Directing Tangible Controllers with Computer Vision and Beholder

Peter Gyory
peter.gyory@colorado.edu
ATLAS Institute
University of Colorado Boulder
Boulder, USA

Krithik Ranjan
krithik.ranjan@colorado.edu
ATLAS Institute
University of Colorado Boulder
Boulder, USA

Zhen Zhou Yong
e0325771@u.nus.edu
Division of Industrial Design
National University of Singapore
Singapore

Clement Zheng
clement.zheng@nus.edu.sg
Division of Industrial Design &
Keio-NUS CUTE Center
National University of Singapore
Singapore

Ellen Yi-Luen Do
ellen.do@colorado.edu
ATLAS Institute
University of Colorado Boulder
Boulder, USA

ABSTRACT

We present Beholder, a computer vision (CV) toolkit for building tangible controllers for interactive computer systems. Beholder facilitates designers to build physical inputs that are instrumented with CV markers. By observing the properties of these markers, a CV system can detect physical interactions that occur. Beholder provides a software editor that enables designers to map CV marker behavior to keyboard events; thus connecting the CV-driven tangible controllers to any software that responds to keyboard input. We propose three design scenarios for Beholder—controllers to support everyday work, alternative controllers for games, and transforming physical therapy equipment into controllers to monitor patient progress.

CCS CONCEPTS

• **Human-centered computing → Systems and tools for interaction design**.

KEYWORDS

Computer Vision, Fiducial Markers, Tangible Interface

ACM Reference Format:
Peter Gyory, Krithik Ranjan, Zhen Zhou Yong, Clement Zheng, and Ellen Yi-Luen Do. 2022. Directing Tangible Controllers with Computer Vision and Beholder. In *SIGGRAPH Asia 2022 Emerging Technologies (SA '22 Emerging Technologies), December 6–9, 2022, Daegu, Republic of Korea.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3550471.3564764

1 INTRODUCTION

We present *Beholder*—a computer vision (CV) toolkit for building tangible controllers for interactive computer systems. This toolkit extends our previous work that demonstrates ArUco markers (CV markers) [Garrido-Jurado et al. 2014] as a material for tangible interaction design [Zheng et al. 2020]. Essentially, CV markers are simple printable graphics that translate physical movements to computational information (e.g. presence, position, rotation) when detected. They can be leveraged to report on a specific tangible interaction when placed strategically within a physical mechanism. In Figure 1 for example, pressing a button completes an ArUco marker for detection, while releasing a button conceals the marker—interacting with this button can thus be detected by the presence of the CV marker. *Beholder* enables designers to map CV marker behavior to keyboard input on the computer (e.g. `'w' keypress`), thereby enabling tangible controllers instrumented with CV marker to interact with any software that takes keyboard events as an input.

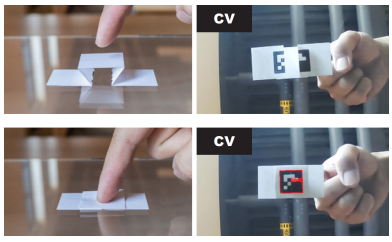

Figure 1: Detecting a button press with CV markers.

CV-driven tangible interfaces have inputs that are untethered, and typically require only a single camera to function (e.g. [Jordà et al. 2007; Savage et al. 2013]). They are simpler assemblies to construct in comparison to tangible interfaces instrumented with electronic sensors and circuits. However, current CV platforms for building tangible interfaces require designers to integrate CV and interaction logic by programming through software libraries (e.g. [Kaltenbrunner and Bencina 2007]). *Beholder* takes inspiration from physical computing toolkits like the *Micro:bit*[1] and *Makey Makey*[2]. These platforms offer "low floors" [Resnick and Rosenbaum 2013] that encourage interaction designers, particularly people unfamiliar with electronics, to try building functional tangible interfaces with physical computing. Notably, the visual editor that they provide and

[1]https://microbit.org/
[2]https://makeymakey.com/

SA '22 Emerging Technologies, December 6–9, 2022, Daegu, Republic of Korea

ACM ISBN 978-1-4503-9472-7/22/12.
https://doi.org/10.1145/3550471.3564764

Figure 2: Example applications for *Beholder*.

the direct mapping to human-interface device (HID) conventions facilitate interaction design for programming novices. *Beholder* aims to serve a similar role for CV—enabling designers to easily adopt CV markers as a material for building tangible user interfaces.

2 *BEHOLDER*: SYSTEM ARCHITECTURE

The *Beholder* toolkit operates on a system that consists of three main components (Figure 3A):

(1) *Actors* are the physical controllers that designers build. These controllers are instrumented with CV markers that report on different interaction events that occur (e.g. Figure 1).
(2) *Spectators* are camera devices (e.g. USB web cameras, smartphones) that observe the actors. These devices transmit the camera feed to a main computer for CV processing.
(3) The *director* is the main computer unit that receives and processes the camera feeds. The *Beholder* toolkit runs on this computer.

2.1 Editor

Beholder's editor runs as a standalone desktop application. It provides a visual node-based scripting environment (Figure 3B) for designers to create logical relationships between CV marker behavior and HID events (e.g. keyboard events like `keyup`, `keydown`, `keypressed`). The editor also provides a real-time debugging by labelling CV marker nodes with their properties (i.e. presence, position, rotation).

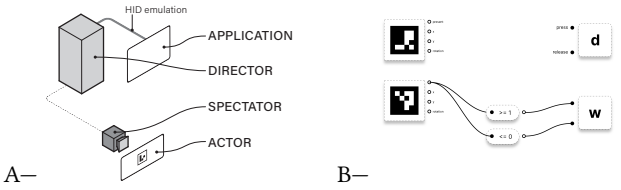

Figure 3: *Beholder*: System architecture and editor UI.

3 EXAMPLE APPLICATIONS

As we were developing *Beholder*, we also deployed it with interaction design students at the National University of Singapore to understand how designers will use it, as well as what they will build with it. From these facilitations, we identified three scenarios where designers might apply *Beholder*:

(1) *Everyday Productivity:* The COVID-19 pandemic has pushed people to work from home, and in turn spurred people to improve their home offices to facilitate this situation. *Beholder* facilitates designers to construct do-it-yourself productivity interfaces that support work (Figure 2A). For example, interfaces with hotkeys that trigger common actions (e.g. copy/paste), or more natural tangible interactions for digital workflows (e.g. a wheel to scrub through the timeline during video editing).
(2) *Alternative Game Controllers:* Video games are usually played on standard HID devices such such the keyboard, mouse, or game pad. *Beholder* enables designers to easily create alternative controllers for video games using simple materials like cardboard and paper [Gyory et al. 2022], which can facilitate a richer tangible playing experience (Figure 2B).
(3) *Augmenting Physical Therapy: Beholder* can be used to instrument existing physical therapy equipment into tangible controllers. Such controllers can be connected to a monitoring system to help therapists keep track of their client's progress; or, they can be directed to control video games, gamifying the (otherwise mundane) physical therapy experience.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No IIS2040489 and National University of Singapore Startup Fund A-0008470-01-00. We also want to thank our students for testing *Beholder* and providing us with invaluable feedback.

REFERENCES

S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, and M.J. Marín-Jiménez. 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* 47, 6 (2014), 2280–2292. https://doi.org/10.1016/j.patcog.2014.01.005

Peter Gyory, Perry Owens, Matthew Bethancourt, Amy Banic, Clement Zheng, and Ellen Yi-Luen Do. 2022. Build Your Own Arcade Machine with Tinycade. In *Creativity and Cognition*. ACM, Venice Italy, 312–322. https://doi.org/10.1145/3527927.3533023

Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. 2007. The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st international conference on Tangible and embedded interaction* (2007-02-15) *(TEI '07)*. Association for Computing Machinery, New York, NY, USA, 139–146. https://doi.org/10.1145/1226969.1226998

Martin Kaltenbrunner and Ross Bencina. 2007. reacTIVision: a computer-vision framework for table-based tangible interaction. In *Proceedings of the 1st international conference on Tangible and embedded interaction* (2007-02-15) *(TEI '07)*. Association for Computing Machinery, New York, NY, USA, 69–74. https://doi.org/10.1145/1226969.1226983

Mitchel Resnick and Eric Rosenbaum. 2013. Designing for Tinkerability. In *Design, Make, Play*. Routledge, Abingdon, Oxfordshire, UK.

Valkyrie Savage, Colin Chang, and Björn Hartmann. 2013. Sauron: embedded single-camera sensing of printed physical user interfaces. In *Proceedings of the 26th annual ACM symposium on User interface software and technology* (2013-10-08) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 447–456. https://doi.org/10.1145/2501988.2501992

Clement Zheng, Peter Gyory, and Ellen Yi-Luen Do. 2020. Tangible Interfaces with Printed Paper Markers. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (2020-07-03) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 909–923. https://doi.org/10.1145/3357236.3395578