Tangible Educational Game for Pre-school Children

by

Yuchen Wang

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Computing Science
University of Alberta

Abstract

Over the last decade, interest and excitement surrounding interactive surfaces and multi-touch tabletops has increased substantially. Although interactive surfaces have many unique and compelling qualities, the interactions they support are by their very nature bound to the display surface. This thesis presents a system that not only enables users to interact on the tabletop surface, but also make use of the space above it and provide useful feedback. An educational game for pre-school children was developed using this system and is presented in this thesis. It aims to benefit children, not only for fun, but also from an educational aspect and helps children with their motor skills and cognitive development. Experiments were conducted to examine the benefits and potential issues of this tangible tabletop educational game.

Table of Contents

List of Tables

List of Figures

...

…

Chapter 1

Introduction

This thesis explores the potential of tangible tabletop user interfaces as a tool to develop educational games for pre-school children. A number of existing systems, Tangible Users Interfaces (TUI) technologies and educational games are reviewed first. Then the thesis presents a system, which includes a surface tabletop as the basis of interaction and a depth camera to provide user information for more feedback and more natural interaction. An educational game for pre-school children was designed and developed on this system following the strategies of game design for children. Finally, we describe a user evaluation study we conducted and give a detailed analysis of the experimental results.

1.1 Motivation

Young children ages three to six play a wide range of digital games, which are now available on large screens, handheld screens, electronic learning systems and electronic toys. Their time spent with games is growing. Current research is beginning to address the questions about the effects that the rules, content, challenges and feedback of digital games make on young children and how games for this age group should be designed and implemented to best serve children's needs.

Well designed games can provide powerful interactive experiences that can foster young children's learning, skill building and overall development [23], while poorly designed games can be time-wasting and even do significant harm to the children, such as rewarding aggressive or anti-social behaviour and instilling anxiety [6]. Many digital games for young children are not developmentally appropriate or evidence based in design. Game designers and publishers who care about the benefits of their games should take the learning goals of the game and the capabilities and interests of the target population of children into account. More research is needed to discover game design strategies that are developmentally appropriate for young children. This is a relatively small research area in the children's media field, but as game playing is becoming more and more popular in pre-school children, it deserves more attention.

Tabletop computers featuring multi-touch input and object tracking are increasingly common platforms for research in Human Computer Interaction. However, such systems are normally confined to sensing activity on the tabletop surface. The notion of Continuous Interaction Space (CIS) was proposed in [52], where the tabletop surface and the volume above it are seen as a unified whole, but the space above the tabletop is still a rich but unexplored canvas for supporting user interaction.

1.2 Thesis Objectives

This thesis describes a TUI that allows multiple users to share and interact with digital content on a computer embedded tabletop and provides information on users' behaviour in the volume above the tabletop for further use. Based on this system, an educational digital game for pre-school children was developed. This thesis also describes theories and game features that not only make children engaged but also support children's learning. A user evaluation experiment on this game was conducted to prove this system has the potential to be an effective tool to foster children's learning and motivate intergenerational communication. The main issues addressed throughout this work include:

- How have TUIs implemented and applied in different fields in the past? What needs to be considered when we want to design a game for pre-school children?
- What hardware is needed for the system proposed in this thesis? How do they work and what functionality do they have?
- How can we use a depth camera to obtain the information of users' hands hovering on the tabletop and make use of that information?
- How should an educational game be designed and how should it be developed? What are the relevant design motivations?

The tangible user interface proposed in this thesis uses a tabletop as the basis of the interaction, which combines the advantages and flexibility of traditional screen displays with the capability of allowing people to share and interact with the digital content. To make use of body information before users touch the tabletop and to extend the interaction medium from the surface of the tabletop to the space above it, we use a depth camera to obtain information on users' hands. Another goal for providing more natural interaction is to integrate the use of tangible objects into our system.

1.3 Thesis Organization

Chapter 2 provides a foundation-building discussion on the modern technologies and techniques that have been used to implement TUIs and how they are applied to various fields. To motivate and validate the idea of using digital games to foster children's learning and overall development, relevant work on game design for pre-school children is discussed and a number of strategies and concern about the game to target population of children are identified.

In Chapter 3, a detailed introduction of the hardwares to be used in our system is given, including their properties, functionalities and limitations. We state the reason of choosing these devices through an comparison with alternatives.

The architecture of the system is outlined in Chapter 4. It describes how we use this hardware to construct the system, and the algorithms we use to solve technical issues are addressed in this chapter. The steps required to estimate user's hand position and recognize tangible objects are described in detail.

Chapter 5 first gives a summary of the strategies and guidance to be considered in game design for pre-school children and a discussion about how we adapt them into our game. Then the chapter introduces the game content, interface, game logic and level settings of our tangible educational game for pre-school children. The props, scripts and the difference to traditional digital games are also highlighted.

Chapter 6 describes the steps to design and deploy a user evaluation experiment of the educational game proposed in this thesis. Then it validates the potency of our system as an effective tool to foster children's learning through a comparison of children's performance using the game developed in this system and a traditional digital game.

Chapter 2

Overview

This chapter gives an overview of the applications of tangible user interfaces and shows how they are applied to different fields. Then some of the most commonly used technologies for multi-touch tabletop and tangible recognition are described. Finally the relevant literature pertaining children's game design is reviewed.

2.1 Tangible User Interface

A tangible user interface (TUI) is a user interface in which a person interacts with a digital system through the manipulation of physical objects. Many tangible interfaces use a tabletop surface as base for interaction, embedding the tracking mechanism in the surface. Tabletop interfaces combine the advantages and flexibility of traditional screen displays with the social implications of allowing people around the surface to have shared access to the entities on the tabletop. Large displays are useful for information visualization when several people must jointly use the information to work together and accomplish a single goal.

TUIs are concerned with providing tangible representations for digital information and controls. They are implemented using a variety of technologies and materials and computationally augment physical objects by coupling them to digital data. These augmented objects often function as input or output devices. Tangible interaction is more intuitive because this hybrid system leverages users' prior knowledge from the real world, so we can combine the advantages of the physical and the digital world [32].

2.2 TUI Application Domains

This section discusses a sample of existing TUIs. The dominant application areas for TUIs are children's education, problem solving and planning, integration with mobile devices, music and performance, entertainment and rehabilitation.

2.2.1 TUIs for Children's Education

One area that has received much interest from tangible interface designers is children's education. First, learning researchers and toy designers have always followed the strategy of augmenting toys to increase their functionality and attractiveness, and second, physical learning environments support the overall development of the child. Marshall [53] takes a critical look at the potential of tangible interfaces to support learning. The paper describes six perspectives, including learning benefits, learning domains, types of activity, integration of representations, concreteness and sensory directness as well as effects of physicality on learning. Marshall highlights the need to empirically demonstrate measurable benefits of physical manipulation for learning.

A large number of TUI systems can be classified as computer-supported learning tools or environments. Digital manipulatives [82] are one type of TUI that build on educational toys such as construction kits and building blocks, which allow children to explore concepts relevant to understanding causality. The blocks can be annotated to represent a real-world system such as virus spread in a population. The blocks light up to show the data flow, and children can probe the current values being propagated through a block by attaching a small display onto it. Tangibles not only are useful resources, but also can slow down the pace of interaction, which is important to guiding children's action.

In designing interactive systems for children, a significant challenge is to address ways of integrating technology with children's social, cultural and physical circumstances [18]. Some tangible interfaces, such as KidPad [69] (See Figure 2.1) and StoryTable [12], enable children to take part in collaborative storytelling activities in a room. Stanton et al. [69] develop an interface using pressure mats and video-tracked physical props to navigate a story. StoryTable [12] uses a multi-touch surface to support children's storytelling activity in groups, in which children are urged to perform operations together.

2.2.2 TUIs for Problem Solving and Planning

Multi-touch displays are convenient tools for information sharing and visualization, and tangible objects endows natural interaction between multiple users, so TUIs are commonly used when collaboration of groups is needed. Code Space [8] includes a large multi-touch display, a depth camera, mobile devices, and a pen-enabled tablet PC. This system is designed for meetings, in which users can share content between two surfaces with direct touching and remote pointing. In Smart rooms [8] [42], users can go beyond the confines of the display and be immersed in a mixture of the real

Figure 2.1: KidPad [12] is an interface that uses pressure mats and video-tracked physical props to assist children to navigate a story. The green plate is used to control vertical movement of characters in the story.

and virtual world. Another promising trend in research deals with how to make better use of gestures and tangible objects to interact with the digital data of surface table more naturally and efficiently. LightSpace [78] uses multiple depth cameras and projectors which project images onto any surface and enables users to operate on objects in a natural way, for example transferring objects by simultaneously touching two separate display and sweeping object into their hand and dropping it on the other interactive wall. Kirk et al. [42] investigate the potential for surface computing in the home and give several guidelines from their findings about the design of applications run on the tabletop used in home. Tangible interfaces can also be used to create virtual sand tables, which are efficient tools for psychological treatment [41] and tactical planning [36]. Usually, tracked wands or gesture are used to interact with the table, for example for updating viewpoints, selection and shaping the ground surface.

2.2.3 TUIs for Mobile Devices

The combination of mobile devices with interactive surfaces has received a lot of attention in recent years [63, 64, 65, 16, 20, 79], and most research focuses on how to make the recognition of mobile devices more natural and convenient. A commonly approach is to use a camera mounted above the surface table. For example, Wilson [79] used a visual handshaking process on the BlueTable to initiate the transfer of photos, where mobile phones have to be placed on the surface and be visible for a camera above the surface. Rofouei et al. [63] associate multi-touch interactions to individual users and their mobile devices by comparing the accelerometer data received from phones and depth camera-based body tracking. The system relies on the fact that the acceleration measured by the phone should match the acceleration of the hand holding the phone, which leads to a limitation that the focused hand cannot be stationary. Also people cannot stand too close for the Kinect to correctly compute skeleton data. Dippon et al. [16] describe a seamless integration of un-instrumented mobile devices into an interactive surface environment using both RGB and depth cameras. The depth camera is used for tracking objects and fingers on or above surface, and the built-in RGB camera provides information for distinguishing multiple devices with the help of visual markers displayed on mobile phones.

Apart from the cameras, there are some others accurate and convenient approaches to connect the mobile devices to the digital table. Echtler et al. [20] developed a system identifying mobile devices via Bluetooth using infrared shadows to differentiate between several objects. In the system proposed in [65], which builds on the touch technique in [63], mobile devices do not need to be put on the surface, but the users have to touch the table to trigger an interaction. Different mobile phones are distinguished by timestamps. Mobile devices can also be used for selection of targets and creating touch events associated with the identity of the mobile, which enables many applications, such as data transfer by natural 'drop and pick', collaborative game playing, and browsing private information.

2.2.4 TUIs for Music and Performance

Dream Medusa [71] is a system allowing several participants to control a performance using voice and manipulating specifically created objects (See Figure 2.2). During the performance, hypnotic visual images are projected on a large screen. The movement of each participant's controller modifies different, mutually orthogonal aspects of the video playback, which determines the appearance of the video visualization. Audiopad [58] is a tabletop interface for musical performance, which uses objects as the music controllers. Ren et al. [60] present a design of a virtual musical instrument system that can be used by a collaborative group of users. The hitting velocity recorded by optical multi-touch interface help to simulate sound.

TUIs can also be applied to other media creations beyond musical performance. Held et al. [29] introduce a 3D puppetry system that allows users to create animations, requiring only a Kinect depth camera and a tabletop. It identifies puppets by extracting features of the images captured by camera and matches the features to a database. Through the segmentation of point cloud generated from color images and depth maps it can estimate the pose of puppets.

2.2.5 TUIs for Entertainment

Compared to traditional video games in which players use keyboard and mouse as input to interact with the digital elements displayed on mobile devices and computer, TUIs have obvious advan-

Figure 2.2: Participants interacting with dream.Medusa., which is an interactive media work [71] through controller objects

tages, with tangible objects offering more natural and convenient interaction and the large screen providing a more immersive gaming environment and support for multi-players. Goh et al. [25] give some guidelines on the design of multi-touch tabletop collaborative games, focusing on how to use the spatial separation afforded by large physical interactive spaces and multi-touch capabilities to motivate children to engage in collaborative activities. PingPongPlus [33] is a digitally enhanced Pingpong game using a tabletop that incorporates sensing, sound, and projection technologies. The embedded sound based ball tracking system connects the physical sports with the digital world, and the graphics and sounds are determined by the position of impact and the rhythm of play.

Interactive tabletops also seem well suited for use in museums [62, 31], because they can present compelling content in a walk-up-and-use form that supports collaborative interaction as visitors gather around the surface, enhanceing children's learning experience. Rizzo et al. [62] explore ways to develop interactive tabletop system for museums, including Virtual Book supporting dragging and dropping images to personalize multimedia book, and Coal Cave, in which users can interact with their fingers, wooden sticks, pens. Rizzo et al. note that physical manipulation in a virtual environment make children feel natural and easy, but also discuss some shortcomings, such as shallow engagement, interference between visitors, and avoiding to read and interpret the instructions. Horn et al. [31] try to apply the concept of Active Prolonged Engagement (APE) from the science museum literature to their system, which is an interactive tabletop exhibit to help visitors learn about evolution in natural history museums. The system is analyzed and evaluated from several aspects, like holding time, off-topic talk, visitor questions, collaborative engagement, and talk among groups.

Figure 2.3: Rehabilitation exercises on the AIR Touch [4]

2.2.6 TUIs for Rehabilitation

The application of virtual reality and tangible tabletops to assessment and rehabilitation has gained considerable impetus [59, 19, 80, 46, 48]. Leitner et al. [46] summarize a process to design tangible tabletops for rehabilitation, including how to adopt and extend the therapy concepts for Tangible tabletops environment and how to set the types of interaction objects and different levels of difficulty. They derive three major concepts for rehabilitation: copy a pattern, compare patterns and compose a pattern. This theory is widely applied in the development of rehabilitation applications, especially the rehabilitation from stroke and Cerebral Palsy. Annett et al. [4] develop a multi-touch tabletop system, the AIR Touch, which combines existing multi-touch technologies with a suite of new rehabilitation-centric applications. The applications are not only fun enough to increase patient motivation and engagement, but can also be customized to meet a patient's abilities and needs and record a variety of objective measurements. Li et al. [48] explore how tabletop games using tangible interaction can support the therapy of children with Cerebral Palsy, in which the training for elbow extension, supination and wrist extension is done by motivating children to use a hammer to hit an electronic board and rotate colored blocks. Dunne et al. [19] use wearable accelerometers to track and discourage compensatory movement strategies during therapy. Patients can interact with the tabletop using tangible objects marked by an optical tag while holding a foam ball to counteract involuntary finger flexion.

The convenience of communication in a group provided by TUIs can also benefit to the improvement of children's social skills. SIDES [59] is an application to take advantage of the tabletop environment to help adolescents with Asperger's Syndrome. They use DiamondTouch table [15], a multi-user touch sensitive tabletop with a top-projected display to develop a cooperative board game encouraging children's conversation.

2.3 TUI Implementation Technologies

Tangible interfaces always use a tabletop surface supporting multi-touch as base for interaction, and they have the capacity to recognize tangible objects, which can interact with the digital content on the tabletop. The most commonly used technologies involved with TUIs are computer vision based approaches using the image obtained from cameras to detect touch and objects, as well as radio-frequency identification (RFID) and micro-controllers.

2.3.1 Computer Vision

Computer vision based approaches are capable of sensing the position of multiple objects on a 2D surface in real time while providing additional information such as colour, size and shape. They can track specifically defined fiducial markers that are attached to physical objects, and sophisticated image processing algorithms are used for interpreting the markers. Since the computer vision algorithms are optimized for a specific marker design, the tag-based systems tend to be accurate and computationally cheap, so they are often used in the development of TUIs. Computer vision TUI systems typically require at least two components, a high-quality camera and a projector or a digital tabletop for providing real-time graphical output. A large variety of TUIs are implemented using tag-based computer vision. Examples include Urp [74], a tangible user interface for urban planning, the reacTable [39], a tangible electro-acoustic musical instrument, and Tangible User Interfaces for Chemistry Education [22]. There are also some TUIs that are not tagged-based, such as Designers' Outpost [44], a vision-based TUI for website design that is implemented using an extensive computer vision and image processing library.

Dohse et al. [17] enhance a multi-touch tabletop with an RGB camera to assign touch events to users and to improve touch detection reliability. They detect arms using skin detection, which leads to a limitation that it only works with bare arms. Also because it does not take use of depth data, so users should avoid overlapping hands, which is a concern in collaborative work. In [70], Takeoka et al. introduce Z-touch, which combines multi-touch and above the table interaction. Their system senses the approximate posture of fingers in the proximity of the surface using multi-layered infrared laser planes that are synchronized with shutter signals from a high-speed camera. The innovation point is to not only use the hand contact points, but also the angle of the hovering fingers to help control parameters in application. The drawback of this setup is that the algorithm for matching different laser layers has a limitation for tracking, so the angle estimation result is not reliable.

Wilson [77] presents a system using a depth camera mounted over the surface as a touch sensor. Depth cameras have been employed in top-view settings in combination with tabletop environments

Figure 2.4: dSensingNI [45] setup with a tabletop and a depth-sensing camera, capable to track users fingers and multi-touch interaction

in many projects. Depth cameras are applicable in tabletop environments where they have for instance been used to create height maps of objects on the surface [46, 76] or to sense user movement [30, 70, 77, 27, 45]. dSensingNI [45] is a framework for tangible interaction using a depth-sensing camera (See Figure 2.4). It is capable of tracking users' fingers and enables multi-touch interactions. Also, the computer vision algorithm can detect physical objects put on the surface, by assigning an oriented bounding box to each object, and use average height and volume to distinguish them. This framework can support object-object interaction (stacking, grouping) and object-hand interaction (grasping, moving, releasing). Hilliges et al. [30] present two different rear projection-vision tabletop setups to detect interaction above the surface. The configuration includes depth camera, projector, as well as a HoloScreen or a SecondLight [34] screen. They approximate the height of hands by their brightness in the image captured by the camera from inside of the table, and then use the fingers position to detect gestures, including pick up' and drop'. They also use shadow rendering to give users feedback about hand and arm position. However, both setups cannot reliably detect touches on the surface, and the algorithm for finger tracking is fragile in this situation, and errors can make objects harder to grasp or drop down

2.3.2 Radio-Frequency Identification

Radio-Frequency Identification (RFID) is a wireless radio-based technology that allows to sense the presence and identity of a tagged object when it is within the range of a tag reader. There are two types of RFID tags: active RFID tags, which contain a battery and thus can transmit a signal autonomously; and passive RFID tags, which have no battery and require an external source to initiate signal transmission. Most RFID-based TUIs employ passive inexpensive RFID tags and hence consist of two parts, a tag reader that is connected to a computational device and a set of tagged objects. The communication between a tag and a reader only occurs when both are close. When a tag is detected, the tag reader passes its identification number to the computer. The TUI application can then determine the application context and provide feedback.

Many TUIs are implemented using RFID technology. For example, Senseboard [35] is a TUI for organizing information using a grid that enables the placement of multiple tagged pucks on a white board that is marked with a rectangular grid; and Smart Blocks [24] is an educational TUI that computes the volume and surface area of 3D shapes built using tagged blocks and connectors. Martinussen and Arnall [54] discuss the design space for RFID-tagged objects, taking account of the size of tags and the shape of the emitted field.

2.3.3 Micro-controllers and Sensors

Micro-controllers are small and inexpensive computers that can be embedded in a physical object, pass the information received from the physical world to the digital world and provide haptic feedback to the physical world through actuators. Their sensors can capture a wide range of physical properties including light intensity, reflection, motion, acceleration, temperature and so on. There are many types of haptic feedback that can be provided, such as motion, attraction and repulsion. Physical Intervention in Computational Optimization (Pico) [57] is an interactive surface based on a tabletop that can track and use actuation to move small objects on top of it. It supports solving complex spatial layout problems through improvisation. The position of these physical objects represents and controls application variables. Users can constrain objects motion with hands and other objects, which leverage the user's mechanical intuition for how objects respond to physical force. The attraction and repulsion between objects are used in the Navigational Blocks project [11], which demonstrates a tangible user interface that facilitates retrieval of historical stories in a tourist spot. The system consists of a set of blocks with each face representing a query parameter. Each block is equipped with orientation sensors and electromagnets, which enables the blocks to attract or repel each other to give actuated feedback on the search result of the parameters configuration.

2.3.4 Comparison of Implementation Technologies

RFID approach can only distinguish objects' identity and presence. Compared with computer vision and micro-contollers approaches, the range of the physical properties it can sense is very limited, because a well designed algorithm in a computer vision based system can recognize objects' shape, colour and position, while a variety of sensors can be used to obtain information of objects' motion, acceleration, location, temperature, reflection, etc. Both RFID and micro-controllers are generally designed for high-performance systems because of their accuracy, efficiency and low computational

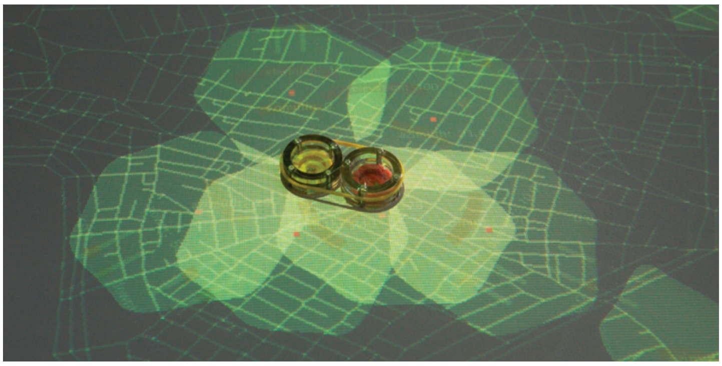

Figure 2.5: Two adjacent cellphone towers in the Pico [57] application. The computer is trying to separate these towers to improve the overall coverage on the map, but is unable to because the towers are physically attached by a rubber band.

cost. While the performance of computer vision method always depends on image quality and consistency of the environment, the tags needed for RFID and computer vision approaches are easily to be embedded in objects or attached to the bottom without altering their appearance. But using micro controllers, developers need to consider more about the position of the wires to minimize their effect on user's interaction and the size and shape of the objects are limited by the sensors and actuators inside.

2.4 Game Design for Children

This section first discusses related research result about the design of children's game, including the use of supporting tools and game design strategies. Then it gives a brief introduction to the resources of the educational games for pre-school children available in the market.

2.4.1 Related Research

Young children ages three to six now have a growing number and variety of digital games available to play on large screens, handheld screens, electronic learning systems, and electronic toys. Furthermore the time spent with games is growing. Current research is beginning to address the questions about what the effects of these games are and how the games for this age group should be designed to best serve the children's needs.

Debra [49] and Rideout et al. [61] have shown that well designed games can provide powerful interactive experiences that can foster learning, skill building, and healthy development in young children, while poorly designed ones can be time-wasting sedentary activities that contribute little to children's learning, skill building, or can even do significant harm to children's development, such as rewarding anti-social behaviour and instilling fear and anxiety. Garrison [23] points out that many digital games on the market are not developmentally appropriate or evidence based in design. They do not recognize or build on the ways young children play and learn. More research is needed to develop game design strategies that work well and are developmentally appropriate for young children. Game designers and publishers who care about the quality and benefits of their games could make use of these discoveries with the help and guidance from experts in the research and design of game-based learning for children. The experts could help turn the research findings into effective evidence-based game design strategies and game implementations, taking into account the learning goals of the game and the capabilities and interests of the target population of children.

Multi-touch tabletop technology offers a dynamic and appealing medium for designers to create collaborative learning experiences for children. Designing computer application for children is different from designing for adults [10]. Many researchers have addressed questions about the impact of technology on children. Methods for designing with and for children have recently become widespread in the design literature.

Halgren et al. [26] found that children click on any visible feature just to see what would happen, and they might click on it repeatedly if it generates sound or motion feedback. This behaviour is related to children's desire for exploration and should be motivated. But it also means that the game mode should be easily understood by young users. The Child Tangible Interaction (CTI) framework [5] is an explanatory conceptual framework that derives abstract design guidelines for tangible and spatial interactive systems from the literature on children's development of abstract cognitive structures. It describes five aspects of interaction: systems as spaces for action, perceptual, behavioural, and semantic mappings, and how systems can provide space for friends by supporting collaboration and imitation behavior. It recommends employing body-based interaction to support epistemic action, and to consider age-related perceptual, cognitive, and motor abilities, and children's understandings of cause and effect relations. The CTI framework further highlights how leveraging children's body-based understanding of concepts and spatial schema for more abstract concepts can provide learning opportunities.

Tangible objects have begun to be used in tabletop digital games over the last decade. Marshall [53] takes a critical look at the potential of tangible interfaces to support learning and presents a novel analytic framework derived from an analysis of work on both tangible interfaces and learning with physical materials. His work identifies a number of latent trends within research on tangible interfaces for learning, which are related to learning benefits, learning domains, types of activity, integration of representations, concreteness and sensory directness, and effects of physicality on learning.

Marshall criticizes that often information could just as well be presented graphically rather than tangibly, and that evaluations of TUIs often do not address the specific contribution of tangibility in terms of which elements of the TUI design are critical for learning. Furthermore, he argues that concreteness and physicality need to be distinguished (e.g., physical artifacts can be abstract), and points out potential negative side-effects of concrete representations, as these can result in decreased reflection, less planning and learning. Most importantly, this meta-analysis of the research area highlights the need for empirically demonstrating measurable benefits of physical manipulation for learning.

As Marshall states in his work [53], there are a number of reasons why the use of tangible interfaces may have learning benefit. One is raised by Triona et al. [73]. If perception and cognition are closely interlinked, then using physical materials in a learning task might change the nature of the knowledge gained. Related to this possibility is the emphasis in Piagetian developmental theory on the manipulation of concrete physical objects in supporting and developing thinking, particularly in young children. It is possible that because tangible interfaces often utilize physical manipulation, they might support more effective or more natural learning [72, 82]. A number of design-focused projects have suggested another learning benefit of tangibles, namely that tangible interfaces might be particularly suitable for collaborative learning. They can be designed to create a shared space for collaborative transactions and allow users to monitor each other's gaze to achieve interaction more easily than when interacting with a graphical representation on a display. They might also increase the visibility of other members' activity, better communicating the current state of their work [69] [21].

2.4.2 Educational Game Sources for Pre-school Children

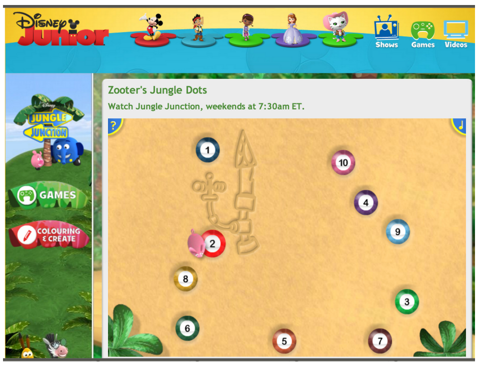

Figure 2.6: Zooters Jungle Dots online game for preschool children

Digital educational games for children are available in many market products, such as V.Smile, Leapster and InteracTV. The games' content covers a wide range, including vocabulary, math, health and science. They also have games designed especially for pre-school children. There are also some websites targeted to preschoolers, such as Disney Junior Games[38], Sheppard Software [68], and the Sesame Street [81]. Many of the games are hosted by popular preschool characters such as Winnie the Pooh, Little Einsteins, and Snow White. For example, Zooters Jungle Dots (Figure 2.6) is an online game developed by Disney Junior and is one of the most popular games on the website. Players can interact with the game using a mouse. The webpage also gives hints and sound feedback. Sesame workshop is a nonprofit educational organization behind Sesame Street. Their projects bring critical lessons in literacy and numeracy, emotional well being, health and wellness, as well as respect and understanding for children all over the world. With the contribution of Sesame Workshop, the games provided by Sesame Street are always well-designed and of high quality, but due to it's commercial background, the style and characters of the games are limited (See Figure 2.7).

Figure 2.7: Sesame Street Featured Games

Chapter 3

System Framework

This chapter introduces the hardware needed for our system, including a surface tabletop used for the game display and the base of the interaction, and a depth camera providing information for more natural interaction.

3.1 Microsoft Surface - PixelSense

Microsoft Surface is an 40" interactive surface computing platform that allows one or more people to use touch and real world objects, and share digital content at the same time. Microsoft Surface has the capacity to see objects placed on the screen in the near-infrared (NIR) spectrum. This technology of PixelSense enables Surface to see fingers and other objects placed on the screen, and also provides for the recognition of Surface tags that are marked with a specific pattern of dots.

Figure 3.1: The appearance of Microsoft Pixelsense

3.1.1 Features

Microsoft Surface has a sleek appearance and is only 4 inches in thickness, which enhances its potential to be used at home with children. With the rich colour saturation from a full HD display and a large screen, the device offers a compelling, immersive and visual experience that draws people in, and is a excellent tool for information sharing among a group of people because of the screen size and its support for multi-touch experience. This device also has the capability to recognize tagged object placed on top of it.

3.1.2 Components - Hardware

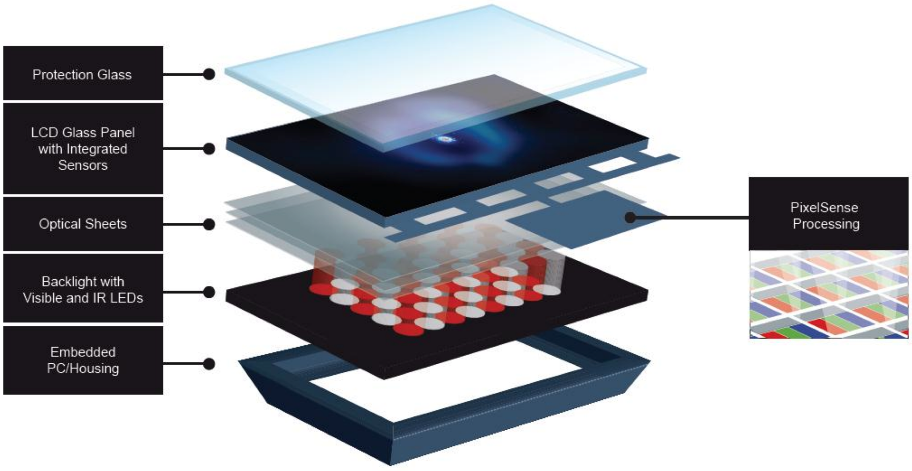

Figure 3.2: The hardware components of Microsoft Pixelsense

The key components that make up the device as shown in Figure 3.2

- The protection glass is used to protect the LCD panel with integrated sensors from getting damaged
- The LCD glass panel is used for display and the integrate sensors interpret the light reflected and convert it into electrical signals.
- The IR backlight unit provides light which hits the objects in contact with the device.
- The output of the processed image created by the sensors is fed into the PC to take any action.

The visible and IR LEDs light up the top LCD glass panel, which displays applications visible to the user. When there is nothing kept on the device or there is no one touching the top glass panel, the entire light just passes through. When a user touches the device or an object is put on the glass panel, the light is reflected back and the pixels just below the touch point or the object sense this via the integrated sensor. The light signal is converted into an electrical signal and the processing engine uses this information to create a picture of what is on the glass panel. Using image processing techniques, the engine then is able to figure out if it is a touch, a multi touch or an object placed on the glass.

The object recognition is managed via tag recognition. Any object that PixelSense needs to recognize should be tagged, and it is essentially the tag that is recognized.

3.1.3 Components - Software

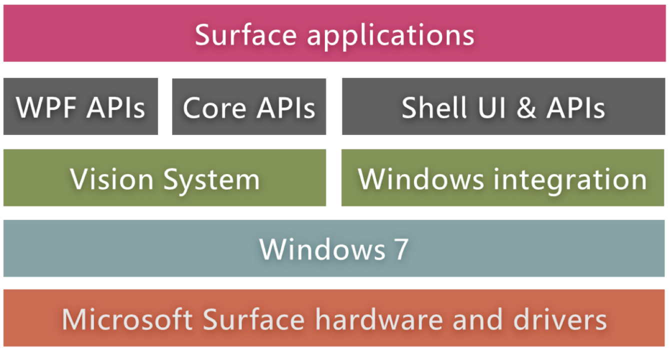

Figure 3.3: The software components of Microsoft Pixelsense

The Surface software components (Figure 3.3) enable developers to respond to multitouch and tagged objects in the applications. These include integration with Windows 7, the Launcher application, and the suppression of user interface (UI) messages from the Windows operating system and other programs running on the platform.

Microsoft Surface system runs on the Windows 7 64-bit operating system. Windows 7 provides all the administrative, security, and directory functionality of the Surface unit. Developers and administrators working with a Microsoft Surface unit have full access to Windows functionality.

The Vision System uses PixelSense^{TM} to process captured touch data by converting raw visual information into images, and the processing result can be accessed through Surface SDK APIs. PixelSense^{TM} allows to detect at each pixel on the surface when a person touches it or when someone moves a finger, tagged object, or untagged object over it. The Surface SDK also provides APIs that enable developers to create advanced applications that access raw images from the vision system.

Two sets of APIs are provided in the Microsoft Surface SDK: Windows Presentation Foundation(WPF) and Core layer, which are discussed specifically in section 3.1.4. The tight integration between Microsoft Surface and the Windows operating system provides system wide functionality on top of the Windows operating system. Developers can use this functionality to support unique aspects of the Microsoft Surface experience, such as managing user sessions, switching between the standard Windows user interface (Windows Mode) and the deployment experience (Surface mode), monitoring critical Microsoft Surface processes, and handling critical failures.

3.1.4 Applications development

Microsoft provides the Microsoft Surface 2.0 Software Development Kit (SDK) for developers to create NUI touch applications for devices with PixelSense and Windows 7 touch PCs. It includes two sets of APIs: the Presentation layer and the Core layer.

Presentation layer The Presentation layer uses Microsoft Windows Presentation Foundation (WPF), which is the standard choice for developing touch-enabled applications because it facilitates many development needs, including custom controls, touch-enabled by default, and many standard user interface (UI) elements, such as buttons, labels, and scroll bars. WPF also convenient the process to use custom graphics to create content-rich applications.

The Presentation layer includes a suite of Microsoft Surface-enabled versions of standard WPF controls and supports the use of XAML files for rapid UI development and exposes routed events and attached properties in a way that is consistent with the WPF model for mouse and stylus input. The Presentation layer also enables a multi-capture system so that one or more contacts on a Microsoft Surface unit can capture each UI element.

Core layer The Core layer exposes Microsoft Surface-specific contact data and events so we can create applications within a user interface (UI) framework. Applications can use the Core layer through the Microsoft XNA development platform, Microsoft Managed DirectX or Microsoft Windows Forms. XNA is commonly chosen to create dynamic and sophisticated graphics by supporting complex two-dimensional and three-dimensional rendering.

The Presentation layer is powerful, but it does not meet the development needs of this project. We want to optimize the process of object recognition, so the Core layer is the only option when the application requires access to raw image data from the vision system. Hence this project is developed using Microsoft XNA platform.

3.1.5 Tagged Object

The Microsoft Surface Vision System captures and processes raw images of a Microsoft Surface screen and can “see” the actual outlines of physical objects that are placed on the screen by using cameras that operate in the near infrared light (at 850 nm), and it also recognizes tagged objects that are marked with a special pattern of dots called tags.

Tags enable objects to drive the Microsoft Surface experience. The criterion for tag visibility falls into two basic categories: tag geometry and IR reflectance. The possible scenarios for tagged objects include printed paper, stickers attached to objects, and laser etched acrylic using IR reflective paint to illuminate the dot pattern of the tag. A tag consists of a specific arrangement of infrared reflective and absorbing areas.

There are other surface tabletops supporting the access to raw image, and many techniques can be used for objects recognition, such as those discussed in Chapter 2. Compared to other techniques, the use of tagged objects for recognition gives our application several advantages: Recognizing tagged objects is more efficient than processing raw images. The tags are small and well-defined, so specialized code in the vision system can locate them quickly, accurately, and efficiently. Each tag represents a distinct binary code value, so an application can distinguish one object from another, which is difficult to implemented using standard image recognition algorithms.

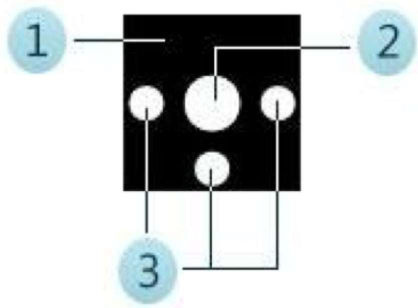

Figure 3.4: Tags used for object recognition: Geometry

Tag Geometry and Tag Values

Each tag that the Microsoft Surface vision system identifies has an orientation and an 8-bit tag value. Figure 3.4 illustrates the basic tag that all other tags are based on. A tag consists of

1. An infrared-absorbing or non-reflective (pass-through) background.

2. One infrared-reflecting circle (0.125-inch radius) in the center of the tag. This circle locates the tag on the Microsoft Surface screen.

3. Three infrared-reflecting circles (0.08-inch radius) located 0.28 inches from the center of the tag in each direction (left, right, and down). These ”guide” circles determine the tag orientation.

There are 256 possible unique tags that are encoded as a byte. The highest order bit (bit 7) is

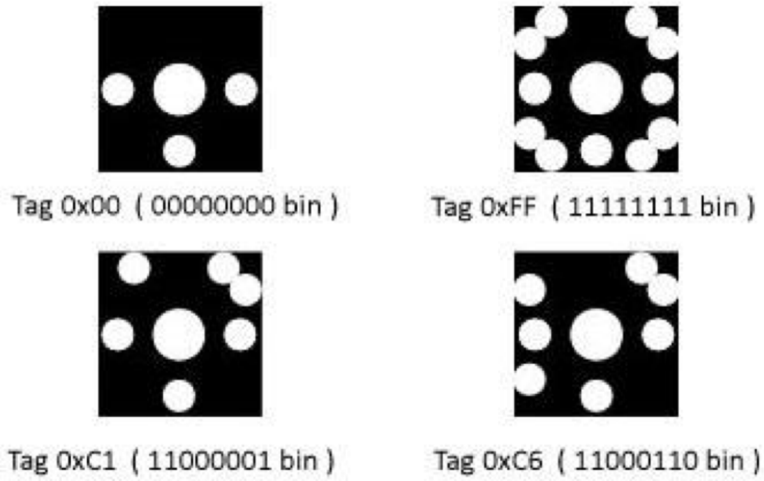

Figure 3.5: Tags used for object recognition

at the 1 o'clock position when you look at the printed side of the tag. Less significant bits are then read counter-clockwise from the 12 o'clock position. An infrared-reflective circle in a bit location represents '1'. The absence of a circle represents '0'.

3.2 Depth Camera

Many depth cameras can be found on the market. Here we introduce two commonly used ones, Kinect by Microsoft and Depthsense 311 by Softkinetic.

3.2.1 Kinect

A Kinect sensor is a physical device that contains cameras, a microphone array, and an accelerometer as well as a software pipeline that processes color, depth, and skeleton data. The physical device is shown in Figure 3.6 and consists of several components:

Color Sensor This video camera aids in facial recognition and other detection features by detecting three color components: red, green and blue. Microsoft calls this an "RGB camera" referring to the color components it detects.

Depth sensor An infrared projector and a monochrome CMOS sensor work together to sense the space in 3-D regardless of the lighting conditions.

Multi-array microphone This is an array of four microphones that can isolate the voices of the players from the noise in the room. This allows the player to be a few feet away from the microphone and still use voice controls.

Figure 3.6: The appearance of Microsoft Kinect

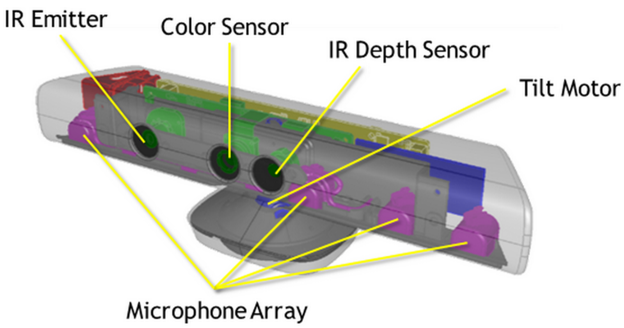

Figure 3.7: The inner structure of Microsoft Kinect

3.2.2 Depthsense

Figure 3.8: The appearance of Depthsense 311

DepthSense is SoftKinetic's time-of-flight depth sensor. The sensor measures the time it takes for infrared light emitted from the DepthSense 311(DS311) camera to return, i.e. the time-of-flight (ToF), and transforms the ToF positional data into real-time depth map images.

3.2.3 Comparison between Kinect and DepthSense 311

Aspect	Kinect	Depthsense
Frame rate (fps)	25-60	30
Depth field of view ($H \times V$)	57.3×42	57×43
Operating range (m)	0.15-1.0 or 1.5-4.5	0.8-4
RGB image resolution	640×480	640×480
Depth image resolution	160×120	640×480

Table 3.1: Comparison between depth cameras: Kinect and DepthSense

From Table 5.1 we can see that both cameras have a broad field of view and high resolution of colour images they generate. However the DepthSense camera has a very low resolution compared to Kinect. DepthSense has two operating modes, Near mode for distances from 0.15m to 1m, and Far mode for distances longer than 1.5m. Our experiments showed that the accuracy is acceptable only in the range from 0.3m to 0.8m, because the 0.5m-1m is the best it can get and the actual range depends on the sunlight and other illumination factors. In our project, the size of the tabletop was about 1.08m $\times$ 0.7m, which implies that the near mode of DepthSense was not enough to cover all the surface of the table. Another reason that we choose Kinect rather than DepthSense for this project is the the low resolution of the depth image, which does not permit accurate computation of hand position.

Chapter 4

System Implementation

This chapter discusses the implementation of the system that is used for the development of the tangible educational games for pre-school children. It is divided into three parts. First, we describe the approach to implement touch detection on the surface tabletop system. Second, the algorithm we used to obtain the position of user's hand through a depth camera is discussed. Third, we describe how we integrate tangible objects into our system based on the tag-recognition capacity of Microsoft Pixelsense.

4.1 Touch Detection

With the help of the vision based multitouch hardware, Microsoft Pixelsense enables each pixel in the display to detect when a person touches it. It is optimized for over 50 unique points of simultaneous touch events, affording opportunities to bring people together. In the Surface SDK, the variable type $ReadOnlyTouchPointCollection$ represents the collection of touch points on the interactive surface, through the traversal of all $TouchPoint$ of the collection, we can get properties of all touch events happened on the display.

4.2 Hand tracking

In the field of Human Computer Interaction, depth cameras have been widely employed to recognize gestural user input in games and interactive displays. For such applications, the users face the camera (front-view) to detect hands, arms and of the whole body, e.g. in the Xbox game station. In contrast, most of the existing tangible tabletop interfaces mount the depth camera above the tabletop facing the surface (top-view). In a top-view setup, a camera can monitor the whole tabletop environment and users dispersed around the table do not occlude each other.

We augmented the tabletop display with a depth camera, to track hands on or over the table using depth segmentation techniques, allowing touch points to be assigned the ownership necessary to support multiple users. Also, more feedback can be provided based on the tendency of arm and hands hovering time above the table.

We experimented with two different configurations, 1) Mounting the camera above the centre of tabletop, and 2) Mounting the camera on the short side of tabletop. In the first configuration, the most challenging issue is the IR noise. The detection of fingers and objects of the tabletop depends on the infrared illuminator and infrared cameras behind the display, which disturbs the depth map generated by infrared-based depth camera. For this reason, the coordinate of users hands (x and y axis) are accurate, but the depth (z axis) is not. The error is hard to be reduced using a statistical approach, because it is related to the environment illumination, the distance between hand and tabletop, and the non-uniform infrared emission of the display. In the case of a side-mounted camera, the height (z axis) and horizontal position (x axis) of hands are accurate, while the distance between hands and the camera on the table (y axis), which is calculated according to depth map and projection matrix, is not as accurate due to noise. In consideration of the purpose and need of this project, we choose to use the first approach in our system.

4.2.1 Model 1: Top-Mounted Camera

Most of the existing tabletop approaches mount a depth camera above the tabletop facing the surface (top-view). In the top-view setup, the camera can monitor the whole tabletop environment and users dispersed around the table do not occlude each other.

So far, infrared depth cameras have not been combined in a top-view setup with tabletop environments that are based on infrared illumination techniques. Therefore, the integration of such a camera into a tabletop environment poses a technical challenge since interference may lead to undesirable effects. The Kinect projects a pattern of infrared light on the scene and calculates the distances based on a disparity map, while the infrared light emitted from the surface of the tabletop may have an effect on the calculated result.

Figure 4.2 and Figure 4.4 show the difference of depth map generated from a depth camera with and without the use of Microsoft Surface tabletop. In Figure 4.2, there is some noise at the edge of the table and arms because with the IR reflection, which is commonly seen in the depth map generated from Kinect. The black dots indicate that the distance of these positions can not be calculated for reasons including illumination, reflection and computation error. The fingers of both hands can be seen clearly in this image. However in Figure 4.4, the depth image is different due to the interference of the IR tabletop. The background is very clear, because the distance from all the

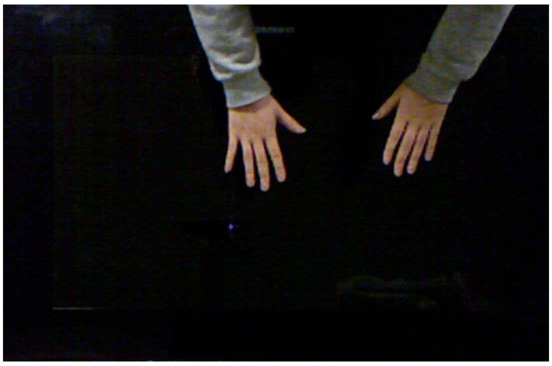

Figure 4.1: RGB images obtained from Kinect when the tabletop is turned off

Figure 4.2: Depth images obtained from Kinect when the tabletop is turned off

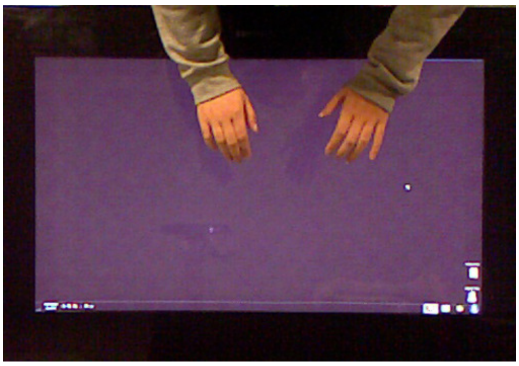

Figure 4.3: RGB images obtained from Kinect when the tabletop is turned on

Figure 4.4: Depth images obtained from Kinect when the tabletop is turned on

points on the tabletop is considered to be "missing". The working IR based tabletop disturbs the detection of the fingers (See Figure 4.4), and the information is not as complete as in the depth map without the surface tabletop. The loss of hand data is the primary reason for the inaccuracy of the hand position.

Contour-based Segmentation

Figure 4.5: Contours found in depth image after noise elimination

In order to detect interaction above the surface, we developed an algorithm that is optimized for the detection of arms, which works on the volume above the interactive tabletop. We obtain the depth map frames generated by the Kinect, with each frame consisting of 640×480 pixels, and then segment the depth image to locate the arms. The segmentation of arms is based on the depth images rather than the RGB ones. Each frame is segmented and the centers of non-contiguous regions in the image are extracted as hands locations. These locations are then compared to the known locations of previous frames and are assigned an identification number to distinguish different users interaction.

For segmentation of hands from the background, many approaches have been explored in the

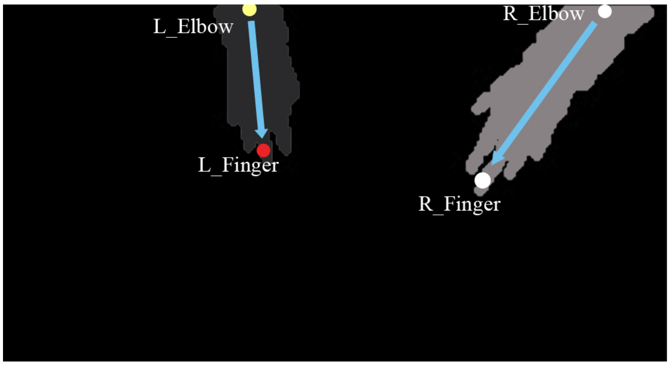

Figure 4.6: Arm information computed from the contours, the position of the fingers and elbows of both hands are marked in the figure, and the blue line indicates the direction of both arms.

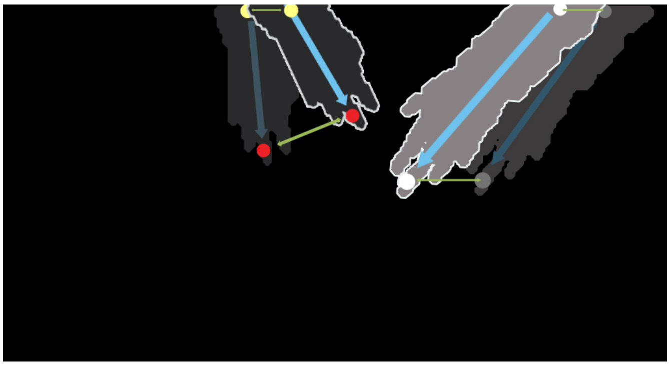

Figure 4.7: The comparison of distance between elbows and fingers when hands are moving, we can see that the moving distance of fingers are much longer than that of elbows.

past, such as segmentation based on skin color, motion detection, and background subtraction with previously trained background models. They all demonstrated relatively good performance, but all of them need several assumption on the scene content, illumination, motion and camera configuration. For example, Dohse et al. enhanced multi-user interaction with tabletop display through hand tracking [17]. In their approach, skin colour is used for segmentation. Due to the table often operating in low lighting and each application causing substantially different colors and intensities on screen, all the hand detection parameters had to be adjusted at runtime. To facilitate accurate tracking of hands, the application needed to maintain a reasonably constant color and intensity across the screen". The benefit of using depth information is that robust image segmentation may be achieved without posing any constraints on the environment or the users.

Having taken the correct rate and efficiency into consideration, our approach only uses pixels depth to segment the images. Hands are detected by creating a binary image of the pixels that fall within the appropriate depth range, which is from the camera to the surface of the tabletop. Each contour of the binary image is found using the cvFindContours function provided by OpenCV [56]. Very small contours are considered noise or belonging to objects put on the screen and are discarded. The remaining elements are assumed to belong to the hands and arms of users. Because the size of users' hands and the objects used in our project are known or easily estimated, we can find an appropriate threshold to distinguish the contours of noise and objects from the ones of arms. In the next step, because the coordinates of table edges are known, by iterating over the points in each contour, we can find the points having an intersection with the edges, which are assumed to be the "elbow" of the users, and the furthest points in the same contour, which are assumed to be the "finger tips". Then the direction of the arm can be computed by the "elbow" and "finger" positions. Although we define the intersection as the "elbow", these pixels may not represent the real elbow positions of the user. The intersection between the arm contour and the edges of the surfaces can represent any position of the arm. The length of the arm captured in the depth map has no influence on the hand tracking result, but if the intersection happens to be on the user's upper arm, the estimated arm direction is slightly incorrect.

This segmentation algorithm is similar to the approach presented in [17], but there only minimum and maximum Y-axis values were examined, which means users were required to stand on the long sides of the table, and the position of the fingers were used to distinguish different users. In our project, all sides of the tabletop are considered to be the edge where users extend their arms from. This gives users more freedom, and it is more suitable for the development of children's games. In addition, we use the "elbow" positions to deal with multi-user cases. From Figure 4.7, we can see that when hands move above the screen, the "finger" moves much further than "elbow". When we get two “finger” positions, it is difficult to determine the owner of each hand especially when the two hands are very close to each other. The “elbow” positions are better to be used as a measurement of the movement. As long as the users sit or stand still at the table, the intersection between their arm and the edge of the tabletop is fairly fixed. No matter where the hand position is, we can determine the minimum or maximum X/Y axis values and use them to assign each contour to a user. Another benefit is that, although users should avoid overlapping hands to prevent occlusions in the overhead camera’s view, it is possible to happen, especially with children playing. Using the finger position, it is possible to incorrectly associate a touch with a user when two hands are overlapping, and errors resulting from the current frame have negative effects on the analysis of subsequent frames. However, if one uses the “elbows” positions instead of the “finger” positions to maintain user identities, the error due to overlapping hands is reduced, and the subsequent frames are not affected even if the current assignment is wrong. Because a hand is continuously tracked, multiple touches created by the same hand can be treated as part of the same event. With this, touches no longer need to be treated as singular events, but can be given a history and associated with previous touches. After the segmentation, the “fingers” points and their users ID are passed to the game application.

Pixel-Based Segmentation

Sometimes there is no intersection between the contour of arm and the edge of the surface table in depth image received from the depth camera (See Figure 4.8b). This happens rarely with children playing, but when adults lean forward, their head may overlap their arms in camera’s view. Depth camera always have a fixed operating range. The Kinect can not detect object at a distance smaller than 0.8m. In this situation, the information about heads and bodies can be lost. Figure 4.8a has clear arms edge, and the depth data of hands. But in Figure 4.8b, the upper arms are too high to be captured, hence the finger data are lost due to interference.

The important part is how to segment the image and distinguish two hands. The segmentation algorithm proposed here is similar to the approach presented in [27]. Here, the segmentation of arms is composed of two steps. First, an initial segmentation is conducted by sequentially scanning the depth image once and assigning each pixel to a segment depending on the euclidean distances in the depth value. In the second step, small segments are merged into larger ones according to their adjacency. The segmentation result is shown in Figure 4.10.

We obtain frames with 640×480 pixels from the Kinect sensor. An initial clustering is obtained by sequentially scanning the depth image, classifying each pixel according to the distance from previously classified pixels in its neighbourhood. It starts from the arms side and scans from left to right. The background has been removed because the depth value of the surface table is known.

(a)

(b)

Figure 4.8: a) Depth map obtained from Kinect including the data of upper arms. b) Depth map obtained from Kinect loosing the data of upper arms

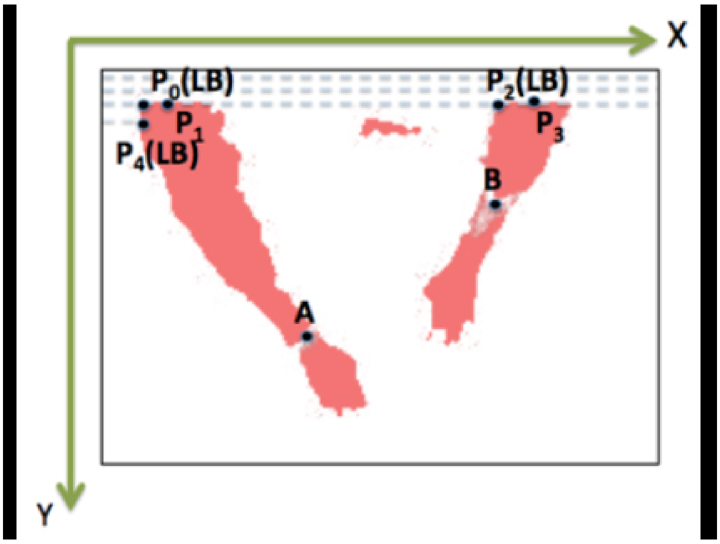

Figure 4.9: Scanning pixels of the depth map for segmentation, the pixels are separated into different clustering according to their distance to previous scanned pixels

(a) (b)

Figure 4.10: a) Segmentation result of the depth map with the data of upper arms. b) Segmentation result of the depth map without the data of upper arms

The clustering process is illustrated in Figure 4.9 and consists of the following:

- P_0 is the first valid pixel encountered when added to an empty segment and is marked as the left boundary (LB) of current segment.

- P_1 is the next valid pixel. It is on the same line as P_0, so we need to calculate the euclidean distance between P_0 and P_1. If the value is within a threshold (T = 50), we regard it as an adjacent pixel of P_0 and add it to segment #1. There may be a gap between "adjacent" pixels, such as in area B, so a threshold is used to test if the gap is acceptable.

- P_2 is the first pixel that does not satisfy the adjacency test, so it is the first pixel to be added to segment #2, with a mark LB (left boundary). Similarly, the pixels after that, such as P_3 are also classified. If the current pixel is the last valid pixel classified as belonging to this segment, it should be marked as RB (right boundary).

Segments consisting of fewer pixels than an empirically found threshold (T = 200) are removed to compensate for segments arising from the noise in the depth data and tangible objects put on the tabletop. After this clustering step, the number of small regions segmented may still be higher than the number of arms, so a merging step is needed. In [40], the merging step weights all three dimensions equally to calculate the inter-cluster variance and total scatter. Since arms have an elongated shape, two segments of the same arm, e.g., forearm and upper arm have a rather high inter cluster variance, whereas two forearms near each other may be merged. In our approach, we use the distance between boundaries segments as a measurement of the adjacency. The boundary pixels of segments are stored in the first step, and the euclidean distance between the pixels belong to two segments are considered to be the segments' distance. The segments with the least distance between each other are merged until the number of the segments is reasonable or the smallest distance between every two segments exceed the higher threshold of adjacency, so as to avoid the over-merging problem. But the merging will be continued if the smallest distance has not reached the lower threshold. In this project, the "reasonable" number of segments is set to two for single player and four for two players. The lower and higher thresholds varies according to representation of depth data, environment illumination and quality of depth map.

4.2.2 Model 2: Side-Mounted Camera

The interference between IR based tabletop and depth camera is the primary reason for the segmentation errors in depth map. In order to reduce the interference effect, another approach has been explored, namely placing the depth camera along the short side of the surface table instead of mounting the camera above.

Because of the limitation on view angle of the Kinect, the camera needs to be put much further away from the surface table if it is used on table's long side rather than the short side, which means that the volume above the surface will be represented by much fewer pixels. Due to the limitation of Kinect resolution, we chose to mount the depth camera on the short side of the tabletop. To estimate the hands' position, three steps are needed. First, we extract the user's arm from depth map. Second, the fingers coordinates are determined in camera view. Finally, the hand positions are projected into screen coordinate.

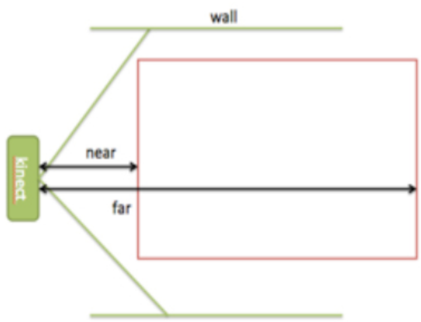

Figure 4.11: Relative position between the camera to the surface tabletop from the top view, near and far planes are used to cut-off extra space

To extract the area above the tabletop from the view of camera, we set the position of upper/bottom/left/right planes to cut off the useless parts of the view box, and the near/far planes are used to limit the object detection distance, so the detection can only be conducted in the centre area of the depth map, which is shown in Figure 4.11 and Figure 4.12.

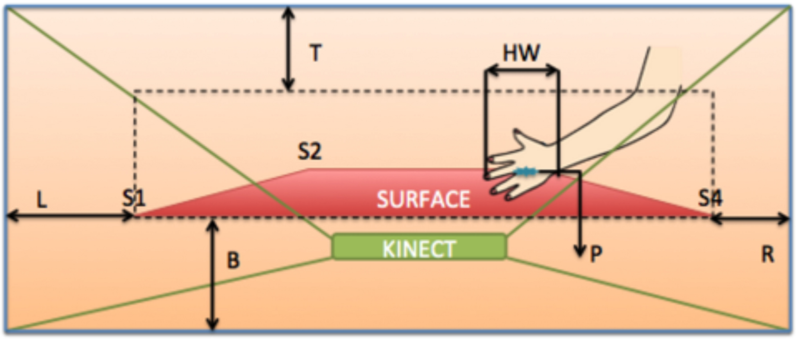

Figure 4.12: Relative position between the camera and the tabletop from the side view, T, L, B, R: distance between valid box on the top/left/bottom/right boundary of view box; S_1, S_2, S_3, S_4: corners of the surface; HW: hand width; P: track pint, i.e. the position of the hand

The segmentation algorithm is similar to the contour based segmentation discussed above. By iterating through the pixels in each contour we find in the depth map, we can get the intersection between the contour and the edge of the view box, and the pixels on the opposite direction of the contours are assumed to be the finger tips of users. Figure 4.13 shows the depth map obtained from a side-mounted camera when the tabletop is turned off while Figure 4.14 shows the one when the tabletop is turned on. The contour has an intersection with the view box at the right side, so the user's hand is supposed to be at the left-bottom part of the depth map. The orange parts of the figures are the extracted hand area and the green dots are the fingers position used in our application. we can find that in Figure 4.14, there is a white area in user's palm, which indicates that the distance between this part to the camera can not be detected due to the interference between IR based tabletop and the depth camera. This interference has a less effect on the detection result when the camera is fixed along the side of the tabletop instead of being mounted above the surface. The interference is strongest in the palm and wrist area due to the infrared reflection.

Because the surface is turned to a trapezoid in DepthSenses view, as shown in Figure 4.12, we need to compute the correspondent position in the surface coordinate system. The process is shown in Figure 4.15. The pixels coordinate received from the camera is $S(cx, cy)$, with the depth value of d, where cx is the value on short edge of surface, cy is the value on the y axis of camera's view, a.k.a, the height between hand to the surface, and d is the distance between hand to the camera. In the screen coordinate system, the short edge of the surface is regarded as the X-axis, and the distance between camera and the hand indicates the value on Y-axis. Point $S(cx, cy)$ is represented as $P(x, y)$, where $x = cx$ and $y = d$. The point $P(x, y)$ is not the accurate position of hand on the screen, because the shape of surface is no longer a square in camera view. We need to apply

$P'(x', y)$.

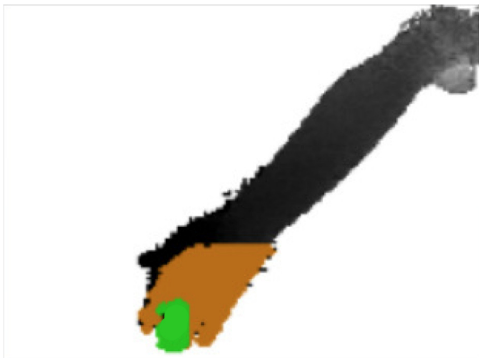

Figure 4.13: Hand detection when surface tabletop is turned off and the camera is mounted at the side of the surface

Figure 4.14: Hand detection when surface tabletop is turned on and the camera is mounted at the side of the surface, the white area is the data loss due to the interference.

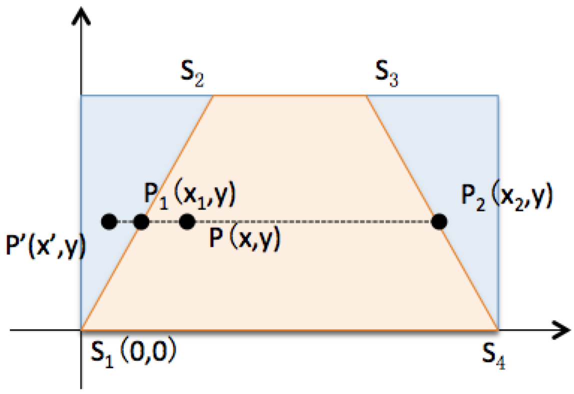

Figure 4.15: Projection from camera view to screen coordinate, the blue area represents the whole surface. The surface is projected to a trapezoid in camera's view represented by the yellow area. We need to project the hand position in camera's view $P(x, y)$ to the one in surface screen coordinate $P'(x', y)$.

some correction on x to project points in trapezoid to make them cover a whole square. We need to project $P(x, y)$ to $P'(x', y)$ as shown in Figure 4.15 where $P(x, y)$ is the hand position in the camera view which is the area in yellow, and $P'(x', y)$ is the real hand position on the surface which is represented in blue square. S_1, S_2, S_3 and S_4 are the corresponding points of the four surface corners in camera's view. To simplify the computation, S_1 is used as the origin of the coordinate. Then $P'(x', y)$ is computed using

$$x_1 = y_1 \times S_{2x}/S_{2y}$$

$$x_2 = y \times (S_{3x} - S_{4x})/S_{2y} + S_{4x}$$

$$x' = (x - x_1) \times S_{4x}/(x_2 - x_1)$$

The hand position in screen coordinate $P'(x', y)$ is sent to the application for further use.

Compared with mounting the camera above the tabletop, side mounting can reduce the effect of IR interference on the segmentation because only a little infrared emitted from the surface tabletop is reflected to the camera sensor. Another benefit is that the height of the hand above the surface is detected accurately. In the top-view configuration, the hand height is measured as the depth value in the depth map, which is not very reliable due to the noise and unstable depth resolution of the sensors. But in the second setting, the height of hand is known as y-axis position, which makes the prediction of the touch event much easier. As shown in Figure 4.16, the hand position we estimate is P_5, and the height of the hand h is known. If we assume that the arm can be regarded as a straight line, we can compute the angle of arm from the intersection point of the contour with the view box (upper arm) and finger's position. Then we can estimate the horizontal distance between fingers and the predicted touch event $d = h/tan(\alpha)$. If the finger information is lost in the depth map due to the interference and the hand's detection is not as accurate as we expected, the prediction of the touch event position is still can be close to the reality. For example, P_6 is considered to be the hand's tips, because the the depth data of the hand is lost. In this case, P_2 is the predicted touch position, which is close to P_1, the predict touch position when there is no interference.

Knowing the position of the forthcoming touch event is very useful. For example, we can use it to provide feedback to the users to avoid invalid behaviour. In a children's game, giving the children a little hint or warning before they make mistake, such as choosing the wrong result of the math problem, is good to keep them engaged as opposed to telling them what they did wrong afterwards. Also, the prediction can help reduce the computation cost of the process related to the position of touch event. For example, we can limit processing of data to a specific region to detect the touch instead of processing the whole screen.

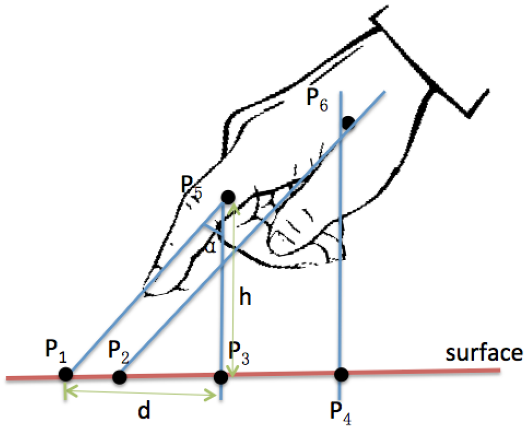

Figure 4.16: Prediction of the touch event based on the arm direction, P_1 is the predicted position when the finger's depth information is reserved, P_2 is the one when the data is lost due to the interference

4.3 Object detection

Microsoft Surface, the tabletop we use in this project, has the capability to detect and distinguish tangible objects put on the table. IR back light is reflected back from objects put on the table and sensed by the integrated sensors. A picture of what is on the surface is generated, which is called "raw vision data" (See Figure 4.19 in Section 4.3.1). Using image processing technique, one can determine touch positions and different contact types, including tag, finger and blob. The touch recognition process in the surface SDK returns the developer all the information related to the touch, including the type (finger, tag or bulb), position, value (for a tag), orientation, and so on. However, this process is very slow, between 0.5 and 1 second. This delay is obvious when an object is being moved, which impacts user experience. The reasons behind the delay are 1) unreliable functionality of the core layer of Microsoft Surface System, 2) light colour background, and 3) low quality printed tags. Figure 4.17 shows the movement track of an object put on it. The object is a stick made from clay with a tag attached at its bottom. Only when the value of the tag can be recognized by the application, the line is updated. To view the result more intuitively, it also draw dots when the object is recognized by the Microsoft Surface. When we move the object on the screen, we found that the track is not smooth, but a mixture of straight lines and curves. There are many gaps between dots, which means that during that period, the tabletop lost the recognition of the tag attached to the bottom of the object.

The object detection process is shown in Figure 4.18. The surface SDK runs a built-in image processing algorithm on the raw image and returns the result, which includes all properties of the

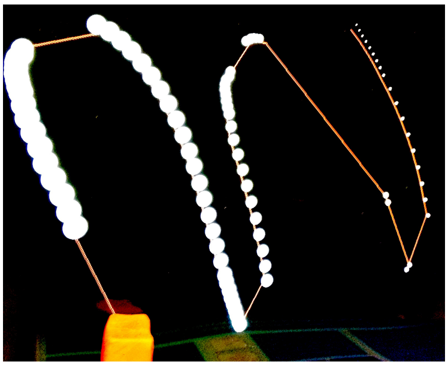

Figure 4.17: Slow response of tag recognition in object's movement, the straight lines indicate the points where the table lose the track of the object, and the dots are shrinking over time

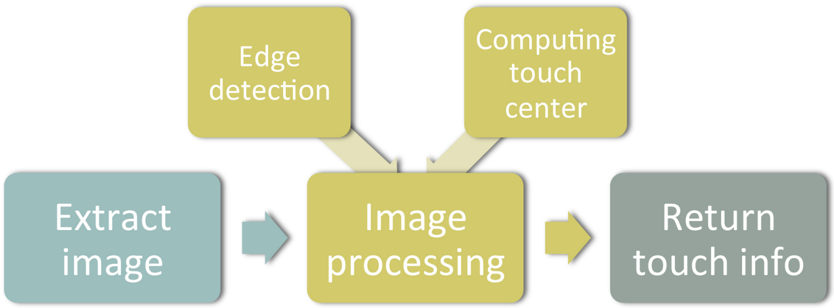

Figure 4.18: Optimized object recognition process

touch. To reduce the effect of the delay, we tried to optimize the object detection process in the project. This approach contains three parts:1) Retrieve of raw vision data. 2) Edge detection and clustering. 3) Touch computation.

4.3.1 Retrieve raw data

Microsoft Surface provides developers the access to raw vision data, which is the image generated directly by the sensors in the LCD glass panel. The raw vision data is a normalized image with 1920×1080 pixels, which is 8 bits per pixel, and can be rendered inside the XNA framework.

Figure 4.19: Raw vision data captured by Microsoft Surface when there are two tags on the tabletop

4.3.2 Edge detection

There are many ways to perform edge detection. The most commonly used approaches may be grouped into two categories, gradient and Laplacian. The gradient method detects the edges by looking for the maximum and minimum in the first derivative of the image. The Laplacian method searches for zero-crossings in the second derivative of the image to find edges. We only need to estimate the position of touch without caring about details, so the tolerance of noise and cost of computation of different methods are more important than fineness and detail kept during the detection. We used the Sobel operator due to its low cost computation and thick lines generated (See Figure 4.20). The operator uses two 3×3 kernels which are convolved with the original image to calculate approximations of the derivatives - one for horizontal changes, and one for vertical. If we define A as the source image, and G_x and G_y are two images which at each point contain the horizontal and vertical derivative approximations, the computations are shown in Figure 4.21.

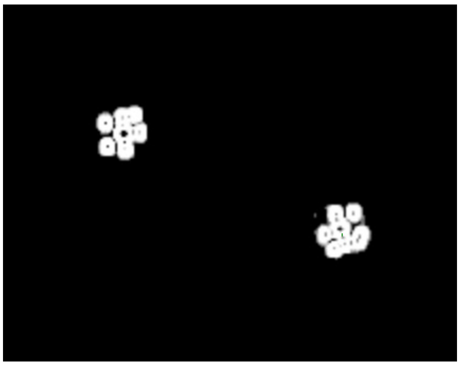

Figure 4.20: Edge detection result using Sobel operator based on the raw image of Figure 4.19

$$Gx = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * A \qquad Gy = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * A$$

Figure 4.21: Sobel operator used in edge detection algorithm, A as the source image, G_x and G_y are two images which at each point contain the horizontal and vertical derivative approximations, and $*$ denotes convolution.

4.3.3 Touch computation

After the edge detection step, the raw data is turned to a binary image with noise eliminated. If there is only one object on the surface, we only need to calculate the centre of all valid pixels and use that to be the position of the object. But if there are multiple object, it is necessary to classify all valid pixels into different groups according to their coordinates and compute the position of each object. By iterating through all valid pixels in the image, we compare the euclidian distance between the current pixel and the centre of all clusters. Because we know the size of the tags, the threshold of the distance comparison can be determined easily. We used tags of size $2cm \times 2cm$, and the threshold of 60 pixels. Then the weighted centres of each cluster is considered to be the objects touch position.

Through the first step, we can retrieve the raw vision data of the surface, which is a normalized image. Then we operate an edge detection algorithm on this image, to eliminate noise and find the active regions on the surface. Finally, we calculate the centre of each region as the positions of touch events. Through the comparison with previous detection results, we can identify each touch event.

This algorithm can update the position of objects much faster than the object recognition process

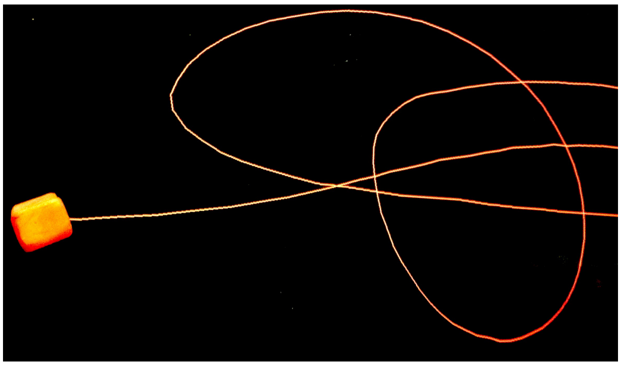

Figure 4.22: Single object recognition with optimized algorithm

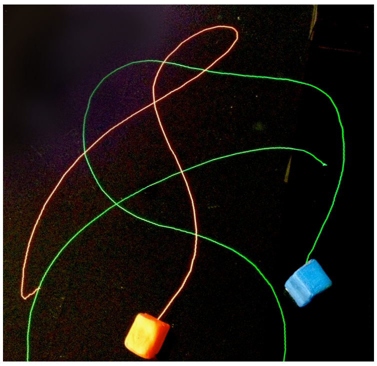

Figure 4.23: Double objects recognition with optimized algorithm

provided by the surface SDK, but it does not give any information about the identity of the objects. We still need the help of the surface SDK to recognize the value of the tag attached to the object, but instead of doing the recognition in every frame, we only need to do this once at the beginning of the object movement. Tags are identified at the beginning, and when players move the object, the identity of the object can be computed based on the relative distance from the positions detected of previous frames. Figure 4.22 and Figure 4.23 show the result of optimized object track approach.

Chapter 5

Game Design and Development

Digital games have the potential to be an effective and common method for intergenerational communication, entertainment, and education. New forms of tangible interaction interaction and supporting technologies, discussed in Chapter 2, can be designed to help children's overall development, providing an opportunity for their parents and guardians to offer guidance, and to improve intergenerational relationships. In order to design playful learning games, we need to understand how children interact with, and understand the representations embedded in tangible objects.

This chapter first summarizes relevant theories from developmental psychology, which should be considered in the design of tangible interaction to support playful and game-based learning. It gives a series of design considerations related to children's cognitive development and discusses how to apply these strategies into the design of educational games. Then it details how the puzzles in the games are designed to meet children's needs.

5.1 Game Design Concepts

Young children are intrinsically motivated to play and explore, and play is one of the main ways they learn [9]. They have an inherent need to exercise their emerging abilities and to make sense of their experiences. Play provides opportunities for exploration, experimentation, and manipulations that are essential for understanding basic concepts and constructing knowledge. Play also helps children develop imagination and creativity, which are key building blocks for future cognitive and emotional development and academic success. If materials for play and learning, including digital games, are designed to serve children's interests and abilities, their desire to play and explore, and their internal need to know, young learners will be more likely to develop and strengthen their initiative, curiosity, attention, self-direction, industriousness, competence, and love of learning [66, 9, 67].

Children in the age range of three to six are very diverse with respect to cognitive skills, physical abilities, temperament, and interests. They have less real-world experience than older children and adults, and they do not consistently distinguish between reality and fantasy, especially at the lower end of this age range [6]. For example, animated characters may seem real to them and magical superhuman powers may not seem impossible to attain. Young children tend to focus on the most salient attribute of a character or setting and usually do not factor in multiple characteristics or detect the subtleties of a character's motives or state of mind [67]. They are also likely to imitate what they see, without recognizing behaviours that might be dangerous or socially inappropriate. Children in the age range of three to six have little prior life experience that they can compare to the message conveyed in game or other media, they sometimes believe that fantasy could happen in real life, they do not always detect or properly interpret psychological nuances or underlying emotions, and they do not yet understand that some messages are designed to persuade or may be misleading or untrue [6, 43].

Tangible systems, such as tabletop games, have the powerful ability to engage pre-school age children in active learning. These new models of interaction can provide children with unique forms of learning. Healy [28] provides support for tangible, physically-based forms of child computer interaction when she states that body movements, the ability to touch, feel, manipulate and build sensory awareness of the relationships in the world are crucial to children's cognitive development. A conceptual understanding of these new forms of tangible interaction for children is needed.

Parents are an important element in young children's learning with digital games. They can share the experience with their children and serve as a guide. Several of the best digital games for young children provide supplementary materials for parents to help them a) learn more about the curriculum goals of the game, b) extend game-based learning into their children's daily activities (e.g., find specific letters in street signs and around the house), and c) be aware of the story lines and characters that have been used to teach these skills in order to help them engage in conversations about the game and what their children have been learning [50].

5.1.1 Learning Goals and Strategies for Young Children

Some of the commercially most popular educational digital games for young children (Sesame Street [81], Disney [38]) teach skills that are appropriate for the targeted age or grade level and are often based on well-established curriculum guidelines. For children ages three to five, educational digital games often focus on kindergarten readiness skills, including reading readiness (e.g. letter recognition, letter formation, letter sounds, simple spelling), math readiness (e.g. number recognition, number formation, counting, grouping), thinking and reasoning skills, perceptual skills, fine motor skills, skills of daily living (such as chores or hygiene), social skills, creativity and selfexpression, and under-standing of concepts (such as family relationships, emotions, healthy foods, safety, science concepts, music and art concepts, occupations, and so on). For kindergarteners and first-graders, the educational content and the mechanics of game play become more advanced, in alignment with established curriculum guidelines and developmental abilities, and so they offer more challenges and complexities.

A variety of instructional strategies can be incorporated into digital games for young children. Commercially popular educational games often fall short in this area, leaning too heavily on drill and practice, which is easy to program but does not take full advantage of the educational potential of the interactivity and experiential learning that digital games can provide [50]. Based on the literature we have discussed in Chapter 2, here we conclude some primary strategies proposed in relevant research that should be considered in children's game design.

1. Demonstrations. Games can provide clear, repeatable verbal descriptions and visual presentations of content. The child can choose to see a demonstration multiple times, as needed, and can be given options to find out more.

2. Stories. The plot of a story can engage the child in thinking about an issue or solving a problem, and it can maintain interest as the child wants to know how the story will end. Stories can conclude with a pro-social and constructive resolution to the problem.

3. Role models. Appealing characters, who are similar to the child and, therefore, especially interesting to the child, can demonstrate behaviour, enact a skill, or show effective ways to relate to others, modelling what the child could do in his/her own life.

4. Providing choice and building on success. In a digital game, a child can engage in creative activities, and in planning and building projects the child has chosen out of personal interest. The game can display the child's progress toward reaching the goal, and the pleasure of succeeding in the project can be inherently rewarding. For children ages five and six, who are beginning to develop attitudes toward learning in academic areas and self- concepts as learners, success in a game can contribute to positive attitudes and self-confidence [47].

5. Learning in a familiar context. Content, skills, and interactive challenges can be presented in a context that the child understands. For example, a digital game for science learning can be located in a playground, which is a familiar setting to many children.

6. Adaptive learning. Game activities should be capable to increase and decrease in difficulty to match the child's fast growing abilities and provide the right amount of challenge.

7. Interactive questioning. Characters can address the child and ask intriguing questions to motivate the child to think about the best answer and practice more. For children ages 5 and 6, the game could require that the child find and apply the answer before the game will allow him or her to proceed.

8. Challenges. Contests, races, mysteries, problem-solving, simulations, and other goal-oriented activities are highly motivating and can cause children to persist in a learning activity until they master it. There is evidence that children's attention to and involvement in learning activities tend to be stronger when learning is presented in a digital game, at least in the short term, compared to more traditional classroom learning approaches.

9. Repetition and rehearsal of skills. Games can be designed to provide extensive opportunities to rehearse and apply new skills. Young children enjoy repetition and are willing to persist, if the game's challenges are not beyond their reach.

10. Interactive encouragement and help. Game characters can address the child during times of inactivity and, when errors are made, to offer appropriate encouragement and assistance.

11. Performance feedback. Hearing spoken words of praise from game characters, receiving virtual items in the game for successful completion of a task, advancing to the next game level, receiving points, and seeing appealing congratulatory animations can show children their progress and achievement. Experiencing success this way can boost children's self-esteem and pride, instilling the self-confidence to engage in more challenging learning tasks. When children do not succeed, the game can provide appropriate hints, help, and support to improve knowledge and skills so that they can ultimately succeed.

12. Social interaction. Games tend to be social activities when opportunities for social interaction arise, and there is evidence that preschoolers will engage in cooperative social interaction while playing digital games in the classroom [55]. Games for young children can be designed to encourage cooperation and social interaction in specific activities and tasks to enhance learning.

During the design of the educational game for pre-school children in our project, we were always paying attention to follow these guidances to make sure our game is appropriate for the targeted age. Our game is set to a context that children are quite familiar with a zoo. We tried to make the game easily accepted by both children and parents through the mimicry of children's playing with parents or guardians to the experience of a family tour of the zoo in the daily life. The game has a completed story line to organize all the puzzles included, to maintain the interest as the child wants to know what happens next. The story plots involved in our game, such as offering help or finding friends, are tightly connected to children's pro-social activities. The story is narrated by a virtual character—a girl in pink dress, who seems to be of similar age as the players and provides feedback and hints throughout the game, following the strategy of interactive encouragement and help discussed above. The content of the puzzles included in the game was chosen to be suitable to the children's skills. There are multiple difficulty levels for each puzzle and they can be adaptively set according to the children's performance. Adaptive learning is appropriate to various children's skills development and children's fast growing abilities.

5.2 Choice of Puzzle Types

5.2.1 Exploration

Young children are intrinsically motivated to play and explore, and play is one of the main ways they learn [67]. They have an inherent need to exercise their emerging abilities and to make sense of their experiences. Exploration should be encouraged because it offers opportunities for making choices and decisions, using one's own ideas and imagination, as well as experimenting and trying out new behaviours. If digital games are designed to serve children's interests and abilities, their desire to play and explore, and their internal need to know, then young learners will be more likely to get engaged and love to learn.

5.2.2 Math

Learning about numbers is a preschooler's first step toward becoming a young mathematician. In preschool, math learning is all about counting, number recognition, and one-to-one correspondence. Coates [14] pointed out that flash cards are not the best method for teaching young children math skills with understanding. Many three-year-olds may be able to identify a number, but they don't understand what the number means. Instead of flashcards, counting things that are familiar to the child as she plays is more helpful. Counting, the ability to recite numbers in order and number and recognition, the ability to visually recognize and name numbers, are the basic steps of children's math skill development. After the children can handle these easily, practicing simple number operations and sorting are good ways to explore their potential. Because children's math skill varies from each other and can get familiar with numbers very quickly, even though they may forget it later, we developed two versions of the puzzle aiming to improve children's math skill.

5.2.3 Memory

One thing that all successful learners have in common is the ability to memorize easily and effectively. Helping the child develop memory skills will bring enormous educational benefits. These memory skills will continue to benefit them through their adult life. Memory games can also be referred as concentration game, which helps them to improve their memory by concentrating and focusing [13]. Playing memory games can be perfect for the children as they help to improve brain power. Classic memory games, e.g. matching pairs, really helps encourage children memory training as it takes a considerable amount of concentration for a young child to watch and remember where each card is [51]. There are many different choices for the pictures on the card. In this project, we use images of animals. First, this fits in the narrative environment and can be seen as part of the story as well as the previous puzzles. Second, animals are commonly seen in children's daily life and they are familiar with them. Children in 3 to 6 years old are not good at reading words, and they vary in math skill, so putting words or numbers on the cards may not be a good idea for this project. The last reason is that very young children might prefer these bright colourful versions of this game, and these images can help them develop pattern recognition skills.

5.3 Game Design: Content & Interface

After an analysis of children's cognitive and psychomotor skills, we designed and tuned a prototype game that is suitable for children aged three to six years old. Children had fun with several puzzles. 1) Exploration. Children are required to find the animals hidden in the scene with the help or the guidance on the screen. 2) Math. Children need to solve the math problem and use the tangible object prepared for them to find the right answer. According to children's math skill, one of two types of problems was chosen, one being about counting and the other about basic computation. 3) Memory. Children may get their memory practiced by trying to open the gates on the screen and match animals behind them. The main aim of the design was to create a fun game for preschool children what would motivate them to practice specific skills.

To test if our system is an effective tool to foster children's learning, there are two versions of this educational game developed for the children. One of the game provides feedback based on players' hands position while the other version does not provide feedback. The interfaces and content of these two games as well as the difference between them will be discussed in following sections. Some of the images used in the interfaces are downloaded from online resources websites [2, 1].

The "apple" mentioned in this chapter refers to a tangible object made for this game. The apple was printed in plastic using a 3D printer, and then painted using child-safe oil painting. The size of the apple is about $5cm \times 5cm \times 5cm$, which is not too big for children to manipulate nor too small to be safe for daily use. A piece of black paper with several white dots is attached at the bottom of the object, which is the tag to be used to distinguish the object by the surface table.

5.3.1 Main Interface

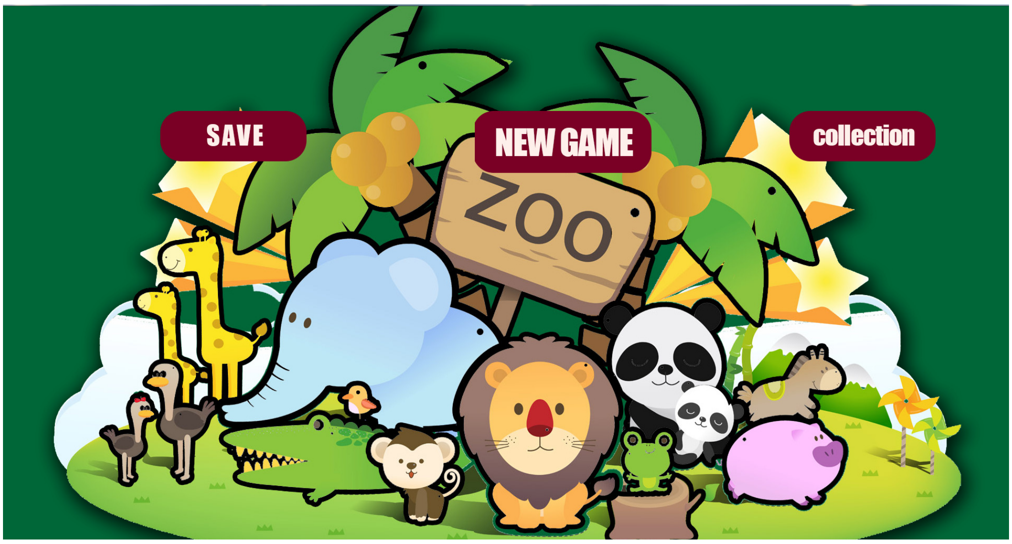

Figure 5.1: Main Interface of the game, providing the options of starting a new game, seeing the collection of animals, saving and quit the game.

When the experiment begins, the first thing users see is the main interface of the game, providing the options of starting a new game or quit the game. Because the content of the game is related to animals, the background of this interface consists of many different animals. This design had two purposes: 1) Let the children have a preparation about what they will see later. Small animals in cartoon style make children feel interested and excited about the game they will play. 2) Diverse animals showing on the screen are an opportunity for intergenerational interaction. Parents may lead children to talk and think by pointing to one of the animals and asking “What is this?”

5.3.2 First Puzzle: Exploration

The first puzzle is designed to motivate children to explore. Figure 5.2 is the interface of the game without feedback. Children are required to find the hidden animal in the South Pole. When the puzzle begins, the girl on the left of the screen tells a short story to the players, including the background of the puzzle and what they need to complete to pass this puzzle. This tutorial gives the guardian a clue about how to lead the child to play with a narrative story.

In the game without feedback, when user touch the ice cube on the screen, the ice melts and a part of the sea lion is seen. If user keeps touching the the ice left, the other parts of the sea lion are revealed (See Figure 5.2 and Figure 5.3). The script used in the tutorial of the first puzzle in game without feedback is as followed.

"Welcome to the South Pole."

"A new zoo is going to be build. BUT we don't have any animals!"

"Can you find some animals?"

Figure 5.2: Interface at the beginning of the first puzzle in the non-feedback game. The character is giving introduction of the puzzle.

In the game with feedback, the background is different. The environment is set to a forest and the user is required to move the stone to the right by touching the stone continuously for 3 seconds in total to find the leopard, and move away three leaves on the tree to discover the bird. During this process, the girl on the screen gives an introduction of this puzzle and feedback based on the child's performance.

When player does the right things to find the animal, such as touching the correct place on the screen, moving the stone or leaves away, the girl gives encouragement like "Good job!", "You are smart!" If the players have not found the right place to touch for a period of time, the girl gives some hints like "The stone is so huge!", "I heard something strange on the tree!"

When user completes all required touches and found all animals, a dialogue window appears on the screen (see Figure 5.5), asking children if they want to repeat this game or skip to next puzzle.

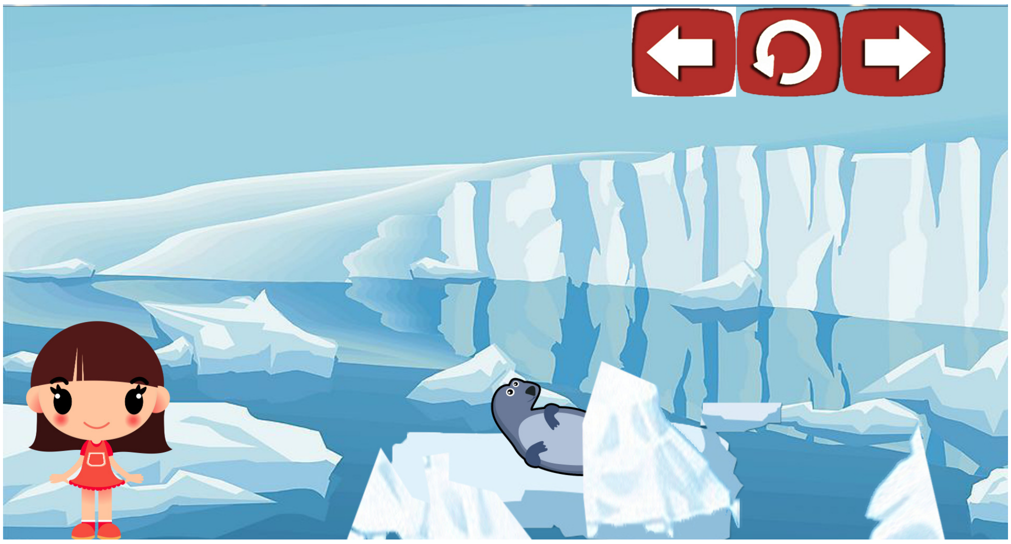

Figure 5.3: Interface of the first puzzle in the non-feedback game. The sea lion has not been completed discovered.

Figure 5.4: Interface of the first puzzle in the game with feedback. Children need to move the stone and leaves away to pass this puzzle.

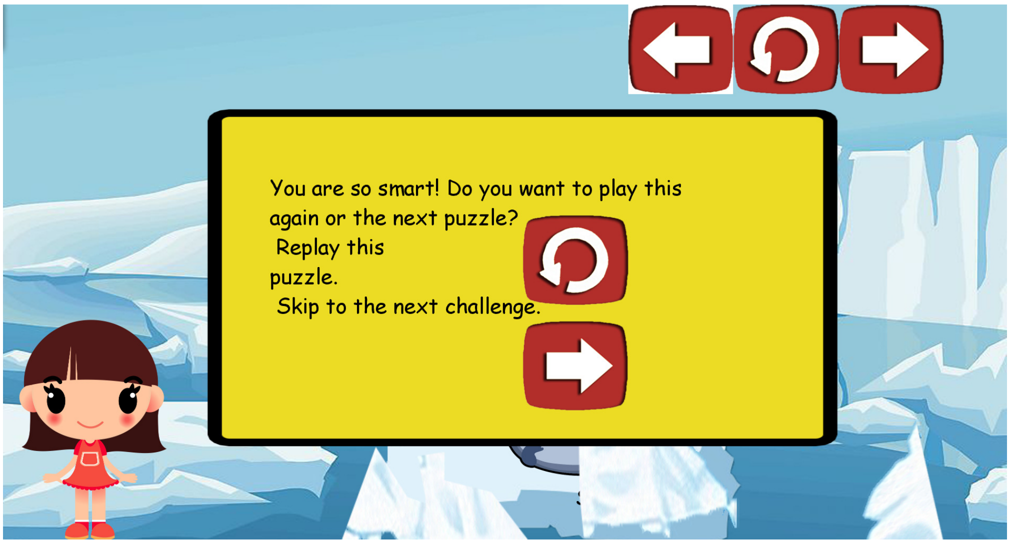

Figure 5.5: Dialogue window appearing when children finish the current puzzle, providing options of replaying this puzzle or skipping to next puzzle.

The buttons are the same as the ones on the right corner which enables users start over or skip to next puzzle during the playing of the puzzle.

5.3.3 Second Puzzle: Math

The second puzzle is designed to help children practice solving simple math problem. Because the math skill level varies a lot in different children with different age, we developed two types of this puzzle: 1) Counting. Players need to have the skill to recognize and count numbers from one to five. There are several plates displayed on the screen and each has a number assigned on it. Children need to put the tangible objects on the plates in order. 2) Computing. Players need to have the skill to compute the result of an addition or a subtraction of two numbers, with all operands and results being smaller than 10. The equation and several plates with numbers on them are displayed on the screen. Children need to put the tangible object on the plate with the result of the equation. Before the experiment, the experimenter can configure the game and choose the puzzle type according to child's age and preference. If the child or parent/guardian indicates that the puzzle is too easy during the experiment, the experimenters can switch the game type and the previous data are discarded.

In the counting version of this puzzle (See Figure 5.6 and Figure 5.8), the virtual character, Mr. Hedgehog, keeps asking the child to put the apple on the plate with smallest number. The plates in this puzzle are randomly distributed on the tabletop. When the child finds the correct plate, the plate disappears and the hedgehog moves to its position. Similar to the first puzzle, the girl in pink dress also gives a short introduction at the beginning for the second puzzle. The script is as follows.

"Mr. Hedgehog cannot find the way to the zoo!"

"Put the apple on correct plate, he will follow you!"

"Remember: only the plate with smallest number is the correct one!"

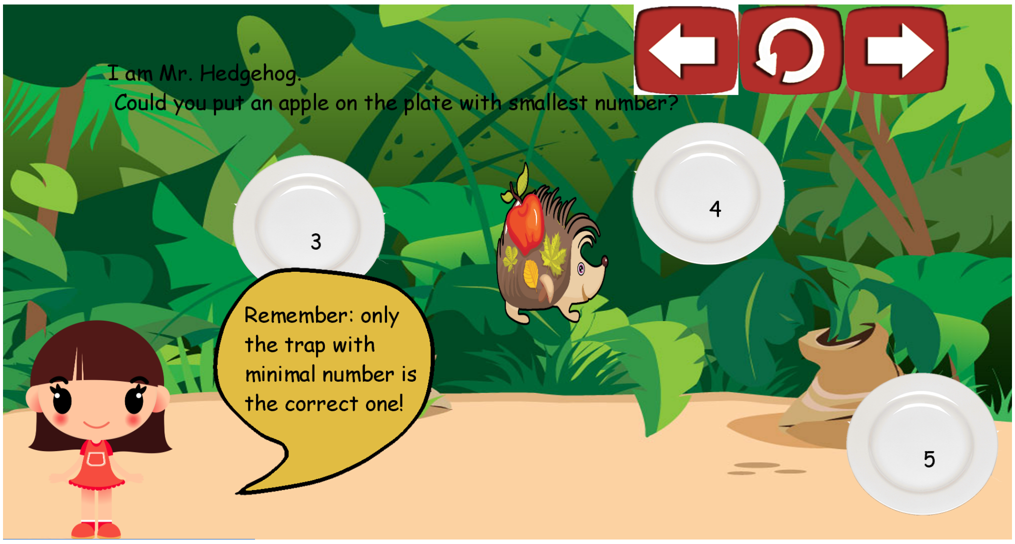

Figure 5.6: The counting version of the second puzzle in the game without feedback

For the computing puzzle (See Figure 5.7 and Figure 5.9), the story told is a little different. Mr. Hedgehog wants some apples and the child needs to figure out how many apples he wants through the equation on the top corner of the screen. When the child finds the correct result, the equation and the distribution of the plates are changed. One puzzle contains three trials and when the child was finished all three trials, a dialogue appears on the screen asking if they want to continue this puzzle or skip to next one. The script is as followed.

"Mr. Hedgehog wants some apples!"

"Do you know how many apples he want?"

"Put the apple on the correct number!"

When child puts the apple on the correct plate, the character on the screen encourages the child with "You are right. Mr. Hedgehog thanks you." or "Good job. You are smart." If the child gives a wrong answer, the girl on the screen says "This is not the plate he wants!" or "This is not the smallest number!"

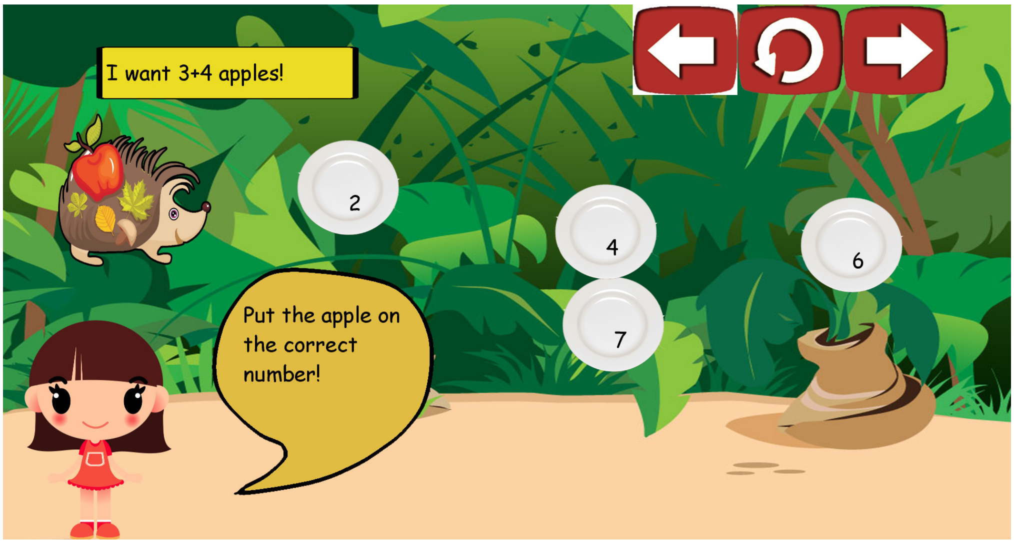

Figure 5.7: The computing version of the second puzzle in the game without feedback

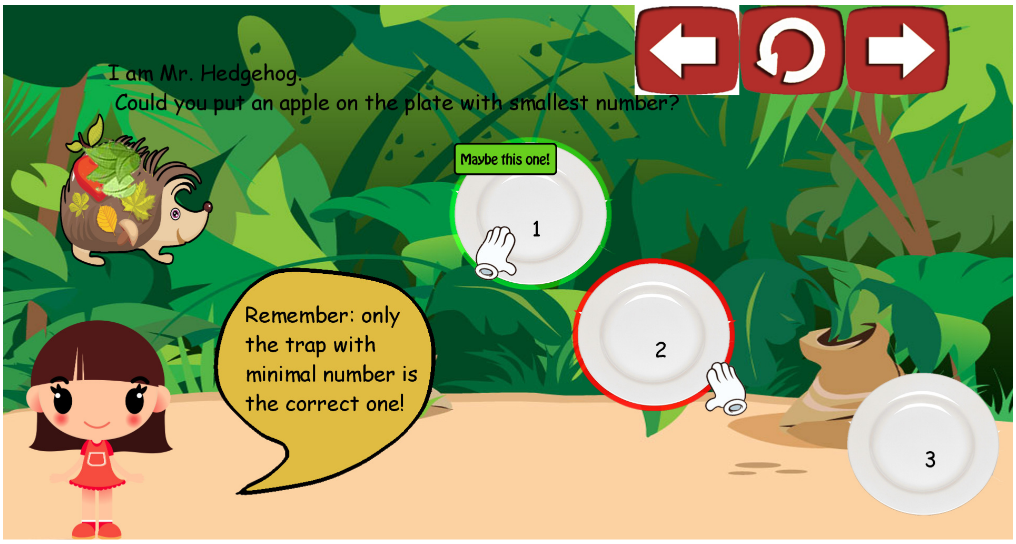

Figure 5.8: The counting version of the second puzzle in the game with feedback. If the child's hand is hovering over the correct answer, it gives a green signal, otherwise it gives a red signal.

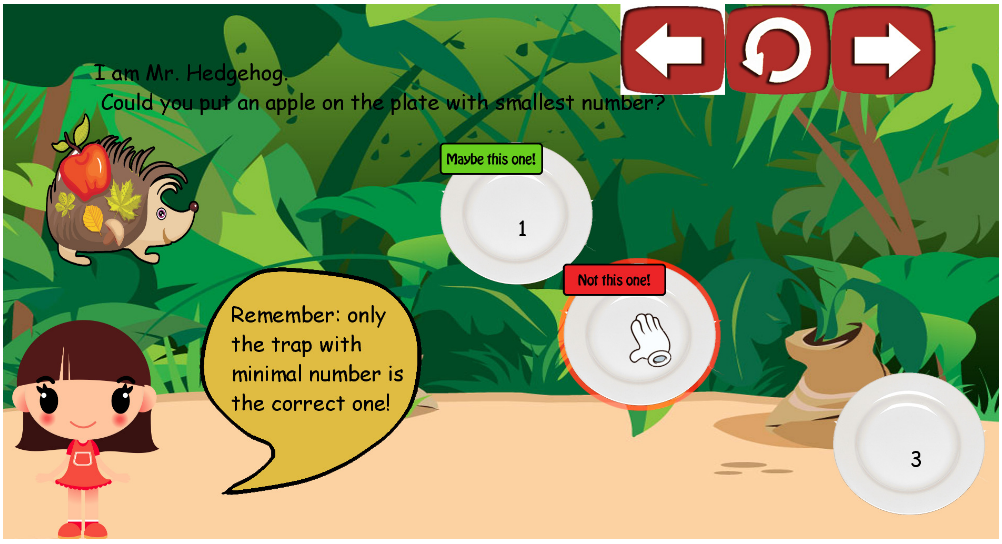

Figure 5.9: The counting version of the second puzzle in the game with feedback. If the child has not found the correct answer within 15 seconds, the game gives a hint saying “Maybe this one”, and if the child puts the object on the wrong plate, it gives warning of “Not this one”.

5.3.4 Third Puzzle: Matching

The third puzzle of this game is designed to help children practice short-term memory. It is an adaption of the classic ‘flipping cards’ memory game. Players need to remember the patterns on the cards and flip the pair of cards with the same pattern. In this game, the patterns are common animals, which children may haven seen in the zoo, a cartoon or a book. Instead of flipping cards, children need to touch the gates on the screen, then the gate opens and the animals behind the gate is seen. The gate does not close until the child opens another gate. If the animals behinds both gates are different, the guardian has an opportunity to interact with child. If the child has never seen an animal before, the parent can explain what it is and what kind of environment it should live in. If the child knows the animal already, the parent may ask “Do you remember what is his name?” or give guidance like “We have seen it in this puzzle, do you remember which gate he is behind?” If the child opens two gates with different animals, both gates close after two seconds. Otherwise the animals disappear and the gates keep open until the end of the puzzle.

At the beginning of this puzzle, the girl in pink dress also gives an introduction to the parent and the child. The script is as follows.

“Animals are hiding behind the door!”

“They want to play with their friends.”

“Can you open the correct door and help them find friends? ”

When the child opens two gates with same animals behind, the girl in the game encourages the player with "Yes! They are the best friends." Otherwise she says "No, they don't know each other." After the child has matched all the animals and open all the gates, this puzzle is completed. The dialogue provides options for repeating the puzzle or going back to the main menu. The child can choose which option to take.

Figure 5.10: Third Puzzle in Non-feedback Game, allowing children to find the gates with matching animals

5.4 Difference between two versions of game

In the game with feedback, the most significant difference between the first game is that users can see the hand-shape signs on the screen which indicates their hand position above the tabletop. Some visual feedback is given according to the position, which makes the game easier to play and more interesting.

The content of the first puzzle of this game is different from the other game, because if the children have known where the animals are hidden, they will feel bored when they are faced with exactly the same puzzle and will spent much less time on the second game they play, no matter which game it is. To make sure the data collected is effective and the bias are eliminated as much as possible, the background of the environment and animal species are changed, but the basic approach of finding those animals is the same as in the game without feedback. For the second puzzle, the game can give two kinds of feedback: 1) when children are hovering above the correct plate, the plate has a green stroke, if the plate being hovered is the wrong one, the colour of the stroke is red.

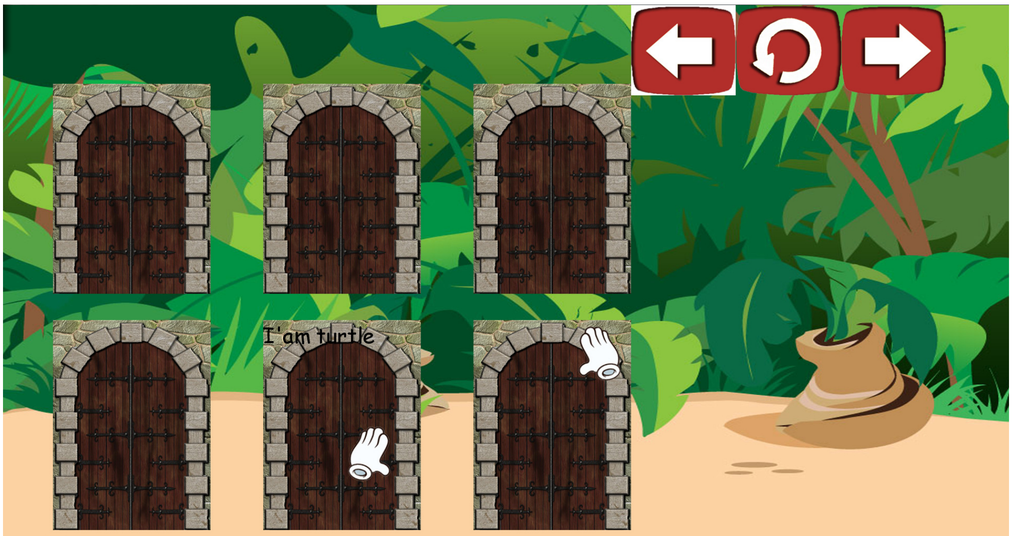

Figure 5.11: Third Puzzle in Game with Feedback

The brightness of the stroke changes with the distance from hand to the plate. When the hand moves nearer to the plate, the colour of the stroke becomes darker. 2) If the child has not found the correct answer within 15 seconds, a hint is shown on the correct plate. If the child places the object on the wrong answer, the plate they choose is marked by a red warning. Similar to the second puzzle, the third puzzle in the game with feedback also provides some hints when children get stuck in the puzzle and guidance when the child tends to choose the wrong answers.

5.5 Game Design: Adaptive Difficulty Levels

Puzzle	Variable	Values			Requirement
		level 1	level 2	level 3	
Counting	Number of Plates	3	4	5	Errors≤1, Duration≤30s
Computation	Range of numbers	Addition with addends and result ≤5	Addition& Subtraction with numbers ≤5	Addition& Subtraction with numbers ≤10	Errors≤1, Duration≤40s
Memory	Number of Animal Pairs	2	3	4	Errors≤ N_{pairs}×2 - 1

Table 5.1: Different difficulty levels of three puzzles in this game

In the second and third puzzle there are three levels with increasing difficulties. Specifically, in the counting version of the second puzzle, the lowest level has only three plates on the table for the children to count, which means it can only teach the child how to count from one to three. And in the highest level the child needs to count from one to five to complete this puzzle. In the computation puzzle, the child needs to know the addition of numbers smaller than five to finish the lowest puzzle and will learn the subtraction of these numbers in the second puzzle. In the third level, the child sees the questions about numbers smaller than ten but bigger than five, such as the result of 8-1 or 7+2. Children are supposed to practice their work memory in the third puzzle, which also has three levels with 2, 3 and 4 paired of animals for children to match. In order to show the puzzle with most appropriate difficult level to children, we need to know if the current level is too easy or too hard for the child, and if the child has got enough practice of current content. So we set a requirement for each puzzle. The game records the child's behaviour during each trail of the puzzle, and if it satisfies the requirement, which means the child made few enough errors or completed the previous trail quick enough, the child is challenged with the puzzle of next level, otherwise it needs to play the puzzle of same level again. All the difficulty and requirement settings are decided based on the development of children from 3 to 6 years old and a test before the experiment.

The purposes of this design are listed as followed. 1) Make the game suitable for children with different skill progress. 2) Keep the puzzle always be a challenge for the children and reduce their tiredness. 3) As a measurement to evaluate the effectiveness of the system proposed in this project.

Children's math and reading skill vary substantially. Age may be a main factor of child's development, but is not the only one. Most of 3-year-old children can count from 1 to 3, but only a few 4-year-old children can do simple addition, while others feel it hard to understand the concept. Children's math and reading skill development are related to factors, such as age, education received, daily practice, preference and parent's requirement. In children's playing, a game that is too easy makes them feel bored and a game that is too hard overwhelms the children easily and they will lose the motivation to play. For this reason, it is necessary to have games with different levels to choose for preschool children.

Children learn very fast. Even if they find the counting from 1 to 4 is very challenging at the beginning of this game, they will learn how to do that and get familiar with number 4 very quickly with their parent's help. And if they find the puzzle still asks them to do the same job when they repeat playing the puzzle, they will soon feel bored and think that "there are no more interesting things in this game". To keep their engagement, we need to challenge them with more difficult question and something they are not familiar with as soon as we think they have got enough proficiency of the current level. So when the children's performance satisfies the requirement of the puzzle, more plates with bigger numbers, more gates or harder math questions are put on the screen for further practice. The location of plates and the animals behind the gates always changed, no matter if the child has finished the puzzle well or not.

5.6 Game Design: Game Usage

Data about young children's digital game usage are a moving target, as their electronic game play changes when new games and technologies become available and as parents' attitudes change about the acceptability of specific types of games. A few recent studies have examined children's time expenditure. One study found that on a typical day, children spent an average of 6 minutes playing digital games at age 3, 10 minutes at age 4, 12 minutes at age 5, and 8 minutes at age 6 [3]. The study also found that 48% of children ages 6 and under have used a computer, and 30% have played digital games; and children ages 4 to 6 spend a little more than one hour per day using screen media other than TV. An earlier study found that children ages 2 to 3 spend an average of about 17 minutes per day on a computer, 19 minutes per day playing digital games, and 5 minutes per day using the Internet, for a total of about 40 minutes per day with these three forms of screen media other than TV [37].

Chapter 6

Experiment

This study presents an exploration into how tabletop interactive games can support pre-school children's learning. After an analysis of children's cognitive and psychomotor skills, we have designed a game with three puzzles that are suitable for children in the age range three to six on a Microsoft Surface tabletop. To test if our system is an effective tool to foster children's learning, we conducted an experiment, in which children played the game with the feedback based on players' hands position as well as a version of the game without any feedback.

During the game playing, the interaction and communication between the parent/guardian and the child was encouraged. It was anticipated that this game cannot only teach children skills, but also provide a motivating and supportive experience of intergenerational communication. Parents can give hints and guidance when the children are stuck in the puzzles and offer encouragement when the children succeed. One benefit of a large tabletop over a handheld tablet and mobile device is that children can be aware of the existence of their accompany and have the chance to communicate with them. This is good for children's social skill development. A collaborative game can lead children to think about others. In this game, children can explain the animals to their parents when they know the animals that appear in the game and can ask for help if they feel it hard to solve the puzzle.

6.1 Participants

This study was approved by the Research Ethics of University of Alberta. All english-speaking children with age three to six accompanied by their parents or guardians were allowed to participate in this experiment. The study focuses on pre-school children's education, so children younger than three years old or older than six years old were not qualified. There were 25 participants in our study, with 9 boys and 16 girls. The youngest child was three years and four months old, and the oldest one was five years and seven months old.

6.2 Experimental Setting and Procedure

The experiment was carried out in the Advanced Man Machine Lab of the Computing Science Centre. The children participated in the experiment under the supervision of a parent or a guardian. They were allowed to freely play with the application developed on the system described in Chapter 4, i.e. a large tabletop system that acts as a touch screen and is capable of recognizing objects put on the surface. A depth camera mounted above the surface computed the position of the children's hands.

Figure 6.1: Child and her parent are playing the game in Early Learning Centre

Each child played two games: the experimental game, which provides the feedback based on hand position and the control game, which provides no feedback. Each game contained three puzzles: 1) Exploration. Children are required to find the animals hide in the scene with the help of the guidance on the screen. 2) Math. Children need to solve math problems and use a tangible object to find the right answer. According to children's math skill, they were given one of two types of problems, one about counting and the other about basic computation. 3) Memory. Children may get their memory practiced by trying to open the gates on the screen and match animals behind them. All the material of the game is positive, and the images are drawn in cartoon style. The experiment lasted for 10-20 minutes, during which children and their parents played with the application freely and parents could give guidance and interact with the children.

The experimenters explained the purpose of the experiment and gave a short introduction of

Target	Variable	Comment
For every repetition of a puzzle	$Duration(s)$	Time spent for every repetition of this puzzle
	$Error$	Number of errors made in this repetition
	$Hint$	Number of hints given in this repetition
	$Level$	Difficulty level of this repetition
For every puzzle	$Repetition$	How many times children repeat the current puzzle
	$Duration_{avg}(s)$	Average time spent on each repetition of this puzzle
	$Error_{avg}$	Average errors made on each repetition of this puzzle
For every game	$Duration_{game}(s)$	Time spent on this game

Table 6.1: Quantitive data recorded for each participant

the puzzles at the beginning of the experiment. Consent forms were signed after all problems were explained and all questions had been answered. Then the participants were let to play both games freely for 10-20 minutes. Once a child began playing the games, the experimenters did not interfer any further. The experiment was discontinued whenever the children indicated that they wanted to quit playing the games. At the end of the experiment, the experimenter made a short interview with the parent or guardian.

6.3 Results

All the quantitive data needed in this study was recorded by the system while the child played each game. Table 6.1 shows the data that was collected, including the time spent on the puzzle (counted in seconds), the number of errors children made during the game, how many times children repeated playing each puzzle and so on.

6.3.1 Quantitive Analysis

Every participant needed to play two games, one was the experimental game providing the feedback based on hand position, the other was the control game without any feedback. Children may feel bored when they are faced with the second game they played, no matter what kind it was, and were not willing to spend much time on it. There was also the possibility that when children played the first game, it costed them some time to get familiar with the device and the rule of the game. All these factors may have an effect on our result. To eliminate the bias, the children were randomly assigned to two groups that played the two games in different order.

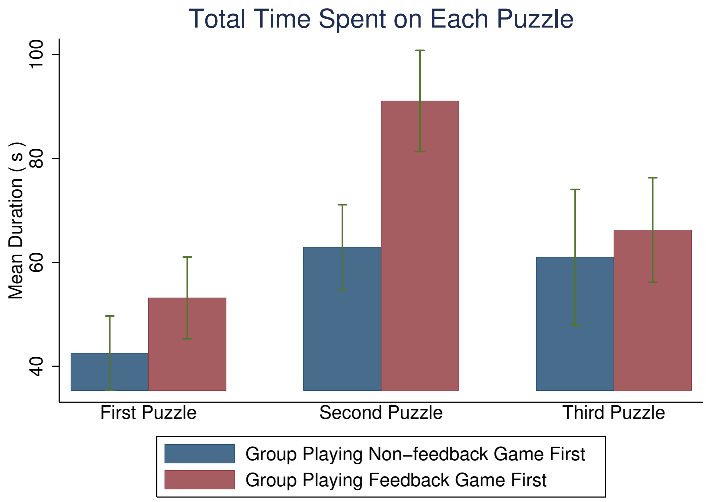

Figure 6.2: Time spent on the game without feedback between the two groups playing both games in different order, with error bars showing the standard errors.

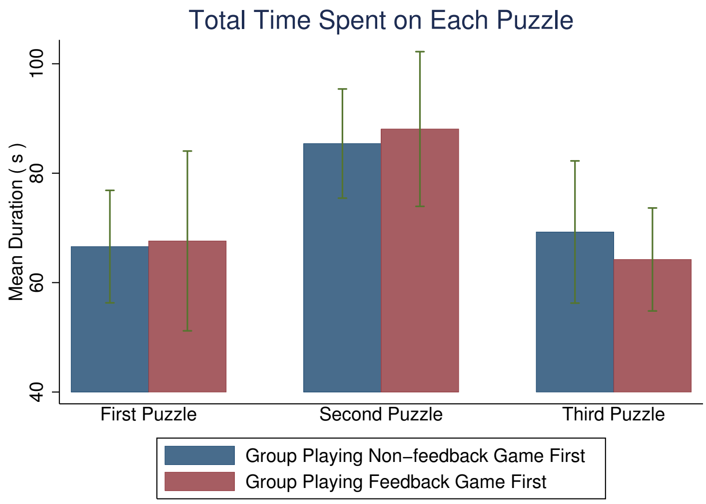

Figure 6.3: Time spent on the game with feedback between the two groups playing both games in different order

All participants were divided to two groups. One group first played the game without feedback and then the game with the feedback. The other group played the games in reverse order. The result (See Figure 6.2) shows that in the game without feedback, the two groups differ in the time spent on the second puzzle $(t(23) = -2.19, p < 0.05)$. There was no difference for the first puzzle $(t(23) = -0.99, p = 0.33)$ and third puzzle $(t(23) = -0.32, p = 0.75)$. Figure 6.3 shows the time children spent in the game with feedback. There was no difference for all puzzles (all $t(23) \leq 0.32, p > 0.5)$. The results indicate that the order of the two games has a very limited effect on the time the children were willing to play.

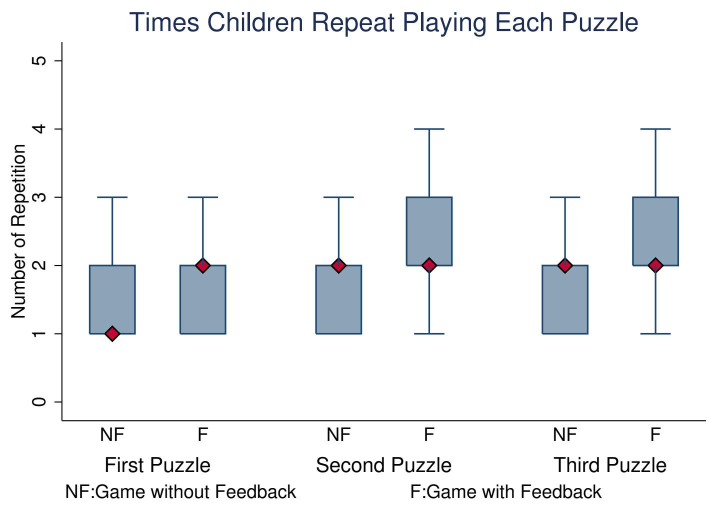

Figure 6.4: Number of children repeat playing each puzzle in both games

Figure 6.4 shows how many times children were willing to repeat each puzzle. The box indicates the lower and upper quartiles and the median is represented by a diamond symbol. The bars indicate the data range. A Wilcoxon signed rank test revealed a significant difference between the no-feedback and the feedback group, $z = -2.42$, $p < 0.05$ for the first puzzle; $z = -3.342$, $p < 0.001$ for the second puzzle; and $z = -2.56$, $p < 0.01$ for the third puzzle. This shows that children were more engaged in the game with feedback and wanted to play these puzzles more often than the traditional game without feedback, especially the second puzzle about math training and the third puzzle practicing children's work memory.

Children can make mistakes in the second and third puzzles by putting the apple on a number that is not the result of the equation or opening two gates whose animals do not match. The difference between the two games is that the game with feedback can offer guidance, when children move their hands and during their thinking. Figure 6.5 shows how many errors were made in each puzzle

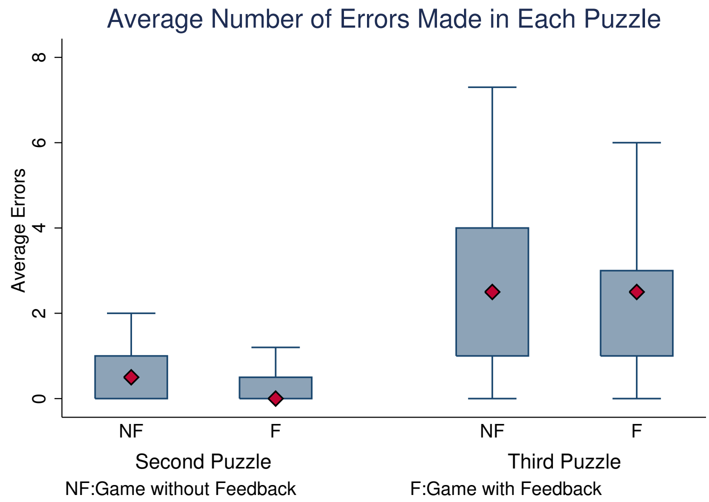

Figure 6.5: Median errors made in each puzzle of both games

of games. For the second puzzle, the result of Wilconxon signed-rank test ($z = 2.15, p < 0.05$) indicates that children made significantly fewer mistakes in the feedback condition than in the non-feedback condition. But in the third puzzle, there is no significant difference ($z = 0.51, p = 0.6$).

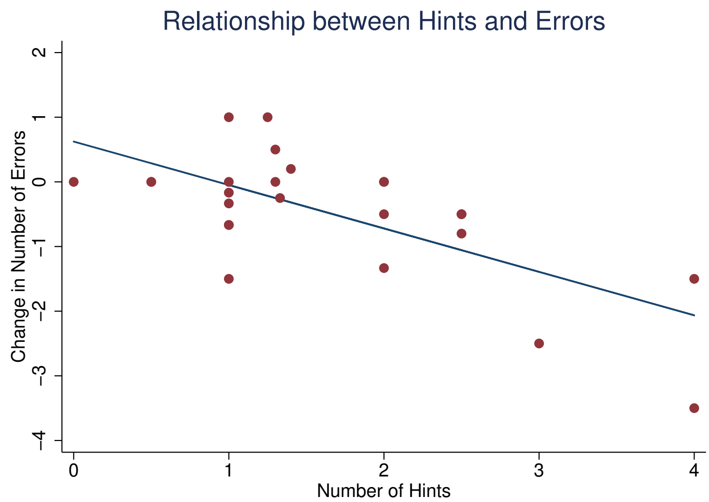

Figure 6.6: Relationship between hints and errors

We hypothesized that the hints given based on the hand position would lead to fewer errors. When we examined the relationship between hints and errors, we computed the average number of hints for each child, i.e., the total number of hints given to a child divided by the number of repetitions. The change of errors was computed as the difference between the average number of errors between the two repetitions of the games. A regression analysis between the errors and the hints given showed a significant correlation ($F(1, 23) = 19.89, p < 0.001$). The relationship between the number of hints and number of errors is shown in Figure 6.6.

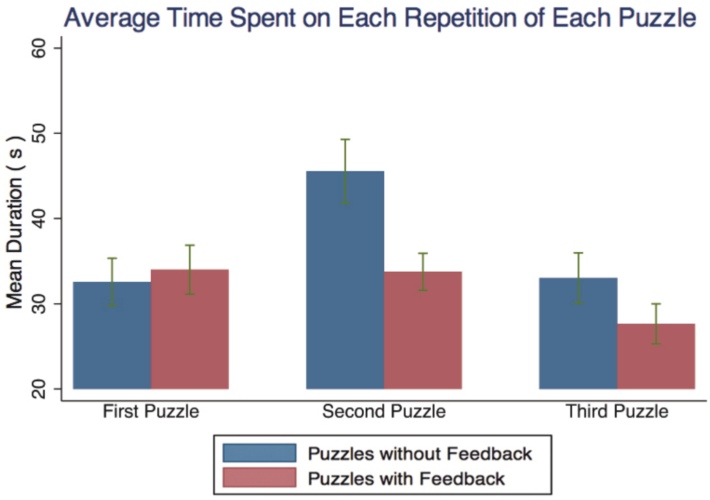

Figure 6.7: Average time spent on each repetition of each puzzle

Even though children were willing to spend more time on the experimental game indicating that means the game with feedback was more attractive, we need to show that this game can foster children's learning instead of just make children spend more time on it. Figure 6.7 shows the average time spent on each repetition of each puzzle in both games. A paired t-test for the first puzzle ($t(24) = -0.49, p = 0.62$) shows that there is no significant difference between the two games. The results of the second puzzle ($t(24) = 3.95, p < 0.001$) and the third puzzle ($t(24) = 2.20, p < 0.05$) indicate that the time spent on the second puzzle and third puzzle of the game with feedback is significantly less than the time spent on the game without feedback. The game with feedback provides more guidance and hints according to children's hands' position in the second and third puzzles, and the analysis proves that the feedback given by the game can significantly reduce the time children need to spend to pass the puzzles.

6.3.2 Qualitative Analysis

We analyzed two groups of qualitative data. First we recorded the conversation during the playing of both games and second, we made an interview with parent/guardian at the end of the experiment.

Variable	Comment
N_g	Number of parent/guardian's discourses
L_g	Number of sentences said by parent/guardian
T	Type of parent/guardian's discourse, enumerating as imperative, suggestive or commentary
N_c	Number of children's discourses
L_c	Number of sentences said by children
N	Number of complete conversations, which means it includes a guardian's question,a child's response and a guardian's comment.

Table 6.2: Statistical data counted from recording

Conversation Analysis

The children were allowed to freely play the games under the supervision of a parent or guardian. The conversation was audio-taped for analysis. Following Gordon et al. [75], we viewed the discourse between children and adults as a process of exchange, and we analyzed the conversation based on the measurements in Table 6.2. L_g is the number of sentences said by parents/guardians during the conversation. N_g means how many discourses the parent/guardian has during the conversation, each of the discourse includes several sentences which were said without interruption, and has a type T assigned to it. Parent/guardian's discourse may have three types.

1. Imperative: The parent/guardian asks the child to complete a task, e.g. "put the apple on the second basket" or "click on the stone". This discourse is defined as Initiating Request for Goods and Services.

2. Suggestive: The parent/guardian intends to give the child a hint or some guidance by asking child a question, like "Can you hear the sound on the tree?" "Can you remember what animal is behind this door?" This behaviour is called Request Information

3. Commentary: Feedback given by parent/guardian based on the child's response or behaviour, like "Yeah, you are right." and "This is not the correct answer, think again." This is defined as Acknowledgement of Information.

Children's participation in this conversation is measured by the number of their sentences. The type of most of them is "Giving Information", e.g. saying "three" in reply to the parent's question "What is the result of 5-2", or telling the parent "It's a tiger!". Sometimes children may also give Request-Information statement, e.g. "what's this animal?"

N represents the number of complete conversations between parent/guardian and the child, which is defined as an ideal-typical sequence initiated by a question, followed by a respond, sometimes finished with an acknowledgement. Tables 6.3 and 6.4 are two examples.

Role	Conversation	Type
Parent	How many animals are there?	Request Information
Child	Four.	Giving Information
Parent	Good boy.	Acknowledgement Information

Table 6.3: First example of a complete conversation

Role	Conversation	Type
Child	I know this, it's a tiger!	Giving Information
Parent	No, dear. This is a leopard.	Acknowledgement Information

Table 6.4: Second example of a complete conversation

Dialogues like these are counted as complete conversations. Figure 6.8 shows that the number of intergenerational conversations is higher in the game with feedback than in the game without feedback, Wilcoxon signed rank test ($z = -2.0, p < 0.05$). An analysis of the details of the recording shows that the difference comes from the fact that there are more conversation in the feedback game following the pattern of the second example given above, indicating that children are more engaged and willing to initiate a conversation. The change of the game does not have much effect on the parents' behaviour, who are equally likely to give guidance and ask question to the child, but it leads to the child expressing more. For example, one of the most common scenarios is the following. The child moves the object above the screen and try to find the correct result of the given equation, and says "Em, not this... is it this one..." while being aware of the real-time hint given from the game. Then the parent replies "yes, that is correct." However in the game without feedback, the child does not have the opportunity to get hints before making a decision, so it can only repeat the process of trying to choose an option, then the game reports error, then trying another option. So the game with feedback can help children to try to find the guidance by themselves and initiate the conversation to express their findings.

Figure 6.9 shows the comparison of average length of the guidance provided by parent/guardian between two games, in which $Avg_{length} = L_g(suggestive)/N_g(suggestive)$. The Wilcoxon signed rank test ($z = 2.0, p < 0.05$) indicates that the game with feedback can significantly reduce the words needed for guidance from the parent/guardian. The reason is that the feedback given by the game based on the child's hand position can replace part of the parent's work.

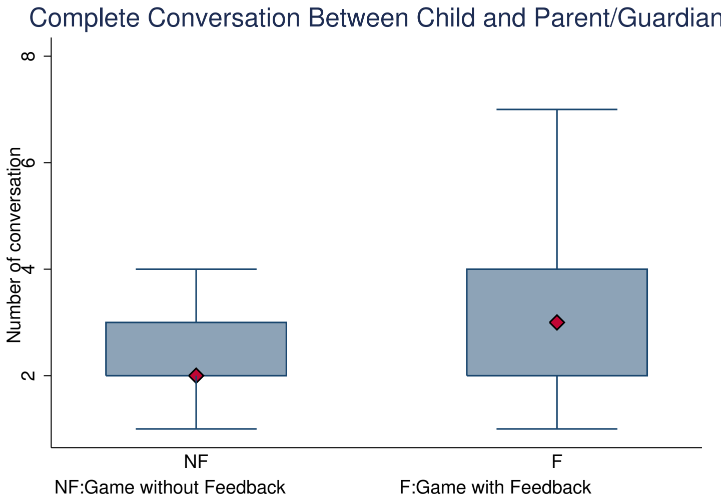

Figure 6.8: Number of Complete Conversations

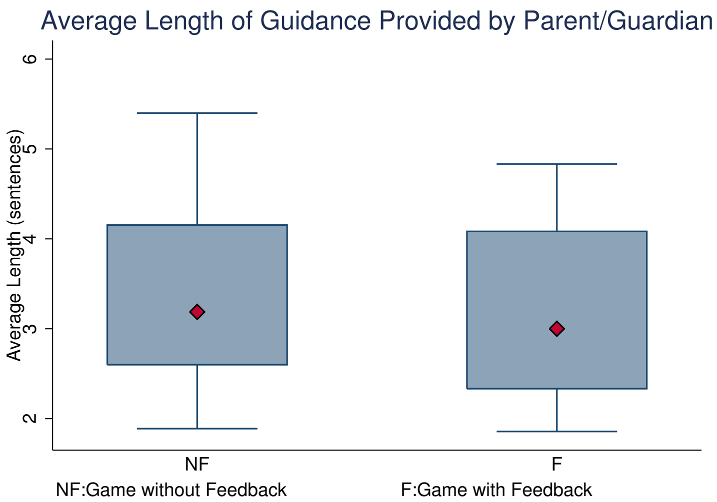

Figure 6.9: Average Length of Guidance given by Parent/Guardian

Parent Interviews

At the end of each experiment, the parent/guardian was invited to participate in an interview. The interview lasted for 5-10 minutes and participants needed to answer the questions listed in the questionnaire (See Table 6.5). They had the right to skip any question that made them feel uncomfortable and were free to leave any comment. The answers were rated on five levels:1) Extremely poor. This design may lead to misunderstanding and lower children's interest. 2) Poor. It is useless design. 3) Fair. It just meets the needs. It is not very helpful. 4) Good. It is well designed and useful. 5) Excellent. It is very helpful and perfectly designed [7].

The questionnaire was divided into three sections, feedback about the game in general, evaluation of the game interface, and feedback about social interaction. The design concepts were visualized using video prototypes to show the physical interaction involved and the movements children were expected to carry out. These prototypes were reviewed by teachers and their feedback was generally positive regarding the playful activities anticipated for the children. They thought that audio and visual feedback would be enjoyed by the children, thus motivating them for training, and that the system could, in principle, offer feedback as to the correct execution of the exercises. All the questions are listed in Table 6.5.

System	1	Do you think the audio and visual feedback have a positive effect on children's learning?
	2	Do you think the feedback based on hand position has a positive effect on children's learning?
	3	Do you think the tangible object has a positive effect on children's learning?
	4	Compared with mobile devices and tablets, do you think this tabletop has more advantages or disadvantages?
Content	5	Do you think the types of puzzles in this game are suitable for your child?
	6	Do you think the difficulty of the puzzles is suitable for your child?
	7	Are you aware that the difficulty of the puzzle increased based on your child's performance?
	8	Do you think this game can motivate communication between generations?

Table 6.5: Questionnaire used in parent interviews

There were 12 parents or guardians that participated this interview at the end of their experiment.

Here we list some of the most common answers of each category.

From the interview result shown in Table 6.7 and Table 6.6, we can infer how users evaluated the system. Most users gave positive feedback about the system design. The visual/audio effect,

Index	Extremely poor	Poor	Fair	Good	Excellent
1	0	0	0	8	4
2	0	0	2	5	5
3	0	0	2	4	6
4	0	0	0	4	8
5	0	0	0	3	9
6	0	0	3	6	3
7	0	0	4	7	1
8	0	0	1	5	6

Table 6.6: Statistical result of the questionnaire in parent interviews, including the number of each rating received for all questions.

integration of tangible objects and the feedback mechanism were considered very helpful to children's engagement, especially the tangible object was seen as a magic toy by the children. But the result also revealed that we need to emphasize the hand's position feedback to make it more noticeable to the children. It was difficult for younger children to pay attention to multiple objects. If the characters or game elements on the screen are very large and coloured, they may not be aware of the hand-shape sign on the screen moving along their hands. As to the game design, almost all the parents were fine with the puzzle types and difficulty levels. They and their children were familiar with these puzzles and did not feel it was hard to adapt to them, which also means for the children who have a lot of game experience, these puzzles are not very innovative. And although three levels of each puzzle can be enough to be adaptive to most of children in three to six years old range, some harder puzzles are needed for the children who have much better skills than their peers. Compared with other handheld mobile devices and tablets, most of the parents think the surface tabletop supporting multi-player has many advantages. They think it very useful to play with and monitor their children when they play the game. Traditional intergenerational games are not fun enough, and the tablet cannot let the parent to join in the game experience, so this large tabletop make a good balance and supports to the communication between generations.

The result of this experiment may be influenced by some potential factors, including the reliability of the system and the background of the participants. The tangible interface proposed in this thesis do not come without their flaws which are due to the limitation of the hardwares and environment, such as hand tracking errors. But these problems did not disturb or interrupt the game playing.

Audio and visual effect	Good	Especially for the memory game, she like the animals a lot.
	Good	For sure it's helpful.
Hand position feedback	Good	It's a good idea but may not be very noticeable for children younger than her.
	Excellent	When she figure out how to use that, she learn more things by herself.
	Fair	I don't think she noticed that.
Tangible objects	Excellent	She even didn't want the apple go.
	Good	She never met this kind of things before and really like it.
	Fair	It has almost the same effect as finger.
Type of puzzles	Excellent	They seem perfectly to the child, and he loved them.
	Good	She likes them, but she play gams a lot and always want to try new ones.
Difficulty of puzzles	Good	At the beginning it's a little to easy, but later it's ok.
	Excellent	I can see how she quickly figure out the problem and she can learn very fast.
	Fair	This game is a little too easy for her. I teach her math everyday.
Comparison between devices	Excellent	It's easier to see what they are doing, that's very useful for our parents.
	Good	Bigger perspective make it more interesting.
	Excellent	The girls enjoy it. They can discuss and play together. If you don't stop them, they will play the game forever.
Communication between generations	Good	The idea of educational games on big table is very good.
	Excellent	I always let my child play games on iPad, but this game is more helpful, because I can play with her.

Table 6.7: Common answers in parent interviews about the system design and game content of the children's educational game

According to the interview, parents showed great satisfaction with the system performance. Hence the system reliability does not have an obvious effect on the experimental result. The participants are children with the age from 3 to 6, so the variance of their previous gaming experience and skill development may lead to different performance during the experiment. However, in our analysis, we compared each child's performance between the two versions of games. So the effect made by children's various background on the experimental result is very limited.

Chapter 7

Conclusions

This chapter first summarizes the contributions of this thesis. Then we discuss the limitations of the proposed TUI system and the design issues of the educational games based on the user evaluation experiment. Lastly, we discuss future work.

7.1 Contributions

This thesis explores the potential of using tangible user interface (TUI) as tool to develop educational games for pre-school children. It presents a system using a tabletop surface as base for interaction, embedding the tracking mechanism in the surface, and a depth camera to obtain user's kinetic information for providing more feedback and natural interaction. This interface combines the advantages and flexibility of traditional screen displays and the capacity of allowing people to share the digital content on the tabletop. Following the guidelines and strategies of game design for young children, we developed an educational game for children in the age range of three to six years.

One of the major contributions of this work to the area of Human Computer Interaction is the exploration of the Continuous Interaction Space (CIS), where the tabletop surface and the volume above it are seen as a unified whole [52]. Tabletop computers featuring multi-touch input and object tracking are an increasingly common platform for research. However such systems are normally confined to sensing activity on the tabletop surface. This thesis argues that the space above the tabletop is a rich and unexplored canvas for supporting user interaction. We presents the technical details of how to combine the depth sensing into a surface tabletop interface, including the segmentation of depth map and algorithms for hand tracking. The information about users' hand in the space above the surface tabletop endows the developers more flexibility in the design of the applications. In this project, we propose an educational game that can provide rich feedback to children according to their hand interactions.

Another exceptional feature of this work is the summary of game strategies that are developmentally appropriate for young children. Though there are a wide range of digital games available on various devices on the market, many are not well designed to provide powerful interactive experiences that can foster young children's learning, skill building and overall development. More research is needed to explore the effects of rules, challenges and feedback of these games on children's development. This thesis designs and presents implementations of an educational game for pre-school children following these game design strategies. They are helpful to engage the children and foster their learning process, and they are developmentally appropriate by taking the capabilities and interests of the target population of children into account. The ideas and solutions involving the game design for young children discussed in this thesis should benefit all researchers who use tangible interface for children's games.

This thesis also pays attention to the effect made by the digital educational games based on this tangible user interface on children's behaviour and intergenerational communication. The majority of games for children are designed for one user working directly with the application and lack any face-to-face interaction. For pre-school children, the practice of appropriate social interaction techniques with peers is very important to their social skill development. Many parents limit the time their children spend on them because using traditional devices, parents loose the monitoring of their children, and leaving the child alone and silent is a concern for parents. Tabletop technology is a unique platform for multi-player gaming that combines the benefits of digital games with the affordance of face-to-face interaction, so it has the potential to be an effective intergenerational communication tool. Children's performance and the feedback from parents and guardians in the user evaluation experiment also shows that the educational game development in our system is significantly more effective to children's skill building and intergenerational communication than traditional educational games.

7.2 Design issues and Limitations

The tangible user interface and the educational game proposed in this thesis do not come without flaws. These flaws are due to the limitation of the hardware and problems encountered during game design.

The system presented in this thesis is using a Microsoft Pixelsense tabletop as the base of the interaction and a Microsoft Kinect for depth sensing. Because the sensors of both are IR based, the integration poses a technical challenge that the interference may lead to undesirable effects in the depth data obtained from the camera, which is the primary reason of the inaccuracy of hands' position computation. Another limitation of this system is that the object recognition supported by Microsoft Pixelsense is tag-based, and thus it can only identify the objects with a tag attached at the bottom and lying on the tabletop surface. The tangible objects are used individually and there is no physical interactions between objects, also the size and shape of the objects are limited.

From the interviews with parents and guardians, we are aware of that there are some shortcomings in the game rules and challenges. Firstly, even though there were multiple levels prepared for each puzzle in the game and the difficulty levels are adaptive to children's performance, the level settings could not satisfy all the children in age three to six due to the diversity of children's skill development, education background and their fast-growing abilities. Secondly, based on the observation of children's performances, we found that the feedback provided based on hand position was not noticeable enough for very young children. It was hard for them to be aware of the signals and warnings and connect what they see with their behaviour because children can pay attention to only very few objects simultaneously.

7.3 Future Work

There are several aspects of the educational game for pre-school children that should be worked on in the future, and further modifications can be done to reduce the usability limitations of the tangible tabletop system.

One of the weaknesses of the hand tracking mechanism lies in the data loss of the depth map due to the interference between IR based surface tabletop and depth sensors. An addition of another camera mounted at the side of the tabletop may compensate the loss of depth data of the top-mounted camera and provide more information to deal with overlapping arms in the segmentation process of the depth map. Another benefit is that the camera placed on the side of the surface can provide accurate estimation of the vertical distance between the hand and the tabletop. This information can be useful to offer better user experience, for example, the feedback may changes according to the height of the hands, or users can control the parameters by moving hands up and down. But this configuration also rises some problems to be solved, such as the calibration of the two cameras.

Also, as discussed in Chapter 2, there are many micro-controllers that can be embedded in tangible objects. They can capture a wide range of physical properties, such as motion and acceleration, and provide haptic feedback such as attraction and repulsion. This integration will enlarge the range of applications that can be developed in this system. For example, children can play block world on the tabletop with each of the blocks digitally connected to the others as well as being recognized by the tabletop. The haptic feedback provided by the actuator can make the games more attractive to the children and the sensors can obtain information about children's behaviour as the input of the game, for progress recording or for rehabilitation use.

On the game design perspective, future work should include adding more types of puzzles and preparing more difficulty levels for each puzzle to meet the needs of children of different ages and skill levels. This may be a reasonable solution in the short term, but as players begin to get familiar with the puzzles, these pre-set games may not match their interest and fast-growing abilities. One solution is to give the parents and guardian's access to customizing the puzzles. According to children's preference and progress, or following the curriculum of the early learning centres, they should be able to choose the packages provided by the system and set up the difficulty levels by themselves. Some modifications can be done on the game interface to make children aware of the feedbacks provided by the game more easily, such as adding sound guidance or using animation to guide children's attention.

Bibliography

[1] Redocn graphic designer forum. http://bbs.redocn.com/, 8 2014.

[2] Zcool online resources for graphic designer. http://www.zcool.com.cn/, 8 2014.

[3] ANAND, S., AND KROSNICK, J. A. Demographic predictors of media use among infants, toddlers, and preschoolers. *American Behavioral Scientist 48*, 5 (2005), 539–561.

[4] ANNETT, M., ANDERSON, F., AND BISCHOF, W. F. Hands, tables, and groups make rehabilitation awesome! In *Proc. of CyberTherapy & CyberPsychology Conference* (2010), pp. 3–8.

[5] ANTLE, A. N. The cti framework: informing the design of tangible systems for children. In *Proceedings of the 1st international conference on Tangible and embedded interaction* (2007), ACM, pp. 195–202.

[6] BARR, R., CALVERT, S., AND WILSON, B. Attention and learning from media during infancy and early childhood. *Handbook on children, media, and development* (2008), 143–165.

[7] BLAIR, J., CZAJA, R. F., AND BLAIR, E. A. *Designing surveys: A guide to decisions and procedures*. SAGE Publications, Incorporated, 2013.

[8] BRAGDON, A., DELINE, R., HINCKLEY, K., AND MORRIS, M. R. Code space: touch+ air gesture hybrid interactions for supporting developer meetings. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (2011), ACM, pp. 212–221.

[9] BREDEKAMP, S., AND COPPLE, C. Developmentally appropriate practice in early. *Childhood Education* (1997).

[10] BRUCKMAN, A., BANDLOW, A., AND FORTE, A. *Human Computer Interaction for kids*. Lawrence Erlbaum, Hillsdale, NJ, USA, 2002, ch. HCI for kids, pp. 428–440.

[11] CAMARATA, K., DO, E. Y.-L., JOHNSON, B. R., AND GROSS, M. D. Navigational blocks: navigating information space with tangible media. In *Proceedings of the 7th international conference on Intelligent user interfaces* (2002), ACM, pp. 31–38.

[12] CAPPELLETTI, A., GELMINI, G., PIANESI, F., ROSSI, F., AND ZANCANARO, M. Enforcing cooperative storytelling: First studies. In *Advanced Learning Technologies, 2004. Proceedings. IEEE International Conference on* (2004), IEEE, pp. 281–285.

[13] CHAUBAL, S. S. Memory games for kids. http://www.studyvillage.com/resources/2771-Memory-Games-for-kids.aspx, 5 2011.

[14] COATES, G. D., AND STENMARK, J. K. *Family Math for Young Children: Comparing.* ERIC, 1997.

[15] DIETZ, P., AND LEIGH, D. Diamondtouch: a multi-user touch technology. In *Proceedings of the 14th annual ACM symposium on User interface software and technology* (2001), ACM, pp. 219–226.

[16] DIPPON, A., WIEDERMANN, N., AND KLINKER, G. Seamless integration of mobile devices into interactive surface environments. In *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces* (2012), ACM, pp. 331–334.

[17] DOHSE, K., DOHSE, T., STILL, J. D., AND PARKHURST, D. J. Enhancing multi-user interaction with multi-touch tabletop displays using hand tracking. In *Advances in Computer-Human Interaction, 2008 First International Conference on* (2008), IEEE, pp. 297–302.

[18] DRUIN, A., AND FAST, C. The child as learner, critic, inventor, and technology design partner: An analysis of three years of swedish student journals. *International Journal of Technology and Design Education 12*, 3 (2002), 189–213.

[19] DUNNE, A., DO-LENH, S., LAIGHIN, G., SHEN, C., AND BONATO, P. Upper extremity rehabilitation of children with cerebral palsy using accelerometer feedback on a multitouch display. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE* (2010), IEEE, pp. 1751–1754.

[20] ECHTLER, F., AND KLINKER, G. Tracking mobile phones on interactive tabletops. In *GI Jahrestagung (1)* (2008), pp. 285–290.

[21] FERNAEUS, Y., AND THOLANDER, J. Finding design qualities in a tangible programming space. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (2006), ACM, pp. 447–456.

[22] FJELD, M., FREDRIKSSON, J., EJDESTIG, M., DUCA, F., BÖTSCHI, K., VOEGTLI, B., AND JUCHLI, P. Tangible user interface for chemistry education: comparative evaluation and re-design. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2007), ACM, pp. 805–808.

[23] GARRISON, M. M., AND CHRISTAKIS, D. A. *A teacher in the living room?: Educational media for babies, toddlers and preschoolers: A background report prepared for Kaiser Family Foundation*. Henry J. Kaiser Family Foundation, 2005.

[24] GIROUARD, A., SOLOVEY, E. T., HIRSHFIELD, L. M., ECOTT, S., SHAER, O., AND JACOB, R. J. Smart blocks: a tangible mathematical manipulative. In *Proceedings of the 1st international conference on Tangible and embedded interaction* (2007), ACM, pp. 183–186.

[25] GOH, W.-B., SHOU, W., TAN, J., AND LUM, G. Interaction design patterns for multi-touch tabletop collaborative games. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (2012), ACM, pp. 141–150.

[26] HALGREN, S. L., FERNANDES, T., AND THOMAS, D. Amazing animation: movie making for kids design briefing. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (1995), ACM Press/Addison-Wesley Publishing Co., pp. 519–525.

[27] HAUBNER, N., LUDERSCHMIDT, J., LEHMANN, S., SCHWANECKE, U., AND DÖRNER, R. Gestural input on and above an interactive surface: Integrating a depth camera in a tabletop setup. ARCS.

[28] HEALY, J. M. *Failure to connect: how computers affect our children's minds–for better and worse*. Simon and Schuster, 1999.

[29] HELD, R., GUPTA, A., CURLESS, B., AND AGRAWALA, M. 3d puppetry: A kinect-based interface for 3d animation. In *Proceedings of the 25th annual ACM symposium on User interface software and technology* (2012), ACM, pp. 423–434.

[30] HILLIGES, O., IZADI, S., WILSON, A. D., HODGES, S., GARCIA-MENDOZA, A., AND BUTZ, A. Interactions in the air: adding further depth to interactive tabletops. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology* (2009), ACM, pp. 139–148.

[31] HORN, M., ATRASH LEONG, Z., BLOCK, F., DIAMOND, J., EVANS, E. M., PHILLIPS, B., AND SHEN, C. Of bats and apes: an interactive tabletop game for natural history museums. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (2012), ACM, pp. 2059–2068.

[32] HORNECKER, E. Beyond affordance: tangibles' hybrid nature. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction* (2012), ACM, pp. 175–182.

[33] ISHII, H., WISNESKI, C., ORBANES, J., CHUN, B., AND PARADISO, J. Pingpongplus: design of an athletic-tangible interface for computer-supported cooperative play. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (1999), ACM, pp. 394–401.

[34] IZADI, S., HODGES, S., TAYLOR, S., ROSENFELD, D., VILLAR, N., BUTLER, A., AND WESTHUES, J. Going beyond the display: a surface technology with an electronically switchable diffuser. In *Proceedings of the 21st annual ACM symposium on User interface software and technology* (2008), ACM, pp. 269–278.

[35] JACOB, R. J., ISHII, H., PANGARO, G., AND PATTEN, J. A tangible interface for organizing information using a grid. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2002), ACM, pp. 339–346.

[36] JOHNSON, T., AND TOWLES, H. A. state, f. wu, g. welch, a. lastra, and h. fuchs. a projector-based physical sand table for tactical planning and review. Tech. rep., Tech. Rep.,(TR09-017), 2009.

[37] JORDAN, A. B., SCHMITT, K. L., AND WOODARD IV, E. H. Developmental implications of commercial broadcasters' educational offerings. *Journal of applied developmental psychology 22*, 1 (2001), 87–101.

[38] JUNIOR, D. Zooters jungle dots. http://www.disneyjunior.ca/en/jungle-junction/games/zooters-jungle-dots/, 6 2013.

[39] KALTENBRUNNER, M., AND BENCINA, R. reactivision: a computer-vision framework for table-based tangible interaction. In *Proceedings of the 1st international conference on Tangible and embedded interaction* (2007), ACM, pp. 69–74.

[40] KESKIN, C., KIRAÇ, F., KARA, Y. E., AND AKARUN, L. Real time hand pose estimation using depth sensors. In *Consumer Depth Cameras for Computer Vision*. Springer, 2013, pp. 119–137.

[41] KIJIMA, R., SHIRAKAWA, K., HIROSE, M., AND NIHEI, K. Virtual sand box: Development of an application of virtual environments for clinical medicine. *Presence 3*, 1 (1994), 45–59.

[42] KIRK, D., IZADI, S., HILLIGES, O., BANKS, R., TAYLOR, S., AND SELLEN, A. At home with surface computing? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), ACM, pp. 159–168.

[43] KIRKORIAN, H. L., WARTELLA, E. A., AND ANDERSON, D. R. Media and young children's learning. *The Future of Children 18*, 1 (2008), 39–61.

[44] KLEMMER, S. R., NEWMAN, M. W., FARRELL, R., BILEZIKJIAN, M., AND LANDAY, J. A. The designers' outpost: a tangible interface for collaborative web site. In *Proceedings of the 14th annual ACM symposium on User interface software and technology* (2001), ACM, pp. 1–10.

[45] KLOMPMAKER, F., NEBE, K., AND FAST, A. dsensingni: a framework for advanced tangible interaction using a depth camera. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction* (2012), ACM, pp. 217–224.

[46] LEITNER, J., HALLER, M., YUN, K., WOO, W., SUGIMOTO, M., AND INAMI, M. Incretable, a mixed reality tabletop game experience. In *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology* (2008), ACM, pp. 9–16.

[47] LEPPER, M. R., AND HENDERLONG, J. Turning" play" into" work" and" work" into" play": 25 years of research on intrinsic versus extrinsic motivation.

[48] LI, Y., FONTIJN, W., AND MARKOPOULOS, P. A tangible tabletop game supporting therapy of children with cerebral palsy. In *Fun and Games*. Springer, 2008, pp. 182–193.

[49] LIEBERMAN, D. A. What can we learn from playing interactive games. *Playing video games: Motives, responses, and consequences* (2006), 379–397.

[50] LIEBERMAN, D. A., FISK, M. C., AND BIELY, E. Digital games for young children ages three to six: From research to design. *Computers in the Schools 26*, 4 (2009), 299–313.

[51] MAGGIE.L. Memory games for kids : Improve your child's memory. http://maggie-l.hubpages.com/hub/ Memory-Games-for-Kids-Improve-your-Childs-Memory, 8 2011.

[52] MARQUARDT, N., JOTA, R., GREENBERG, S., AND JORGE, J. A. The continuous interaction space: interaction techniques unifying touch and gesture on and above a digital surface. In *Human-Computer Interaction–INTERACT 2011*. Springer, 2011, pp. 461–476.

[53] MARSHALL, P. Do tangible interfaces enhance learning? In *Proceedings of the 1st international conference on Tangible and embedded interaction* (2007), ACM, pp. 163–170.

[54] MARTINUSSEN, E. S., AND ARNALL, T. Designing with rfid. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction* (2009), ACM, pp. 343–350.

[55] MULLER, A. A., AND PERLMUTTER, M. Preschool children's problem-solving interactions at computers and jigsaw puzzles. *Journal of Applied Developmental Psychology 6*, 2 (1985), 173–186.

[56] OPENCV.ORG. Open source computer vision. http://opencv.org/, 7 2014.

[57] PATTEN, J., AND ISHII, H. Mechanical constraints as computational constraints in tabletop tangible interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2007), ACM, pp. 809–818.

[58] PATTEN, J., RECHT, B., AND ISHII, H. Audiopad: a tag-based interface for musical performance. In *Proceedings of the 2002 conference on New interfaces for musical expression* (2002), National University of Singapore, pp. 1–6.

[59] PIPER, A. M., O'BRIEN, E., MORRIS, M. R., AND WINOGRAD, T. Sides: a cooperative tabletop computer game for social skills development. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work* (2006), ACM, pp. 1–10.

[60] REN, Z., MEHRA, R., COPOSKY, J., AND LIN, M. Designing virtual instruments with touch-enabled interface. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (2012), ACM, pp. 433–436.

[61] RIDEOUT, V. J., VANDEWATER, E. A., AND WARTELLA, E. A. Zero to six: electronic media in the lives of infants, toddlers and preschoolers.

[62] RIZZO, F., AND GARZOTTO, F. The fire and the mountain: tangible and social interaction in a museum exhibition for children. In *Proceedings of the 6th international conference on Interaction design and children* (2007), ACM, pp. 105–108.

[63] ROFOUEI, M., WILSON, A., BRUSH, A., AND TANSLEY, S. Your phone or mine?: fusing body, touch and device sensing for multi-user device-display interaction. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (2012), ACM, pp. 1915–1918.

[64] SCHMIDT, D., CHEHIMI, F., RUKZIO, E., AND GELLERSEN, H. Phonetouch: a technique for direct phone interaction on surfaces. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology* (2010), ACM, pp. 13–16.

[65] SCHMIDT, D., SEIFERT, J., RUKZIO, E., AND GELLERSEN, H. A cross-device interaction style for mobiles and surfaces. In *Proceedings of the Designing Interactive Systems Conference* (2012), ACM, pp. 318–327.

[66] SCOTT-LITTLE, C., KAGAN, S. L., AND FRELOW, V. S. Creating the conditions for success with early learning standards: Results from a national study of state-level standards for children's learning prior to kindergarten. *Early Childhood Research & Practice 5*, 2 (2003), 1–27.

[67] SINGER, D. G., AND SINGER, J. L. *Imagination and play in the electronic age*. Harvard University Press, 2009.

[68] SOFTWARE, S. Preschool and kindergarden games. http://www.sheppardsoftware.com/preschool/preschool.htm.

[69] STANTON, D., BAYON, V., NEALE, H., GHALI, A., BENFORD, S., COBB, S., INGRAM, R., O'MALLEY, C., WILSON, J., AND PRIDMORE, T. Classroom collaboration in the design of tangible interfaces for storytelling. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2001), ACM, pp. 482–489.

[70] TAKEOKA, Y., MIYAKI, T., AND REKIMOTO, J. Z-touch: an infrastructure for 3d gesture interaction in the proximity of tabletop surfaces. In *ACM International Conference on Interactive Tabletops and Surfaces* (2010), ACM, pp. 91–94.

[71] TAYLOR, R., BOULANGER, P., AND OLIVIER, P. dream. medusa: A participatory performance. In *Smart Graphics* (2008), Springer, pp. 200–206.

[72] TERRENGHI, L., KRANZ, M., HOLLEIS, P., AND SCHMIDT, A. A cube to learn: a tangible user interface for the design of a learning appliance. *Personal and Ubiquitous Computing 10*, 2-3 (2006), 153–158.

[73] TRIONA, L. M., AND WILLIAMS, C. Point and click or build by hand: comparing the effects of physical vs. virtual materials on middle school students' ability to optimize an engineering design. In *In Proc. of CogSci2005* (2005), Citeseer.

[74] UNDERKOFFLER, J., AND ISHII, H. Urp: a luminous-tangible workbench for urban planning and design. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (1999), ACM, pp. 386–393.

[75] WELLS, G., MONTGOMERY, M., AND MACLURE, M. Adult-child discourse: Outline of a model of analysis. *Journal of pragmatics 3*, 3 (1979), 337–380.

[76] WILSON, A. D. Depth-sensing video cameras for 3d tangible tabletop interaction. In *Horizontal Interactive Human-Computer Systems, 2007. TABLETOP'07. Second Annual IEEE International Workshop on* (2007), IEEE, pp. 201–204.

[77] WILSON, A. D. Using a depth camera as a touch sensor. In *ACM international conference on interactive tabletops and surfaces* (2010), ACM, pp. 69–72.

[78] WILSON, A. D., AND BENKO, H. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology* (2010), ACM, pp. 273–282.

[79] WILSON, A. D., AND SARIN, R. Bluetable: connecting wireless mobile devices on interactive surfaces using vision-based handshaking. In *Proceedings of Graphics interface 2007* (2007), ACM, pp. 119–125.

[80] WILSON, P. H., DUCKWORTH, J., MUMFORD, N., ELDRIDGE, R., GUGLIELMETTI, M., THOMAS, P., SHUM, D., AND RUDOLPH, H. A virtual tabletop workspace for the assessment of upper limb function in traumatic brain injury (tbi). In *Virtual Rehabilitation, 2007* (2007), IEEE, pp. 14–19.

[81] WORKSHOP, S. Sesame street games for children. http://www.sesamestreet.org/games.

[82] ZUCKERMAN, O., ARIDA, S., AND RESNICK, M. Extending tangible interfaces for education: digital montessori-inspired manipulatives. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2005), ACM, pp. 859–868.