OASIS: Creating Smart Objects with Dynamic Digital Behaviors

Ryder Ziola[1,2], Shweta Grampurohit[2,3], Nate Landes[2,3], James Fogarty[1], Beverly Harrison[2]

[1]Computer Science & Engineering
DUB Group
University of Washington
{ryder, jfogarty}@cs.washington.edu

[2]Intel Labs Seattle
beverly.harrison@intel.com

[3]Visual Design
University of Washington
shweta@uw.edu
nate.landes@me.com

INTRODUCTION

Our daily environment is full of physical objects that are unaware of and incapable of engaging with the digital world. These everyday physical objects lack the benefits of being digital, such as access to vast stores of information, complex interactive behaviors, and rich multimedia experiences. Although barcodes, RFID tags, and processors are being built into an increasing percentage of the objects that surround us, there will always be a large class of objects that it is either not possible or not economical to augment or retrofit with direct access to digital information.

Recent technological advances suggest the possibility of creating an infrastructure that allows normal physical objects to take advantage of the benefits of the digital. Specifically, processor power advances and improvements in machine learning promise robust and real-time object recognition. Increasingly affordable micro-projectors and depth cameras can be combined to create an input/output node that can observe physical objects and display information in-place. Together, these advances can be used to create islands of interactive infrastructure, where everyday physical objects can retain their original use or purpose while simultaneously being augmented with dynamic information, behaviors, and projected interactive content based on the context of their use.

This paper discusses OASIS, an infrastructure for Object Aware Situated Interactive Systems that we have designed to explore interaction with this notion of a smart object [2]. We developed two OASIS prototypes: Kitchen OASIS augments food objects on a kitchen counter with relevant information and functionality, and Lego OASIS augments Lego models with interactive digital behaviors. From these experiences, we developed a framework for understanding the general problem of representing objects and managing identity in an object-aware interactive system.

OASIS

OASIS combines a depth-camera and projector to recognize uninstrumented objects, receive user input, and display situated feedback on a table surface. Depth information is used to isolate hands and foreground objects in the scene. Each object is analyzed by object recognition algorithms and assigned a label based on features of their color gradient, depth gradient, and size [1]. The hand is identified by its angle of approach, and its depth information is used to provide touch events on the surface.

Kitchen OASIS

The Kitchen OASIS prototype was designed to explore object interaction in an environment with a large number of objects from a single category (i.e., food). When a food item is placed on the counter, the system projects an interaction point next to it and a label with its recognized identity. In the event of misrecognition, the identity can be touched to choose among a ranked list of alternate identities or to enter a new identity. Touching the interaction point brings up a menu of contextually relevant actions. For this prototype, we built a shopping list (adding/removing this item from a shopping list) and a recipe fetcher (displaying a list of recipes that incorporate this item). When objects are moved close to each other, they form a group. Objects in the group are visually connected and a single interaction point is provided for the entire group. This interaction point gives access to functionality scoped to the context of the set of grouped objects (e.g., recipes that include all ingredients in the group). Kitchen OASIS also demonstrates several standalone applications, such as a full shopping list. These are also sensitive to the context provided by physical objects (e.g., placing a food item on the shopping list filters the list to show only that item). Finally, an interaction point can be dragged off and released, creating a virtual copy of an item or group. This virtual copy saves that context from which it was captured and provides access to the same

Figure 1. Kitchen OASIS: A banana is recognized and the user adds it to the shopping list.

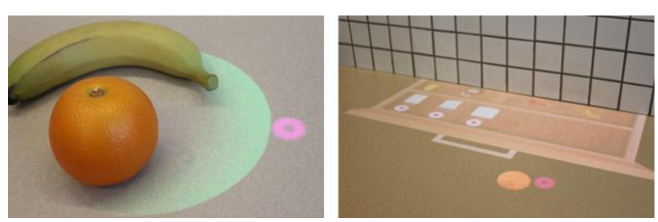

Figure 2. Kitchen OASIS: Objects are grouped by proximity and virtual copies can be stored in a digital drawer.

Figure 3. Lego OASIS: A dragon breathes flames, a house catches fire, and virtual train tracks are interactively placed.

functionality as the physical object or group. It can also be stored in a digital drawer for future use.

Lego OASIS

The Lego OASIS prototype enhances play with Lego toys by projecting environments around them and augmenting them with digital behaviors. In contrast to Kitchen OASIS, Lego objects do not show an explicit object labels or menu-like lists of functions upon being recognized. Instead, object identities trigger behaviors specifically tailored to that object. For example, placing a model of a train on the table causes the system to project a short length of train tracks in front of the train. The front end of the tracks can be dragged across the surface to draw a set of tracks for the train to use. A dragon triggers a fire breathing behavior and roars. The most interesting scenarios are triggered when objects combine based on proximity. For example, a fire-breathing dragon influences other nearby objects: when moved close to a house (or other 'flammable' Lego object), the dragon's flames can be directed at the house and set it on fire. Similarly, a fire truck moved close to the burning house sprays water to put out the fire. In addition to the physical Lego models, scenarios can involve virtual models. A projected virtual train station can stand in for a missing physical model that a child does not own or has not built. The virtual train station can still take part in the same interactive scenarios (e.g., when the physical train pulls up to the virtual train station, people are shown disembarking and refueling the train). In this way, behaviors relating multiple objects can be linked to physical objects, to virtual objects, or to any combination of both.

A Framework for Interactive Object Recognition

Building upon our experiences with these prototypes, we have distilled a conceptual framework that captures the subtleties of managing object identity in such object-aware systems [2]. The framework describes seven entities that combine to define an object and its identity over time.

A sensor frame contains one or more *detected objects*, the isolated sensor data received from an object. The detected objects in successive frames are mapped to the same *instance*, providing the object with continuity over time. An instance is assigned a *role* that captures the identity or category of an object (e.g., apple, steak, dragon). A role can be associated with multiple *capabilities* that encapsulate the things an object can do or that can be done with it. Information about the physical structure of an instance is contained in a *model* consisting of multiple *representations*, each capturing a distinct way of considering the structure of the instance (e.g., as a point cloud, as a CAD drawing). A model can be tagged with multiple *annotations* of semantically meaningful components (e.g., a train's "front", a dragon's "mouth", a fire truck's "lights"). The process of interactively managing object identity consists primarily of manipulating the relationships between these entities.

WORKSHOP PARTICIPATION

There are at least two aspects of our work with OASIS that we hope will be of interest to the workshop.

First, the design of our OASIS prototypes has focused on the natural meanings of everyday objects. Objects generally are not co-opted as arbitrary tokens, as generic actuators, or otherwise assigned a metaphorical meaning (e.g., as one might use a physical widget in a tangible interface). Instead, existing physical objects retain their original meaning while benefiting from a new layer of digital augmentation. This approach to augmenting existing objects without fundamentally changing their meaning allows us to extract context from the presence of objects and reason about activities while avoiding the need to provide explanation about an artificially constructed digital environment.

Second, our experience using a sensing infrastructure to add smart capabilities to existing objects has revealed some of the strengths and subtleties of this approach. Most notably, the decision to put computation in infrastructure rather than the objects themselves increases the scale of the world knowable by the system. We believe the workshop would be an appropriate forum for examining the contrasts between an infrastructure sensing approach to enabling smart objects versus directly building augmented objects. Further, the conceptual framework we have distilled in our work with OASIS provides a useful structure for understanding interaction with object-aware systems. It can be applicable to any sensor-based infrastructure for tracking and modeling real objects over time to build up the illusion of smartness. We look forward to discussing how it can provide concrete insights for designing and implementing systems that use probabilistic sensing and identification techniques to interact with uninstrumented objects.

REFERENCES

1. *Bo, L., Ren, X., Fox, D. (2010).* Kernel Descriptors for Visual Recognition. Proceedings of NIPS.
2. Ziola, R., Grampurohit, S., Landes, N., Fogarty, J., and Harrison, B. (2011). OASIS: Examining a Framework for Interacting with General-Purpose Object Recognition. *In Preparation.*