consists of detection and tracking the assigned marker in real-time that allows for interaction for user.

Phase 3 consists of dynamic 3D projection mapping which projects out the relevant virtual content and maps it onto the 3D object. This will ensure that the projections will follow the 3D physical object when it is placed in a new position on the platform.

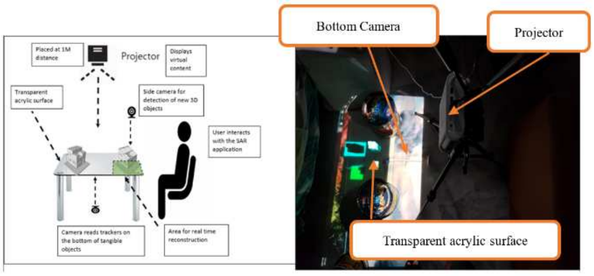



**Fig. 3.** The design of the Physical set-up

To support the SAR platform, physical set-up is designed as shown in Fig. 3. The main hardware resources required are projector (Epson EB-W06), cameras (Acer webcams, Ezviz C1C). These hardware are easier to get and low-cost when compared with more advanced devices. A cheaper projector can also be used at the cost of brightness and resolution of the projection. However, it is recommended to use a projector of 1500 lumens or more to ensure brighter projections. Alternatively, the room can be made darker to make the projections more visible. The projector is used for projection mapping and the camera placed on the bottom of the transparent acrylic surface is used to track markers which are placed on the bottom of the Lego blocks. These markers can be the Lego base's patterns or Fiducial trackers placed onto the bottom of Lego bases. An area is also defined to support for real-time reconstruction where a camera is placed. The users are allowed to change the shape and build the Lego block in this region. By reconstructing the Lego block in this region, it supports coverage by the cameras and reduces the obstructions that may occur. In the next section, the technique adopted by the SAR system to complete the process mentioned in phase 1 to 3 is described.

### 3.2 Framework for 3D reconstruction

The 3D reconstruction framework involves of the detection of individual Lego block and building it up as described below:

*Input: Side camera view image*

*Output: Reconstruction of Lego 3D model*

*Algorithm :*

*1. Detect individual Lego block in the $x$ (defined) region.*

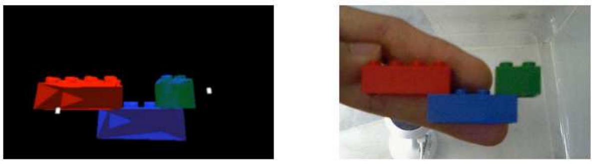
*2. Track the Lego block and record the placements when joint with another block.*

*3. The area for reconstruction is defined as mentioned in the physical set-up to improve the tracking capability and limit camera view obstructions.*

*4. Assign model to tracker base in real-time.*

*5. Apply undistort function using intrinsic parameters and map the coordinates to switch to bottom camera.*

The individual Lego block is continuously tracked using colour tracking, and their placements are recorded each time they are joined with another block. Trigger areas are assigned on each possible Lego slot (using HitTest VVVV function) to determine where they should join. If the joining Lego block's center touches one of the trigger areas the function with return a Boolean true data type and the index of trigger area will allow it to lock onto that respective Lego slot. This will allow for the support for real-time reconstruction of the 3D model. However, difficulties may arise in the ability to continuously track the Lego blocks. Therefore, to ease the tracking capability the area for reconstruction is defined as mentioned in the physical set-up. Fig. 4 shows the 3D reconstructed virtual model on the left and the real object on the right as seen from the webcam. Once the user has finished building the Lego shape, he/she desires it is then assigned to a marker base so that the tracking can be carried out from the bottom camera. This is done to prevent the projections from interfering with the colour tracker from the top camera.



(a) Digital reconstruction by the system (b) Image from webcam feed

**Fig. 4.** Current tests of real-time reconstruction

Based on the description given above, several techniques are required to build the proposed framework, and in general they can be divided into three phases (i) trackers, (ii) 3D reconstruction, and (iii) projection mapping. Moreover, there are several ways (or algorithms) to implement the modules mentioned above.

Hence, some preliminary evaluations were carried out to determine which way (or algorithm) works best for each phase. The evaluation metrics as well as the evaluation results are presented in the next section.

### 3.3 Measurement techniques or evaluation

Different system tests are required to validate the prototype. These include separate tests for object detection, marker assignment and tracking, dynamic 3D projection mapping and integration which are described in the following Table 3. Additionally, user testing is proposed to evaluate the effectiveness of the system from different range of users. It involves the observation techniques to capture the various strategies and approaches that users may take when performing basic interactive tasks with the proposed system.

**Table 3.** System testing and user evaluation.

| System & Usability Tests | Relevant objectives | Description |
|---|---|---|
| System test for object detection | 1, 2 | To test the time required for the system to correctly detect and identify individual Lego blocks and the final shape built. As the objective for this reconstruction system is to be in real-time minimal latency is required in this process. |
| System test for marker assignment and tracking | 1, 2, 3 | To test the accuracy of the marker tracking as well as the identification of each marker which has been assigned to its respective object. |
| System test for dynamic 3D projection mapping | 1, 2 | To test the accuracy of the projection mapping in relation to the coordinates retrieved by the marker tracking. This test will be performed after calibrating the system. |
| System test for integration | 1, 2, 3 | To ensure that the entire system is well integrated and that each component works and flows correctly with other components of the system. |
| Usability test for end user behaviour | 4 | To be conducted on non-expert users (n=5) to demonstrate the real-world applications. The users will be asked to perform basic interactive tasks such as building basic Lego shapes and moving markers on the platform. |

## 4. Implementation and analysis

The experiments and analysis were carried out to investigate the practicality of the selected techniques. The system is capable of performing colour detection, tracking, real-time real construction of shapes involving 2x2, 2x3, 2x4 bricks. Based on the coordinates retrieved from colour detection and tracking, it is able to perform projection mapping in real-time as well. However, there are still a few issues that need to be resolved which will be further explained in the following section.

### 4.1 Resources for implementation

The required hardware and software resources were selected in accordance to support, for the objective of using easily accessible hardware and software to develop the proposed SAR platform as described in Table 4. Most hardware resources stated are easily accessible such as cameras, projector and physical Lego blocks. In addition, all of the software used are free and open-source programs.

**Table 4.** The required resources for development of SAR platform

| Hardware | Software |
|---|---|
| Cameras (Acer webcams)/IR camera (Ezviz C1C) | VVVV (2021) |
| Projector (Epson EB-W06) | OpenCV (2021) |
| Physical Lego objects | Blender (2020) |

VVVV (platform) (VVVV group, 2021) is used for development as it is a real time interactive live programming environment/toolkit. The real-time aspect of VVVV helps greatly in producing the prototype as changes made in code can be seen in real-time compared to other traditional programming languages where it requires building and compiling. The Open Source Computer Vision (OpenCV) library (2021) is selected for image processing. As the image processing required in this platform involves of object detection and marker tracking. Furthermore, image processing is used in calibration of camera and projector to support for dynamic 3D projection mapping. Lastly, 3D modelling software such as Blender (2020) are chosen to create and edit 3D shapes.

### 4.2 Phase 1: Implementation of object detection

In this phase, the main aim is to recognise the object constructed by user using a set of Lego blocks. Based on the scope decided, a user can choose from a group of Lego blocks with the size of 2x2, 2x3, and 2x4. To ease the process of recognition, different colours are used for Lego blocks with different sizes. For example, all the 2x2 blocks could be in white colour, 2x3 blocks could be in yellow colour and so forth. With this, colour detection can be adopted to determine the location or where the user has placed the block in the object.

The entire process is continuously monitored. When a user picks a block and adds that to the existing structure, the location is determined and recorded in real-time, so that we can see the 3D reconstruction of the object in real-time on the screen. The trigger area technique is used by setting up trigger areas (using HitTest VVVV function in the shape of a circle) in each of the slots available on a Lego brick. And if the joining Lego block's center touches one of the trigger areas it will lock onto that respective Lego slot as shown in Fig. 5. The translucent green circles represent the trigger areas, and the small white square represents the center point of the Lego object as captured from the webcam. When this center point touches one of the translucent green circles it will turn blue and shift the joining Lego object to its respective slot.

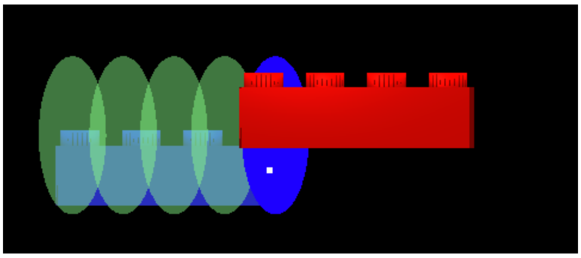


**Fig. 5.** Implementation of objection detection using trigger areas

### 4.3 Phase 2: Detection & tracking the assigned marker in real-time

In this phase, the aim is to detect and track a marker continuously, in real-time. Marker-based tracking was implemented by assigning the base pattern and colour of a Lego block to a marker. Both the colour-based tracker and Kernelized Correlation Filter (KCF) tracker were tested in preliminary experiments and results showed satisfactory accuracy in tracking under good lighting conditions as shown in Fig. 6. However, accuracy of tracking suffered under low light conditions, especially in the case of KCF tracker, where it was found to increase the offset in tracking.

Due to these challenges, the system uses col-our-based tracking and place an IR camera under the physical table to better detect the marker features. Additionally, as described in Table 5 in section 4.5 the Aruco marker and fiducial markers were also tested by switching to the bottom camera. This requires re-mapping of the coordinates as the object shape and size information detected from the top camera must now be tracked by the bottom camera. Undistort functions are also applied to the both the top and bottom cameras to reduce camera distortions using the intrinsic parameters gathered from camera checkerboard calibration.

For the colour tracking function a WithinRange (OpenCV function) is used to filter out the desired HSV (hue, saturation, value) colour range. In addition, a threshold value and some gaussian noise is also added to reduce noise. After this the individual colour masks are then applied separately to the video capture. The contours are then detected using the OpenCV function and the convexhull is calculated in a ForLoop based on the contours. Using the convexhull, it then draws the approximate polygon to display the object shape. This coordinate information must be then mapped to VVVV as most of the OpenCV function reside in VL which uses a different coordinate system.

(a) Colour based tracking (b) Kernelized Correlation Filter (KCF)

**Fig. 6.** Implementation of marker-based tracking

### 4.4 Phase 3: Implementation of dynamic 3D projection mapping

In this phase, it aims to map the virtual projections to the moving physical object in real-time. To support for dynamic 3D projection mapping in real time, the system must first be calibrated using the checkerboard to allow for accurate conversions of virtual coordinates to real world coordinates. The 10 x 7 checkerboard was printed and attached to a solid board. Twenty images were taken to in various positions covering all the x, y and z axis. VVVV (2021) recommends the camera calibration reprojection error to be less than 0.5. Here, the camera calibration reprojection error of the system scored 0.4. The camera calibration reprojection error is calculated by using the OpenCV camera calibration function.

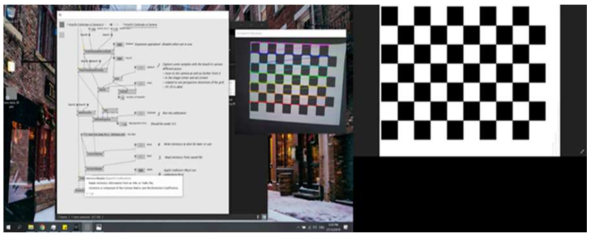


**Fig. 7.** Implementation of Dynamic 3D Projection Mapping (3D rendering)

The patch shown in Fig. 7 was tested to support for this by using the camera intrinsic and extrinsic parameters. In the preliminary tests, dynamic 3D projection mapping in real time was able to be produced however, the accuracy of the mapping can be improved as sometimes the projections would have some offsets. This is most likely due to the conversion of coordinate systems as shown in Fig. 8.

Additionally, manual calibration technique has also been tested. The manual calibration of the projector involves of marking the physical scene and matching the VVVV renderer to the camera viewpoint using homographic transform applied onto a quad. Homographic transform function is applied here instead of normal transform function as it will allow for individual control and placement of each corner of the renderer, which allows for finer adjustment. Through manual calibration the projector's renderer view will be the same as the camera view.

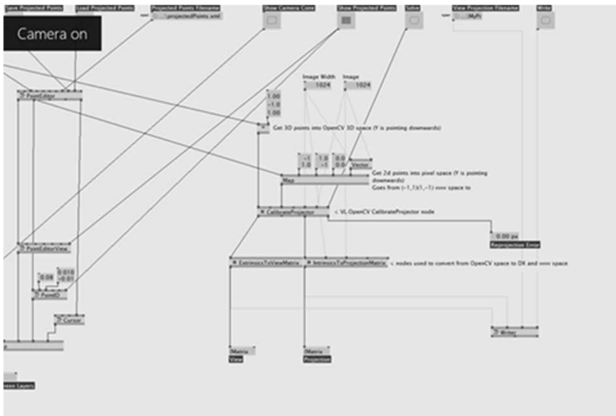



**Fig. 8.** Implementation of Dynamic 3D Projection Mapping (Mapping Coordinates)

### 4.5 Evaluation and analysis

This section elaborates details of the test results with comparison of different trackers, techniques for 3D reconstruction and projection mapping.

Table 5 displays the results of the tests conducted on different techniques of the system. The tests for tracking accuracy and consistency involved the subject being tracked to be moved around the scene in different speeds. The tracker must be able to seamlessly track its target and not lose the target when the target is being moved across the scene in three (3) mode of speed. The speed is considered as slow if it is less than 5cm/sec, medium if it is between 5 to 15cm/sec while fast speed was considered to be between 15 to 30cm/sec. The camera and FPS (frames per second) were kept consistent as a control variable.

The tracking precision is calculated based on the concept of intersection over union (IoU) (Khandelwal, 2020) of ground truth bounding box (i.e the actual target) and the predicted bounding box. This allows the assessment of the correct overlaps between the actual target and predicted bounding box. An IoU score higher than or equal to 0.5 is classified as a true positive and an IoU score lesser than 0.5 is classified as a false positive. This test was conducted three (3) times using different objects for each type of tracker and an average IoU score was calculated.

In the detection after loss of tracking test, the object would first be placed on the scene where it is being tracked. After this the object would be removed from the scene and then placed back in the scene. If the tracker is able to detect the object once the object is placed back in the scene, it passes the test.

The colour tracker passes the test for multiple objects of the same colour if it is able to detect two or more objects of the same colour placed part from each other as their own individual objects.

To support real-time reconstruction and record the relative placement of Lego bricks, two techniques were compared. The distance calculation technique involved of comparing the distance and angle from the center of one brick to another. Depending on the this, the Lego brick will lock onto one the adjacent slots available. Whereas the trigger area technique involves of setting up trigger areas (in the shape of a circle) in each of the slots available on a Lego brick. And if the joining Lego block's center touches one of the trigger areas it will lock onto that respective Lego slot. The evaluation of the accuracy score depends on the numbers of errors to determine the rating scale of 1-5 (from low to high). The high score of 5 would be rated when there is no or one error only; if 2-3 errors it would be rated as 4; if 4-5 errors it would be scored 3; if 6- 8 errors the score would be 2 and when there are more than 8 errors, the score would be rated as low score 1. These errors involve of errors in offsets, wrong slot joining and constant flickering between slots.

For projection mapping two main techniques were tested. In the coordinate lock technique, this involved of locking the camera coordinates retrieved from the colour tracker after the desired object has been created. This allows the projection of the desired shape of the object however, since the colour tracker is still being used the projections can interfere with it and affect the accuracy of the colour tracker. The other technique involves of locking the texture itself once the shape has been constructed. This no longer requires the colour tracker to be active once the shape is created therefore, the projections cannot interfere with the colour tracker.

**Table 5.** Comparison of various system techniques tested.

| Comparison of marker and markerless trackers | | | |
|---|---|---|---|
| | **KCF Tracker** | **Aruco Tracker** | **Fiducial Tracker** |
| Tracking Accuracy & Consistency | Slow, Medium Speed (<10 cm/sec) | Slow Speed (<5 cm/sec) | Slow Speed (<5cm/sec) |
| Tracking Precision | 0.71 | 0.95 | 0.74 |
| Detection after loss of tracking | No | Yes | Yes |
| Requires marker | No | Yes | Yes |
| **Comparison of colour trackers** | | | |
| | **Freeframe Dshow9 colour Tracker** | **Proposed Colour Tracker** | |
| Tracking Accuracy & Consistency | Slow Speed (<5 cm/sec) | Slow, Medium, Fast Speed (<30 cm/sec) | |
| Tracking Precision | 0.82 | 0.85 | |
| Detection after loss of tracking | Yes | Yes | |
| Tracking Multiple objects of same colour | No | Yes | |
| **Comparison of techniques for 3D reconstruction** | | | |
| | **Distance calculation** | **Trigger area** | |
| Accuracy | Low (2/5) [Requires fixed distance from the camera] | Medium (3/5) [Requires fixed distance from the camera] | |
| **Comparison of projection mapping techniques** | | | |
| | **Coordinate Lock** | **Texture Lock** | |
| Requires switching to another camera view | Yes | Yes | |
| Projections interfering with colour tracker | Yes | No | |

### 4.6 Implementation strengths & issues

The strength of the current project is 3D projections allowing the users to move tangible objects. This can improve the user's interaction as it can allow for a more natural interaction. Additionally, the support for real-time reconstruction allows for users to build their own shapes instead of just using the pre-defined shapes. In the tests conducted, the system was able to correctly detect and project the desired texture onto the specific structure constructed of up to 10 Lego bricks placed in random positions.

However, there were some issues faced in conversion of 3D virtual coordinates onto real world coordinates where the projections are mapped. This is due to the dynamic nature of the 3D projection mappings and therefore the accuracy in projections is a limitation. Furthermore, some issues faced in tracking of individual Lego blocks to be able to constantly detect the final shape due to camera occlusions. To improve the visibility of the projections most projection mapping applications are carried out in a low lighting environment however, this can be challenging for the camera to perform tracking of marker features.

## 5. Conclusions

The tracker-based real-time interactive SAR system was successfully developed after testing and analysis. The system projects the SAR virtual contents correctly mapped onto the real-world 3D objects with support for real-time 3D reconstruction. This has introduced a new interaction method that allows to create object detection and tracking in a real-time in the physical SAR platform. Moreover, it would have designed and developed the physical SAR platform suitable for dynamic projection mapping. The benefits from the proposed SAR platforms are stated below.

1. Static Vs Dynamic: Most projection mappings are static and therefore with the addition of dynamic content it can involve users in a more effective manner

2. 2D Vs 3D: It is easier to visualize complex data and ideas/plans through 3D projection mapping compared to 2D

3. AR Vs SAR: SAR provides a more seamless blend between the virtual and the real world compared to AR as the virtual content is projected to the real-world

4. Touch Vs tangible interaction: Tangible interaction provides natural ways of user interaction compared to touch interaction.

## 6. References

Al-Megren, S. and R. A. Ruddle, 2016. Comparing Tangible and Multi-touch Interaction for Interactive Data Visualization Tasks, TEI '16, Eindhoven, Netherlands.

Ball, C., 2018. Augmented/Mixed Reality for Events: The Pros and Cons for Each Medium and the Likelihood of Adoption, The meeting technology professionals. Available online at: https://www.corbinball.com/article/36-mobile-and-wireless-technology/243-vrvsar (accessed on July 10, 2021)

Blender. 2.83, 2020. Available online at: https://www.blender.org/ (accessed on July 10, 2021)

Cortes, G., E. Marchand, G. Brincin, and A. Lecuyer, 2018. MoSART: Mobile Spatial Augmented Reality for 3D Interaction with Tangible Objects, Front. Robot. AI, 5(93). Available online at: https://doi.org/10.3389/frobt.2018.00093 (accessed on July 10, 2021)

Fleischmann, O. and R. Koch, 2016. Fast projector-camera calibration for interactive projection mapping, IEEE.

Guo, Y., S. C. Chu, Z. Liu, C. Qiu, H. Luo, and J. Tan, 2018. A real-time interactive system of surface reconstruction and dynamic projection mapping with RGB-depth sensor and projector, International Journal of Distributed Sensor Networks, 14(7), Available online at: https://doi.org/10.1177%2F1550147718790853 (accessed on July 10, 2021)

Khandelwal, R., 2020 January 6. Evaluating performance of an object detection model, what is mAP? How to evaluate the performance of an object detection model? Towards data science, Available online at: https://towardsdatascience.com/evaluating-performance-of-an-object-detection-model-137a349c517b (accessed on July 10, 2021)

Kim, H., I. Takahashi, H. Yamamoto, S. Maekawa, and T. Naemura, 2014. MARIO: Mid-air Augmented Reality Interaction with Objects, Entertainment Computing, 5(4), 233-241, Available online at: https://doi.org/10.1016/j.entcom.2014.10.008 (accessed on July 10, 2021)

Laviole, J., 2012. Spatial augmented reality for physical drawing, UIST 12: The 25th Annual ACM Symposium on User Interface Software and Technology Cambridge Massachusetts USA, October 7

- 10, Available online at: https://dl.acm.org/doi/10.1145/2380296.2380302 (accessed on July 10, 2021)

Makarov, A., 2021, February 1. 10 augmented reality trends to watch in 2021: The future is here. Available online at: https://mobidev.biz/blog/augmented-reality-future-trends-2018-2020 (accessed on July 10, 2021)

Mendes, M., J. Almeida, J., H. Mohamed, and R. Giot, 2019. Projected Augmented Reality Intelligent Model of a City Area with Path Optimization, Algorithms, 12(7), 140, Available online at: https://www.mdpi.com/1999-4893/12/7/140 (accessed on July 10, 2021)

Milgram, P., and F. Kishino, 1994. Taxonomy of Mixed Reality Visual Displays, IEICE, Transactions on Information and Systems, vol. E77, 1321-1329.

Mousavi, H. H., M. Khademi, L. Dodakian, S. C. Cramer, and C. V. Lopes, 2013. A Spatial Augmented Reality rehab system for post-stroke hand rehabilitation. Stud Health Technol Inform. 184:279-85. PMID: 23400171

Open. CV., 2021. Available online at: https://opencv.org/ (accessed on July 10, 2021)

Park, M. K., K. J. Lim, M. K. Seo, S. J. Jung, and K. H. Lee, 2015. Spatial augmented reality for product appearance design evaluation, Journal of Computational Design and Engineering, 2, 38–46.

Park H. B., 2017. Available online at: https://www.youtube.com/watch?v=r4DTMedEH4U (accessed on July 10, 2021)

Wilson, A. D., 2005. PlayAnywhere, ACM Press. Available online at: https://dx.doi.org/10.1145/1095034.1095047 (accessed on July 10, 2021)

Winder, J., and K. Larson, 2017. Bits and Bricks Tangible Interactive Matrix for Real-time Computation and 3 D Projection Mapping, Paper presented at Future Technologies Conference (FTC).

VVVV. group., 2021, January 24. VVVV a multipurpose toolkit. Available online at: https://vvvv.org/ (accessed on July 10, 2021)

## AUTHOR BIOGRAPHIES

**Aye Chan Zay Hta** is a master's student at the Department of Computing and Information Systems, School of Engineering and Technology, Sunway University, Malaysia. He graduated with a BSc (Hons) in computer science (2019), validated by Lancaster University and a Diploma in information technology (2017) from the same university.

**Yunli Lee** is an associate professor in the Department of Computing and Information Systems, School of Engineering and Technology at Sunway University. She was awarded a BIT (Hons) in software engineering in 2002 from Multimedia University, Malaysia. In 2004, she received hermaster's degree in software from Dongseo University, South Korea and obtained her PhD in engineering (Digital Media) in 2009 from Soongsil University, South Korea. She is currently a senior member of the IEEE society, Professional Technologist of MBOT and Malaysia Director of the International Association for Convergence Science & Technology (IACST). Her current research interests include ultrasound imaging, the time series of FOREX data, and augmented reality technology.

**Wai Chong Chia** is a senior lecturer in the department of computing and information systems, school of engineering and technology at Sunway University. He received his BEng(Hons) in Electrical and Electronics Engineering from University of Nottingham Trent in 2006. He then continues to pursue his studies and receive his MSc in Electronics Engineering and Doctor of Philosophy from The University of Nottingham in 2008 and 2013 respectively. His main research is focused on visual processing for wireless sensor networks, embedded systems and mobile applications.