Mixed-Reality and Project Based Curriculum Development: Empowering STEM Learners and Educators in Southwest Virginia Appalachian Region

Denis Graćanin*  B. Michael Stewart†  Timothy Duncan‡  Mohamed Handosa§  Hendrik Schulze¶
Virginia Tech, USA  Washington County VA Public Schools, USA  Virginia Tech, USA  VRVis Research Center Austria

ABSTRACT
The recent advances in mixed-reality (MR), mobile and related technologies indicate a tipping point in our ability to go beyond the standard educational, and training settings and study adaptive processes during complex physical, social, and educational interactions in realistic complex environments. Such technologies are exciting because they permit educators to move beyond artificial experimental settings that have limited ecological validity. We hypothesize that using these technologies in a real-world context, where computer generated stimuli are blended with the real-world stimuli, will provide necessary affordances in support of higher-fidelity embodied interactions resulting in more effective skill development, experiential learning, training and education. That will, in turn, help address prohibitive social impacts (e.g., resource and other disparities) to provide an exciting opportunity to re-define education and learning experiences while maintaining utility and effectiveness. To illustrate the challenges and opportunities we describe the development and deployment of a HoloLens application for Nurse Aide skills training. The application is used by nursing students at Washington County (Virginia) Career & Technical Education Center. The goal is to augment the student’s experience in the classroom settings and to provide a rich set of educational contents in an MR environment.

Index Terms:  Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality; Applied computing—Education—Computer-managed instruction

1 INTRODUCTION
Technological advances, changes in manufacturing and economic factors result in new approaches to how products and services are identified, developed, produced, distributed, and maintained [13]. Globalization and technological advances have changed the demand for production workers and the nature of production jobs. Information technology continues to transform our society. Advances in automation, robotics, artificial intelligence, and machine learning require new models of human engagement [19].

The US does not have formal collaboration mechanisms among governments, educators, labor representatives, and employers addressing workforce development policies and practices. Those efforts are polycentric in nature, supported by private and public investments [17]. We need better understanding of factors that influence career pathways, expanded certifications and other non-degree credentials and better indicators of K-12 STEM education results [15].

Many businesses observed the lack of needed mathematics, computer, and problem-solving skills in high school graduates. It shouldn’t be a surprise that 75% of 8th graders are not proficient in mathematics [22]. Implications are that significant investments need to be made in K-12 STEM education. Recent efforts focused on the integration of the individual STEM disciplines [21] and the development of new framework and assessments for the next generation science standards [23, 24]. The developed framework has three dimensions of science learning, the core ideas of individual disciplines, the key cross-cutting concepts linking the individual disciplines and the practices used by scientists and engineers [24].

Connections among the formal education system, afterschool programs, and the informal education sector could improve STEM learning [25, 28]. Consequently, teachers must be provided opportunities to teach STEM in new ways and to interact with their peers facing similar challenges [14]. However, teachers have limited engagement in policymaking, partly due to many job-related requirements during the school year [18].

Collaborative actions can improve the education and skills of the US workforce, particularly in manufacturing and high-tech industry. “Businesses, local school districts, labor, community colleges, and universities should form partnerships to help students graduate from high school, earn an associate’s degree, and take part in continuing education in the workplace” [13].

Most of the responsibility for and activities needed for economic growth are local. Therefore, strong local college-university-industry collaboration is essential to enhance student learning and to prepare the STEM workforce of the future [16]. The challenge is how to provide better course offerings, lab activities, experiential and work-based learning programs [3, 4], and other activities. In other words, the goal is to enable students to acquire knowledge, skills, and attributes they need to be successful in the STEM workforce.

Achieving this goal requires additional resources to supported hands-on, experiential activities and to embed assessment activities into engaging classroom activities [20]. However, regional socio-economic disparities and lack of financial resources make that goal very difficult to achieve.

Rural regions, like the Appalachia region, face particularly significant challenges regarding the availability and adequacy of resources and services due to a variety of factors, including economic challenges, geographic isolation, reduced professional resources and capacity, educational opportunity, and cultural characteristics. Rural communities are also affected by the “digital divide,” between those who have ready access to information and communication technologies, such as computers and internet, and those who do not.

In spite of those challenges, local communities are committed to investing in the future of its students by providing a comprehensive, wide-ranging scope of education that will prepare them for the challenges they will face as 21st-century citizens. The students should have access to the technical skills and tools they need from pre K-12 environment and throughout the remainder of their educational career. That requires environment (where students learn, to occur in a variety of places including home and in cross-curricular ways), engagement (evidence of students being involved in engaging work, where the teacher serves as a facilitator), tools (students have

*e-mail: gracanin@vt.edu  †e-mail: mstewart@wcs.k12.va.us  ‡e-mail: tsduncan@wcs.k12.va.us  §e-mail: handosa@vt.edu  ¶e-mail: hschulze@vrvis.at
ready and equal access to technology tools to meet 21st century curricular needs, and results (accountability requires that technology expenditures are tracked and assessed for impact on students).

The recent advances in mixed-reality (MR), mobile, and related technologies indicate a tipping point in our ability to go beyond the standard educational and training settings and study adaptive processes during complex physical, social, and educational interactions in realistic complex environments. We build on an ongoing partnership with Washington County Public Schools (V A) to use MR technology for project- and skill-based curriculum for nursing students at the Washington County Career & Technical Education Center (WCCTEC). This pilot project provides a grounded design approach leading to the initiative to overhaul STEM curriculum.

2 RELATED WORK

MR technologies allow for providing users with an environment that blends the physical surroundings with virtual objects. The recent advances in MR technologies provide a great opportunity to support deployment and use of MR applications for training and education. Users can interact with virtual objects that can help them be more engaged and acquire more information compared to the more traditional approaches.

Virtual reality (VR) and Augmented Reality (AR)/MR technologies [2] technologies/applications are becoming more affordable and thus more accessible to general users. While there is over forty years of research in this area, we still need more findings to better understand challenges when it comes to developing MR applications. MR [10] / AR [1], and VR [12] applications have been used mostly in higher education (science, humanities, and art), unlike vocation education. In the medical/health domain, MR has been mostly used for medical treatment, surgery, rehabilitation, education, and training, but not universally to other medical fields [5].

Cubillo et al. [6] describe a learning environment based on AR, ARLE (An Augmented Reality Learning Environment). ARLE can be used by both the teacher to develop AR educational resources and students to acquire knowledge in the area. They found a significant learning improvement for users who used ARLE compared to the users who did not use ARLE. Underwood and Kimmel [30] conducted a study of the use of MR simulation to introduce social justice and understand culturally relevant pedagogy to pre-service school librarians. Computer simulations can enhance traditional instruction (especially laboratory activities) but more work is needed to explore the possible impact of teacher support, the lesson scenario, and the computer simulation’s place within the curriculum [29].

The advantages of MR include learning gains, motivation, interaction, and collaboration. Better learning performance, motivation, and engagement demonstrate the effectiveness of MR [1]. However, there is the need for longitudinal studies to study the evolution of knowledge and skills over time and to inform about the suitability of MR to support significant learning. In order to support continuous and natural user interfaces and interactions in such an environment, there should be means to capture user inputs. Different MR devices tend to support user interaction in a variety of ways, usually using voice-based, gesture-based, and tangible-based user interfaces.

While being less immersive may be an inherent problem for MR technology, meanwhile it also proposes an interesting question for how we can expand the application scope for full utilization of this technology. One of the main challenges of MR is the limited field of view. As such, how to visualize large chunks of (or big) data is questionable [27]. With technology progressing forward, it is expected that the field of view can be enlarged even beyond of the human field of view in the near future [32].

MR systems inherently depend on the surrounding space to support user interactions. Our cognitive processes are dependent on how our body interacts with the world (affordances) and how we off-load cognitive work onto our physical surrounding (embodied cognition). Embodied interactions demonstrate the importance of the body’s interactions with the physical world. Interaction of our body and the surrounding physical world affect our cognitive processes and embodied cognition [31]. Embodiment cognition leverages the notion of affordances, potential interactions with the environment, to support cognitive processes [8]. Embodied interaction and embodied user interfaces lead towards invisible user interfaces and move the computation from the computer desktop to physical space and place. MR provides support for embodied learning [11].

3 CHALLENGES

The main goal of MR is to enrich the actual physical environment with digital (virtual) entities. To achieve full immersion, such an environment should react to a users behavior appropriately and interaction should be as natural and intuitive as possible.

MR systems inherently depend on the surrounding space to support user interactions. Compared to the use of gaze and gesture or the use of voice commands to interact with virtual objects, the support for embodied interactions can provide a more natural way to interact with the surroundings and allows for developing a rich user interface. Embodiment cognition leverages the notion of affordances, potential interactions with the environment, to support cognitive processes [7].

The spatial awareness of an MR device, such as Microsoft HoloLens, allows a great degree of freedom regarding recognition, movement, and exploration of confined spaces and physical objects, enriched with virtual objects. However, the interaction capabilities of the HoloLens are limited by multiple factors which play an important role regarding the natural feeling of an immersive environment.

The interaction concept is based on gaze tracking and hand tracking in combination with a predefined set of hand gestures recognized by the HoloLens, as well as voice commands. This concept yields limitations of the device.

We focus on limitations emerging from the gaze tracking and hand tracking in combination with hand gestures. Both hand tracking and hand gesture recognition need to be activated by extending an arm forward, so the active hand is within the HoloLens’ field of view while pointing a finger upwards.

The first limitation is the possible space of interactions. The hand gestures recognition is possible only within the HoloLens’ field of view. Additionally, interaction with a specific object or a specific part of an object requires the user to gaze at it. Both of these preconditions create the need for not necessarily natural behavior patterns in order to interact with objects in a given environment.

The second limitation is the range of interactions. The HoloLens’ interaction techniques include gaze tracking in combination with hand gestures. Any hand forming the ready gesture within the field of view of HoloLens can get tracked). However, the main problem is that there is no way to tell whether the detected hand is the left or right one. Moreover, there is no way to tell whether the detected hand belongs to the user or to another person.

The hand is treated as a disjoint object with no information about its side or which body it belongs to. From a programming point of view, the HoloLens API triggers an event whenever any hand is detected. The hand is then assigned an Id. When the hand goes out of the field of view or abandon the ready hand gesture, it is no longer tracked. When the same hand is used and detected again, it is assigned a new Id and is treated as a new hand. Custom ‘gestures’ that resemble natural interaction patterns as well as performing tasks which require interaction with both hands at the same time or interaction with multiple objects simultaneously are therefore not (or only to a certain degree) possible.

These limiting factors prevent users, especially the inexperienced ones, to perceive actions and interactions within MR environments as seamlessly and immersive as possible. Using tracking devices enables the awareness, recognition, and processing of all body parts.
(not only hands) as well as the orientation of body parts. Utilizing additional information allows the development of more complex interaction schemes, involving multiple body parts and a higher level of detail for a broader range of recognizable gestures.

The recognition of the entire skeleton allows for example interaction with objects outside of the HoloLens’ field of view or interaction with both hands at the same time. Moreover, interaction is not restricted to gestures performed with hands but also with other body parts. The skeletal information, in combination with the spatial awareness of the HoloLens, makes it possible to derive and interpret joint movements as context-aware (semantically charged) gestures. For example, the user can use a foot to push an object out of the way without actually looking at the object.

4 APPROACH

When interacting with a smart space, especially in remote interaction scenarios, users may prefer to interact with the virtual objects in a more engaging manner using their limbs almost the same way they interact with the physical objects. For instance, a user may control a remote appliance by touching virtual buttons on its corresponding virtual model. Unfortunately, this kind of embodied interaction is not supported by the current MR devices. However, support for such user interactions can be provided with the aid of other devices.

Compared to the use of gaze and gesture or the use of voice commands to interact with virtual objects, embodied interactions can provide a more natural way to interact with the surroundings and support a rich user interface. Moreover, the embodied interaction can support interaction scenarios that cannot be supported under the limits of the interaction techniques incorporated in the HoloLens device. For instance, it can support simultaneous interaction with different virtual objects (e.g., using both hands).

Some tracking devices can track several users simultaneously. Consequently, MR devices may receive several tracking data sets for several persons. In that case, the MR device may need to identify which data set belongs to the user. Given that MR devices are head-mounted devices, the current location of a device in the MR coordinate system gives a good indication of the current location of the user’s head. Comparing the device location with the registered tracking data sets can reveal which data set belongs to the user.

We use another approach to extend the range of embodied interactions. This approach extends the functionality of Microsoft HoloLens to support an extended range of embodied interactions in an MR space (Figure 1 top) by using the Microsoft Kinect V2 device (Figure 1 center). The developed technique allows HoloLens MR applications to have access to the tracking information of user’s skeleton and joints. Figure 1 bottom shows a HoloLens-rendered virtual skeleton aligned with the corresponding physical body. That can provide additional support for embodied interaction.

Compared to the use of gaze and gesture or the use of voice commands to interact with virtual objects, the extended support for embodied interaction can provide a more natural way to interact with the surroundings and allows for developing a rich user interface. Moreover, the embodied interaction can support interaction scenarios that cannot be supported under the limits of the other interaction techniques incorporated in the HoloLens device.

Based on the proposed approach, we implemented a system that integrates HoloLens devices with Kinect devices. A Kinect server application tracks the user skeleton using the Kinect device. The HoloLens application obtains tracking data from the server through its Kinect client module before the registration module maps it to the HoloLens coordinate system using a transformation matrix.

We use a four-step process (Figure 2) to determine the transformation matrix. The goal of each step is to collect two corresponding points, one from the Kinect coordinate system and another from the HoloLens coordinate system. The four-point pairs are collected by asking the user to place a hand at four different positions in space that are indicated by virtual objects. Once the user’s hand is in position, hand tracking information is collected from both Kinect and HoloLens to form a point pair. Afterwards, an algorithm is applied to obtain the transformation matrix and save it for later use.

5 HOLOLENS APPLICATION

We applied our approach to the development of a HoloLens application for Nurse Aide skills training [26]. The goal is to augment the student’s experience in the classroom settings and to provide a rich set of educational contents in an MR environment.
Students need to learn a set of skills with quite specific steps in a certain order. This requires not only the theoretical knowledge of how a specific skill needs to be done but also practical application to manifest the exact workflow required. With limited space and limited availability of hospital equipment in schools, the number of workstations to actual practice the skills are limited as well.

The application simulates an actual care-giving situation with all the equipment necessary to perform the assigned skills (Figure 3). There are 22 skills ranging from “Hand Hygiene (Hand Washing)” to “Transfers from Bed to Wheelchair using Transfer Belt”.

The developed HoloLens application recreates the scenery of a hospital room (Figure 4). Within this virtual hospital room are the required objects and props to perform the skills in a ‘close to reality’ environment. Figure 5 demonstrates an embodied interaction with digital entities, a denture, and a toothbrush. Our application not only guides the student through the steps of a skill but also requires the student to perform specific and detailed interactions within the MR environment to proceed to the next step of a skill.

Almost all skills require at some point a more detailed user tracking than the HoloLens alone can provide. For example, a crucial part of proper hand washing requires the student to keep the hands and forearms at a downward angle to prevent ‘contaminated’ water to run down the arms. With the HoloLens alone, there is no possibility to check this condition. With the additional data about the entire user skeleton provided by the Kinect, we can track all the user’s actions required to perform the skills.

While navigating through the menus of the application and setting up the system (registration of the Kinect position and the HoloLens user position, proper positioning of the sceneries) relies on the actual HoloLens hand gestures, the interactions required to train a skill are done based on the Kinect skeleton data.

As most of the students had little to no experience with MR devices or virtual environments in general, it took some ‘warm-up’ time for the students to get used to the new experience and to move around and interact comfortably with digital entities. Initially, instead of actually walking around and exploring the boundaries, the students simply turned their heads. For example, rather than walking to the sink and open the faucet like in a real environment, the students started bending forward and reaching for the faucet by extending the arm. While we had to provide a lot of instructions and guidance for the parts of the application where HoloLens hands gestures are required, after the initial learning phase the students completed the skill training without or with only limited guidance. With the integration of the Kinect into the system, we eliminated the need for unnatural gestures. This reduced the threshold for the students to comfortably interact with virtual objects.

5.1 Interactivity

A user should receive instant feedback as they interact with objects in the MR space. A noticeable delay in response to the user’s commands can degrade the user experience. Therefore, the user’s skeletal information should be delivered to HoloLens with minimum latency to ensure the responsiveness of the system. The responsiveness is determined by the delay (latency) between the time at which the user makes a given gesture/move and the time at which the user receives the corresponding feedback through the MR device.

The HoloLens application is based on a custom developed framework that provides a template and related functionality to implement skills. Rather than hard-coding individual skills, each skill is implemented in a structured way by implementing a sequence of steps for each task in the skill.

Figure 5 shows a user squeezing toothpaste on a toothbrush before washing a denture. Figure 5 top shows a side view of the user in the physical environment (classroom). Figure 5 bottom shows the MR environment, as seen by the user. The user uses both hands to achieve this task. It is essential that the latency between the user
gestures and the resulting changes in the MR environment is limited.

The overall latency $d \geq d_1 + d_2 + d_3 + d_4$, where $d_1$ is the time it takes the Kinect device to capture a frame and send its data to the workstation; $d_2$ is the time it takes the workstation to extract skeleton information from the received frame data producing a skeleton information message; $d_3$ is the time needed to send the skeleton information message from the workstation to the HoloLens; and $d_4$ is the time it takes the HoloLens to provide a feedback based on the received skeletal information.

The overall latency was determined by capturing multiple MR video recordings of user gestures, specifically the closed hand gesture. Exploring the frames of the captured videos revealed that it takes at most four frames for the HoloLens to provide a feedback after the user gesture takes place. Although the Kinect-based recognition involves several processing and communication steps, results have shown that its performance is comparable to that of the built-in HoloLens recognition. In fact, Kinect-based recognition can often perform faster than the built-in HoloLens gesture recognition.

The total system delay ranges between 66 and 134 milliseconds. This estimated latency is caused by the processing and communication steps that take place between a change in user skeleton state and providing the corresponding feedback. Delays $d_1$ and $d_4$ are device specific and beyond our control. The measured average values of $d_2$ and $d_3$ are approximately 0.157 and 0.476 milliseconds, respectively. Compared to the overall latency, both $d_2$ and $d_3$ are negligible.

Current technical constraints and hardware limitations make a comprehensive solution that combines the multitude of required sensors with MR devices difficult to achieve. However, there are many opportunities for combining multiple conventional and unconventional data sources (not only tracking devices) in a comprehensive framework. When the data sources provide overlapping coverage of the MR space, these sources can be used for internal data alignment, error correction, and increased accuracy.

5.2 Pilot Study

The initial requirements for the HoloLens application were determined through discussions with WCCTEC teachers and students. They were asked to try the HoloLens device using several health-related applications from the HoloLens app store. Those requirements were used to create the application.

This first version of the application was then tested at WCCTEC to gather preliminary feedback and to refine the user requirements. The refinements were mostly due to the calibration process, user interactions and the perceived versus actual latency. The second (current) version of the application has greatly improved performance in terms of interactivity, latency and ability to adjust the position of the MR environment within the physical world.

We are currently conducting a pilot study at WCCTEC. The study is conducted by teachers as a part of the curriculum activities. So far, nine subjects have completed the pilot study (average age 16.67, all females, all nursing students, six are 11th grade, three are 12th grade, playing video games on average 1.5 hours a week). Each subject was asked to complete the first two skills, hand washing and cleaning denture. After completing the skills, the subjects completed a survey with seven questions:

1. How well does the virtual space proportion match the actual physical space size and proportion?
2. How well is the navigation in the virtual space?
3. How would you rate the interaction with the virtual objects?
4. How easy was it to use the application in general (navigation, voice commands etc.)?
5. How would you rate the overall realism of the application?

6. How would you rate skill 1 (Hand Washing)?
7. How would you rate skill 2 (Cleaning Denture)?

The current results are encouraging. Two subjects observed that the sink should be raised and one subject had difficulties putting on the HoloLens device. Once the pilot study is completed and the application is revised, the formative study will be conducted.

6 Discussion

The extended range of interactions is particularly useful for training, experiential learning and skill development. For example, Nurse Aide skills training is usually done in a training facility or a lab equipped as a hospital room. However, using the described approach, a regular classroom can be converted into a virtual training facility. The surrounding physical structure (walls, floor, etc.) provides a physical location where digital entities (healthcare equipment) are placed. Thus, a modest investment in MR and tracking equipment can successfully replace an order of magnitude more expensive physical equipment and infrastructure. Compared to the immersive VR, the MR approach allows the user to see hands (body) while performing skills and facilitates embodied learning.

While the skeletal information lacks finer details, such as information about individual finger joints, it provides sufficient granularity to support a wide variety of tasks in many domains. For example, using a toothbrush to clean a denture (Figure 5) is a representative of tasks in various domains requiring holding and manipulating objects using one or both hands, without finger movements (e.g., playing drums, using hammer, moving or throwing objects). Furthermore, using the whole body skeletal information allows using body posture and movements for skills development. Probably the most illustrative example of opportunities are all the console video games that use Kinect, from dancing and sports to adventures and role playing.

The main thrust of the research is to explore the research issues and challenges related to the deployment of an MR based wearable technology framework within the context of STEM education to empower learners and educators to break the socio-economic limitations in a rural school division with limited resources. There are three main challenges. The first challenge is how to develop the supporting technology to support rapid prototyping and deployment of new content. We need to address cognitive and functional seams that are potential obstacles to the use of MR technology and empowering educators to develop new content. The second challenge is how to bridge technology and education silos to create a truly interdisciplinary framework that connects educational psychology and instructional design with MR technology. The third challenge is how to accurately measure and quantify benefits for learning and professional development.

A body of relevant research [9] indicate that the more complex the teaching activity is, the more opportunities there are to learn by teaching. The corresponding explanatory framework can support the incorporation and extension of practices that provide opportunities for students to learn by teaching their peers, such as cooperative learning, peer tutoring or peer assessment. The most effective approaches include active participation of teachers who engage in the analysis of examples of effective instruction and the analysis of student work, a content focus, alignment with school policies and practices, and sufficient duration to allow repeated practice and/or reflection on classroom experiences.
MR devices provide an affordable opportunity to develop immersive applications. However, their limited support for user input constrains the possible interaction scenarios. We described an approach to overcome this limitation by integrating MR devices with tracking devices. An MR application can rely on the user tracking information to extend its ability to capture user actions and to support interaction scenarios that were not possible before.

The pilot project for Nurse Aide skills development demonstrates that close collaboration with K-12 institutions, teachers, and students makes it possible to investigate and deploy innovative MR enabled curriculum within the educational and work settings. The project results and findings provide a solid foundation for future research and provide an exciting opportunity to re-define education and learning experiences while maintaining utility and effectiveness. We can use Microsoft HoloLens device to effectively augment the student’s experience in the classroom settings and to provide a rich set of educational contents in an MR environment. Developing and deploying such MR education tools requires identifying the set of services and topics to be supported. The Nurse Aide pilot project provides a grounded design approach to address the STEM curriculum needs in K12 education. The ongoing efforts focus on selecting curriculum topics to introduce the developed MR-based approach, as well as to explore the underlying psychological and physiological mechanisms, i.e., how will pedagogical approaches be influenced by deploying MR technologies. That can inform design practices that are age- and socio-culturally appropriate.

7 Conclusion

MR devices provide an affordable opportunity to develop immersive applications. However, their limited support for user input constrains the possible interaction scenarios. We described an approach to overcome this limitation by integrating MR devices with tracking devices. An MR application can rely on the user tracking information to extend its ability to capture user actions and to support interaction scenarios that were not possible before.

The pilot project for Nurse Aide skills development demonstrates that close collaboration with K-12 institutions, teachers, and students makes it possible to investigate and deploy innovative MR enabled curriculum within the educational and work settings. The project results and findings provide a solid foundation for future research and provide an exciting opportunity to re-define education and learning experiences while maintaining utility and effectiveness. We can use Microsoft HoloLens device to effectively augment the student’s experience in the classroom settings and to provide a rich set of educational contents in an MR environment. Developing and deploying such MR education tools requires identifying the set of services and topics to be supported. The Nurse Aide pilot project provides a grounded design approach to address the STEM curriculum needs in K12 education.

The ongoing efforts focus on selecting curriculum topics to introduce the developed MR-based approach, as well as to explore the underlying psychological and physiological mechanisms, i.e., how will pedagogical approaches be influenced by deploying MR technologies. That can inform design practices that are age- and socio-culturally appropriate.

References