COMPUTER VISION AND AUGMENTED REALITY TECHNOLOGIES IN EXPERIMENTAL ARCHITECTURAL DESIGN EDUCATION AT THE AA

Immanuel KOH
Architectural Association School of Architecture, London, United Kingdom
artplusik@gmail.com

Abstract. This paper aims to investigate the potential of both open source software and new media (esp. computer vision and augmented reality) as tools for architectural design and education. The examples illustrated in the paper would be drawn mainly from students’ projects done as part of their AA Media Studies Course submission at the AA School of Architecture (AA) during the academic years from 2011/2012 to 2012/2013. The paper outlines the main approaches, which students have chosen to implement, both directly and indirectly, these new media and tools into their studio work at the AA. Section 1 briefly introduces a range of currently available open source computational design toolkits that are deemed useful for quick implementation of computer vision and augmented reality technologies. The related programming languages, softwares and hardwares would also be introduced and described accordingly. Sections 2 and 3 are accompanied with a visual catalogue of students’ projects to better illustrate the diversity in the understanding and implementation of computer vision and augmented reality technologies in architectural design. Section 4 serves to conclude the paper by first discussing briefly the feedback from students at the end of the course before clarifying the context of the research and thus its relation to recent work done by others using similar technologies.

Keywords. Computer vision; augmented reality; generative design; interaction design.

1. Introduction

The use of Computer Vision and Augmented Reality Technology as a tool for spatial design exploration has become more accessible to both designers and the general public in recent years. In addition to hardware affordability, the various widely available open source software tools and a rapidly growing community of complementary online forums or downloadable tutorials, where technical know-how are being shared freely, have given rise to a new mode of architectural design exploration both in cutting-edge academia and design practices.
The possibility to extract accurate spatial/depth data of real physical environments/objects and further embed their augmented forms virtually suggest great potential of such tools to be included in the architect’s standard digital toolbox of CAD/CAM technologies. However, the necessary programming skill-sets needed to best utilize these newly acquired tools proved rather challenging to the typical architectural students, despite the progressive inclusion of scripting courses in recent architectural design syllabus.

In order to facilitate a more conducive learning experience, the choice of software to be taught in the classroom is critical to the success of this endeavor in introducing new hardware and software knowledge simultaneously. At present, among the various creative coding open source software environments, there are generally 2 types of environment: 1) Visual/Graphic Programming and 2) Text-based Programming. The former is more visual whereby students use ‘patches’ to make more complex ‘definitions’ to perform the computation. These include Max/MSP, Quartz Composer, Rhino Grasshopper, Scratch…etc., each is well-supported with their corresponding Kinect wrapper/libraries. On the other hand, the latter includes Processing, Openframeworks, Cinder…etc. These are typically based on higher-level programming languages (e.g. Java, C++, Python…etc) with or without their own IDE. The decision to eventually use Processing in the course is due to various reasons from a teaching and learning perspective. It is the intent to provide a greater range of possible entry levels for students with different or no programming experiences and thus allowing them to more easily migrate to other programming languages in the future. At the same time, the existence of a thriving online community of Processing developers and resources also suggest the ongoing popularity and usefulness of the software concerned. The excellent API/documentation and thus the widespread usage of Processing among other design disciplines is another main factor that is considered, other than the relatively gentler learning curve of the programming language. Although a C++ based programming tool would be deemed much faster in real-time computation, the difficulty of the language is the main disadvantage as an introductory programming course for architectural students.

2. Kinect

2.1. TRACKING SKELETON DATA: SPATIAL ENVELOPE AS STRUCTURES & FASHION WEAR

The Microsoft Kinect Sensor has the capability in extracting and tracking multiple users’ skeleton data in real-time. The joints position and their corresponding joint orientation allows one to compute spatial envelopes which reconfigures itself based on the spatial relationships among the tracked body joints. The available skeleton data includes feet, knees, hips, torso, shoulders, elbows, hands and heads.
In particular, Project 1 (Figure 1) has used this capability of the technology to great effects. The students set out to generate a series of prosthetic armatures using the joints’ spatial relationships. A volumetric mesh ‘container’ is generated in real-time, which then becomes the ‘search space’ for the deployment of a branching algorithm. The branching geometric construct is an attempt to create a generative network of structural support that would eventually be 3d-printed with ABS or fibre-carbon material. Variations in density, profile, directionality and thickness of the branches are being generated based on the volume enclosed by the designated joints. Since the spatial dimension of each joints’ position would varies according to different users’ body frame and body language, this tracking prototype suggests the potential to generate customised prosthetic profiles for every individual. In addition, this prototype is programmed to output the geometries as standard CAD/CAM file format, thus facilitating a smooth workflow incorporating digital fabrication processes. Although the output and subject of this project may seemed more akin to a product design exercise, however, the skill sets and design approach adopted here is applicable and scalable in an architectural context.

In the case of Project 2 (Figure 2), the student constructed the entire virtual envelop of the body to create a prototype for an interactive fashion wear. A variety of cross-sectional profiles are aligned both in position and orientation to the corresponding body joints, where a continuous set of splines would then thread through them, forming a complex interweaved envelope. Each individual cross-section is a polygon that varies its scale and number of sides based on the proximity of its opposite joints (i.e. the distance between the left limbs and right limbs), thus changing the density of the overall network. Although the constant reconfiguration of the proposed fashion wear
is visualized graphically, the concept can be further implemented with elastic fabric, LEDs, micro-controllers and sensors (e.g. LilyPad Arduino and flex sensors, etc.).

2.2. TRACKING USERS’ 2D PIXELS/3D POINT CLOUD: DANCE PERFORMANCE AS AN INTERACTIVE 2D SILHOUETTE PROJECTION AND A DYNAMIC 3D TRACE

In both Project 3 and Project 4, professional dancers become the co-creators, alongside the students, of the generated virtual space/form. Both attempt to visualize the energy of each dance move and further enhance the experience.

In Project 3 (Figure 3), the intention is to enhance the dance performance by projecting visual effects onto the dancer’s body in real-time. This is achieved by embedding interactive particle-systems within the extracted 2D silhouette of the dancer. The movement of the particles would response according to the outline of the projected silhouette, thus any contrasting and abrupt dance move would immediately intensify the collision behaviour of the particles, especially those along the boundary of the silhouette. The Kinect is used here to first extract the dancer’s 2D pixels and translating it into a Gaussian blur image before undergoing a blob detection algorithm. The algorithm aids in the extraction of the body’s silhouette as an enclosing polyline. This closed polyline would act as the boundary constraint where particles with varying collision behaviour interact. Since the computation and visualization occur in real-time, the performing dancer responses directly to the visuals and intuitively alter his/her subsequent dance move. The project proposed an improvised space of encounter.

In Project 4 (Figure 4), another student attempts to ‘freeze’ the dancer’s move in 3D space. The Kinect sensor is able to extract the depth map of the dancer at any point in time and project these as a 3D point cloud. In this way, the dance could be viewed in a 3D virtual space whereby one could navigate the view using the cursor.

Figure 3. Dancer improvising based projected particle system on his body silhouette.  
Figure 4. 3D Point cloud traces of a dancer.
to pan, orbit and zoom. Similar to the Project 3, the dancer is able to response to his/her previously ‘frozen’ point cloud and improvise based on the newly generated spatial entity. In fact, the code is programmed in such a way that whenever there is a prolonged pose during the performance, a set of the dancer’s point cloud would be deposited on the 3D scene. In addition, certain designated joints would continuously trace their movement and generating 3D meshes. This project illustrates the use of the Kinect as a potential tool to record and evaluate body movement in 3D space.

2.3. GESTURAL RECOGNITION: PHYSICAL HANDS AS VIRTUAL CRAFTSMAN’S HANDS

The Kinect sensor is able to recognize a list of human gestures, such as waving, raising and swiping the hands. Unlike the earlier projects, which require an initial full body pose calibration in order to extract any skeleton data, here the sensor only requires gesture detection without even a full view of the body. In Project 5, the students begin with a concept to reflect the nature of the Kinect’s scanned data. The scanning technique used by the Kinect is a variation of the standard Structured-Light 3D scanning system. Instead of projecting stripes of visible lights and measuring the deformation of the reflected light, the Kinect projects patterns of infrared light. However, regardless whichever variation of the system, the data collected is always based on a specific view rather than a full 360 degrees scan. The students understood this technical limitation and decide to scan and create objects based on the simple principle of axial and planar symmetry. For instance, a table lamp has an axial symmetry while a human head has a planar symmetry. The final proposal combines the Kinect’s capability in gesture recognition and the concept of symmetry to create a new design interface. With this interface, any user could use their hands to virtually sculpt a series of symmetrical scanned objects. For instance, the Digital Potter’s wheel interface (Figure 5) allows users to sculpt the objects in way resembling a pot-
ter ‘spinning clay’. In the Digital Mirroring interface (Figure 6), users only sculpt one side of the object and the other side would generate an inverse form along the plane of symmetry. Once the user is satisfied with the newly sculpted versions, the files could be directly sent for 3D printing or CNC milling.

3. Augmented Reality (AR)

3.1. AR MAKERS AS AUDIO MARKERS

AR markers not only allows one to map virtual 3D objects onto a physical plane, it could also be used to trigger audio samples. In Project 6 (Figure 7), the student uses both ‘Processing’ and ‘Ableton Live’ (an audio sequencer) to reinterpolate an architectural representation of a plan. By envisioning the plan of Chungking Mansions’ ground floor in Hong Kong as a drum machine, each of its hallways becomes a button that triggers the repetitive sound specific to the portion of the building. One could

Figure 6. Digital Mirroring interface detecting hand gesture to sculpt scanned head with planar symmetry.

Figure 7. Series of AR Markers act as an audio sequencer to produce different sound scenarios.
then start to play and overlap different sound scenarios or even to begin a new path through the building’s hallway by separately triggering each of the hallways. The traditional ‘mute’ architectural plan is here embedded with the dimension of sound.

3.2. AR MARKERS AS SOCIAL MARKERS

Project 7 explores how might augmented reality be used as a tool for measuring the interaction intensity of a group of people in real-time. Each student in the class is assigned with an AR Marker and a webcam is used to track their proximity among themselves (Figure 8). The greater the interaction, the more connections are visualized. It is a simple attempt to use AR technology as a tool to understand spatial organization based on the actual interaction of people within the space observed.

3.3. AR MARKERS AS SPATIAL MARKERS

Project 8 (Figure 9) uses AR markers as physical and tangible controllers to gain dynamic control of the overall spatial organisation. By tracking the various AR markers with the webcam and designating them as input points and attractors, the student has proposed an alternative way of architectural modelling.

4. Discussion

4.1. STUDENTS’ FEEDBACK

The students’ feedback of the course at the end of each term suggests a number of pedagogical issues for considerations. On the positive aspects, students find
learning scripting in an alternative context (i.e. beyond a typical CAD–based implementation context) very liberating and creative, while still allowing them the opportunity to learn the fundamental principles of programming and their actual application in a hands-on manner. The use of open-source software also encouraged students to appreciate the current collaborative/making/hacking phenomenon in our contemporary culture. On the negative side, due to the nature of the course structure (i.e. 3 hours per week with 8 sessions in total), students find learning both the scripting language and the new technologies challenging when asked to appropriate them with their current design studio work. One of the attempts being later implemented to reduce these difficulties among students is to allow them to work as teams of maximum 2–3 people who share similar approaches and techniques, though slightly differing subject matters. In this way, the learning curve of the more technical aspect of their project is greatly reduced while encouraging more brainstorming among them within the teams. Students also find that they could afford more time for the actual design of their proposed projects.

4.2. RESEARCH/TEACHING CONTEXT & RELATED WORK

The Xbox Kinect sensor and Augmented Reality technologies have been increasingly utilized as tools in recent years by researchers to explore a range of highly specific research agenda. Most notably is the Xbox Kinect sensor being used for gestural tracking and 3D scanning. For instance, Fox and Polancic (2012) utilized the Kinect sensor to more accurately tracked the gestures of users as part of their research in developing a gesture-based catalogue of remote controlling dynamic architectural spaces. Here, the technology played the role of evaluating their initially formulated gesture vocabulary. At a more generic level in the use of the
Kinect’s gesture tracking capability, Pak et al (2011) used the Kinect mainly as a low-cost and portable touchless interface in forming part of their textile-based portable immersive environment proposal. Roupé et al (2012) presented a similar use of the Kinect but focuses on the body as an interface for virtual reality navigations. All the three references, directly or indirectly suggest that the gesture-based interactions is a more natural interface for spatial manipulation, communication and reasoning. At the other end of the spectrum, Prousalidou (2012) discussed the use of the Kinect as primarily a 3D scanner to extract material data directly from physical model in order to better inform the digital model simulating the design of fabric formwork panels. In this case, the Kinect helps to provide a higher resolution of 3D data in the research.

Each of the references cited above has a clear outline of its research’s problems, objectives and methodologies from the outset of its investigation. Typically, the Kinect is a means for them to further validate their hypotheses during the research’s real-world testing and evaluation stage. In contrast, the projects illustrated in this paper here attempts to problematize the use of Kinect/Augmented Reality in relation to architectural design as an applied research via the structure of a complementary teaching module, instead of a full-fledged research project spanning a few years within the laboratory. In other words, the process is almost the inverse here. This paper is a direct outcome of the various experiments done within the context of a term long media studies module originally designed to first and foremost to both introduce students to scripting as an important design skill set and to complement the concurrent core studio work of the students from both the AA’s undergraduate and graduate schools. Thus, on the outset, each student would already have his/her own research topic and agenda shaped by their respective core studios, each taking on differing trajectories. The course tutor (i.e. author of this paper) would then guide the students, both conceptually and technically, in appropriating the technology (i.e. Xbox Kinect or Augmented Reality) into their intended research specifics, creating a dynamic feedback loop with their core studio work. In this way, the tutor aims to ensure a diversity of approaches being proposed in the exploration of the technology as an applied generic architectural design tool within the limited span of the term-long curriculum. However, some students have chosen to use the course as an independent study, while others did manage to integrate the work with their core studio work in varying degrees of success. The common objectives shared by the projects illustrated in this paper is the experimental appropriation of the selected new media, as well as, the minimal level of scripting competency to be achieved by the students at the end of the term.

This research conducted both in an open framework and in an applied design manner thus insists that it is the task of the architect or designer to make sense of
these new technologies within their design workflow as a means to enrich their design activities and to push the discipline in unexpected directions.

References


