A Gestural Interaction Model to Classify Deformable Everyday Objects

Master's Thesis
submitted to the
Media Computing Group
Prof. Dr. Jan Borchers
Computer Science Department
RWTH Aachen University

by
Ravi Kanth Kosuru

Thesis advisor:
Prof. Dr. Jan Borchers

Second examiner:
Prof.Dr. Jochen Müsseler

Registration date: 04.08.2015
Submission date: 03.03.2016

Eidesstattliche Versicherung

______________________________ ______________________________

Name, Vorname Matrikelnummer

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Arbeit/Bachelorarbeit/ Masterarbeit* mit dem Titel

selbständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Für den Fall, dass die Arbeit zusätzlich auf einem Datenträger eingereicht wird, erkläre ich, dass die schriftliche und die elektronische Form vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Ort, Datum Unterschrift

*Nichtzutreffendes bitte streichen

Belehrung:

§ 156 StGB: Falsche Versicherung an Eides Statt

Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

§ 161 StGB: Fahrlässiger Falscheid; fahrlässige falsche Versicherung an Eides Statt

(1) Wenn eine der in den §§ 154 bis 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.

(2) Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Abs. 2 und 3 gelten entsprechend.

Die vorstehende Belehrung habe ich zur Kenntnis genommen:

Ort, Datum Unterschrift

Contents

List of Figures

List of Tables

Abstract

We re-purpose everyday objects in our daily life by substituting them for similar objects which may not be readily available. This ability to exploit everyday objects in different ways inspired us to look at one such interaction paradigm, deformations. We consider deformation as a change in physical shape brought about by the users hands. We look at everyday objects in a home environment to understand possible deformations that could be leveraged as an extra interaction dimension for digital interfaces.

We conducted a house-to-house survey to come up with a common list of deformable everyday objects. We categorized these multitude of objects primarily on shape based on the survey. We then conducted a gesture elicitation study on a representative subset of deformable everyday objects. This enabled us to have better a understanding of how users perceive and exploit deformable properties of everyday objects. We identify a generalized set of parameters on how our hands are employed to perform a variety of deformations. We observed that hand based deformation gestures itself could be expressive enough to convey enough information about the object under use. We created a more focused classification of objects based on common gestures. We followed up with a motion capture study to quantify deformable gestures on everyday objects. We used machine learning on the motion capture data to classify deformable everyday objects. This in turn also verified our observation that deformable everyday objects could be identified using deformation gestures.

The study of deformable everyday objects reveals new interaction possibilities via deformation gestures. We believe this could lead to opportunities of reusing these everyday deformation paradigms to interact with digital devices.

Acknowledgements

First and foremost, I would like to thank all the users who participated in my studies and allowed me to survey their homes as well. Your time and feedback have been a very valuable to this thesis.

Secondly, I would like to thank Christian Corsten, M.Sc., my supervisor. Thank you for keeping your door open and being available whenever I have needed it. Working with you has been a great learning experience which has enriched me, and I believe has made me a better researcher.

I would like to thank Prof. Dr. Jan Borchers and Prof. Dr. Jochen Müsseler. The working environment at the lab has been a big motivation. It has been an immense pleasure being a part of the Media Computing Group.

Special thanks to Chatchavan Wacharamanotham who gave valuable advice through my work.

I would also like to thank my friends Devashish, Subramanyam, Stephen and everyone else for all the patience and advice.

Finally I would like to thank my parents for their constant support in my life.

Thank you very much!
Ravi

Conventions

Throughout this thesis we use the following conventions.

Text conventions

Definitions of technical terms or short excursus are set off in coloured boxes.

EXCURSUS:
Excursus are detailed discussions of a particular point in a book, usually in an appendix, or digressions in a written text.

Definition:
Excursus

Examples are set off in blue boxes.

AN EXAMPLE:
Examples explain certain concepts or conventions used in the thesis.

Example:
An Example

The whole thesis is written in American English. For reasons of politeness, unidentified third persons are described in female form.

Chapter 1

Introduction

"We use our hands as general purpose devices, to pick up objects, to point, to climb, to play musical instruments, to draw and to sculpt, to communicate, to touch and feel, and to explore the world."

—MacKenzie and Iberall [1994]

Our hands are one of the primary instruments to interact with the everyday world. We could use them to grasp and pick up a coffee mug or leaf through the pages of a magazine. Our hands can also convey a lot of information about the objects they explore. For example, we can feel the texture of a soft velvet cloth or the sharpness of a knife's edge. The way we use our hands depends on the properties of the object that suggests how it can be handled. These may depend on the shape like a cylindrical bottle that can be grasped with the palm or material like textile that can be folded. Donald Norman referred to these properties as *affordances*.

Our hands are a primary instrument to interact with the world around us.

AFFORDANCES:
"[...] Affordances refers to the perceived and actual properties of the thing, primarily those fundamental properties that determine how the thing could be possibly used." – Norman [2013]

Definition: *Affordances*

Objects with similar affordances could be substituted for one another.

Similar affordances on different objects can help us substitute the function of one missing object with another or what is termed as appropriation[1]. For example, many a time when we are stumped while trying to locate a bottle opener, we resort to the one of the most creative ways of appropriating everyday objects. We have noticed how a cigarette lighter, to a cutlery knife and at times, a second bottle end up doing the job of being an ad-hoc bottle opener. The pointed end of a knife has a similar affordance to a screw driver while the blade shares its affordance with a letter opener. This knowledge of affordances could help us readily substitute a missing screw driver with a knife.

Instant User Interfaces are a concept of re-purposing everyday objects for digital control.

Corsten et al. [2013] applied the concept of re-purposing everyday objects to the digital domain and introduced Instant User Interfaces (IUI). They presented the use of everyday objects that mimic affordances of digital control devices as instant alternatives that decrease our dependence on dedicated devices. They allow for a spontaneous replacement of a missing or a broken controller along with an added enhancement of haptic feedback. This haptic feedback arising from tactile clues provided on everyday objects also supports eyes-free interaction. They note how people look for everyday objects that mimic GUI controls such as knobs and buttons.

The everyday world provides objects with a wide range of affordances.

The everyday world provides us with an extensive canvas filled with alternatives for interacting with the digital world. IUIs motivated us to take a deeper look into the everyday world on how we interact with the objects around us. Everyday objects come in varied shapes, sizes and materials contributing to a multitude of possible affordances. For example, coffee mugs afford touching and translations. Some objects like a sponge afford squeezing resulting in a change of its shape, also called as deformations. In contrast to the interactions with rigid bodies explored by Corsten et al. [2013], these flexible objects bring the possibility of new gesture paradigms into context.

[1]"The process of using a designed tool or object in a different context and for a different purpose than intended, is called appropriation."- Corsten [2012]

1.1 Deformation as an Interaction Paradigm

The Oxford dictionary defines deformation is:

DEFORMATION:
"Alteration of form or shape; relative displacement of the parts of a body or surface without breach of continuity; an altered form of."

Definition: *Deformation*

Deformation has been the focus of interaction research for a while. Deformable user interfaces in ongoing research comprise of various custom built devices that allow the recognition of deformations such as bending, twisting etc. (Schwesig et al. [2004], Kildal et al. [2012]). Deformations add an extra dimension of information that give an additional input capability that could help simplify a complicated workaround.

Custom deformable user interfaces have been a part of ongoing research.

As with everyday objects such as a sponge, which can be squeezed or a hand towel, which can be folded, we see a wide range of possible types of deformation. A single object can also offer multiple affordances that allow deformation. This versatility of everyday objects opens up a wide range of possible deformations. We believe this would be helpful in designing new user interfaces that take advantage of different deformation interactions.

Everyday objects could offer an insight into designing new deformation based interfaces.

1.1.1 Towards Deformable Everyday Objects

We were interested in the deformations that are possible using our hands. We extended the definition of deformation in relation to hand based interactions to formally define deformable objects as:

We define deformable everyday objects.

Definition:
Deformable Objects

Deformable Objects:
Objects, that can be manipulated with one or more fingers, resulting in a temporary/permanent physical shape different from the original. The resulting force needed to deform should not result in a physically damaged (e.g., cracked or broken) object.

Damaged objects are not covered under our definition.

In line with the definition of deformation we do not consider objects that are broken as a result of the deformation as it results in the breach of continuity. Deformations themselves could be visualized or characterized in multiple ways such as a squeeze, twist, pull etc.

The gesture performed by a user to affect deformations is termed *deformation gesture* in our work. We formally defined it as:

Definition:
Deformation Gesture

Deformation Gesture:
A deformation gesture is the combination of the use of hands (fingers) and the action performed to deform the object.

The subtle changes in how we employ our hands makes it possible to exploit the varied affordances of deformable everyday objects.

1.1.2 The Human Hand

The study of the human hand has an interesting advantage in HCI research.

So far we have been discussing the human hand as the primary tool in our interactions with everyday objects. The study of how we employ our hands has interested HCI researchers as much as psychologists. Vatavu and Zaicti [2013] showed how size and shape could be inferred from how users explored an object with their hands. They outlined the one of the advantages of understanding the use of our hands when grasping an objects, as an identification tool for the objects themselves.

This is also reflected in the contribution of our work that explores the feasibility of gesture based classification of deformable everyday objects.

1.2 Research Questions

We are interested in studying deformations in everyday objects in a home environment by trying to answer the following questions:

1. What are the deformable everyday objects in a home environment? We consider a combination of household survey and an informal interview to list and broadly categorize deformable everyday objects.
2. What are the different affordances offered by everyday objects that enable deformations? What are the common perceived affordances across users on a given deformable everyday object? These affordances could depend on various parameters such as form, material, and size. Multiple gestures could be performed on a given deformable every day object. We designed a gesture elicitation study which is influenced by guessability study (Wobbrock et al. [2009]) and gesture elicitation studies like Vatavu [2012]. Unlike the referred studies the user does not have defined end digital task that needs to be accomplished. We also use this to identify parameters which could be quantified to have an interaction model for everyday objects.

   We explore the world of deformable everyday objects by first identifying a representative set of objects followed by a gesture elicitation study.
3. What would be a representative model for interacting with deformable everyday objects? Could we create quantified representation of the gestures that could be used to classify the everyday objects around us. We use motion capture to have quantified data on the parameters identified by the previous studies. We then use a machine learning technique to verify the classification of objects based on the tracked parameters.

   We validate our gesture set for deformable everyday objects with analyzing motion capture data.

We look at deformations as a combination of hands, respective fingers, along with the action to exploit object affordances. We will provide a set of categories wherein the majority of everyday objects could be binned based on the gestures. We then use machine learning to identify objects using deformation gestures. We gathered varied deformation gesture data on everyday objects for our work. This data captures a lot of information beyond the scope of the analysis of our work that may prove useful for further research into understanding deformation gestures.

We do not deal with relevance of the deformation gestures for digital tasks.

We would like to point out that the relevance and mapping of tasks to deformable objects or gestures is not in scope for our study. Various studies, especially those including custom deformable devices provide an overview of the type of tasks that suit deformations such as bending (Ahmaniemi et al. [2014]). Corsten [2012] gives an extensive insight into the user experience of working with everyday objects for IUIs and provides a set of guidelines for the interaction flow. We assume the existence of systems that make everyday re-purposing possible and can capture deformation gestures on everyday objects.

Throughout our work we shall be using *deformable affordances* or *deformation affordances* to refer to affordances that allow deformations.

1.3 Outline

The thesis is structured as follows:

Chapter 2: We present an overview of the related work on everyday objects and deformable interfaces.

Chapter 3: This chapter includes the results of the house to house survey to list deformable everyday objects. We also broadly categorize the objects and create a limited representative set of objects to be used for the subsequent studies.

Chapter 4: We describe the gesture elicitation study on the representative set of objects. We discuss how users exploit the deformable affordances and their preferences.

Chapter 5: This chapter is a small excursus on the tools used for gathering quantified data and its analysis.

Chapter 6: We give detailed information on how the quantified data was collected and how it is structured. We also present the results from the data analysis which include the accuracy of the machine learning based classification.

Chapter 7: This concludes the thesis with a summary of the thesis and future work.

Chapter 2

Related Work

As the title of the thesis suggests, during our literature review we had two main branches of research to look into:

1. Everyday Objects
2. Deformable Interfaces

We look into existing literature that deals with everyday objects and deformable interfaces.

This is the broad categorization of the projects we reviewed. Any project that dealt with re-purposing everyday objects even if they included deformation of any form was included in the *Everyday Objects* group. Under *Deformable Interfaces* we looked at all projects that included custom built deformable devices. In the end, we also take a brief look at the work on automatic recognition of object size and shape by Vatavu and Zaicti [2013].

2.1 Everyday Objects

There has been extensive research into the working with everyday objects and re-purposing them as digital controllers. We analyze some of the existing work on everyday objects and discuss them based on the following criteria:

We analyze literature on everyday objects mainly based on the exploited affordances.

Location: We look at the area of everyday lives that the work targets. For example, is it confined only to the workspace desk or is area agnostic.

Affordances: We also examine the affordances that are exploited in the reviewed work. We shall be especially looking for deformation based gestures.

Technology: We take a brief foray into the technology implemented in the projects especially with regards to object detection.

2.1.1 OnObject: *Chung et al. [2010]*

OnObject allows custom sound coupling to any object tagged with RFID.

The OnObject system presented by Chung et al. [2010] allows the coupling of a custom sound tag to a gesture on any device or object. This is achieved through a device containing a RFID reader and an accelerometer for a sensor along with an integrated microphone and speaker. Any object could be tagged with an RFID tag and coupled using the system to a custom recorded voice or sound sample (cf. Figure 2.1). An object could be assigned multiple sound tags one for each supported gesture.

The device recognizes pre-defined gestures.

The system is coupled to the RFID tag and not the object itself.

The device has an in-built gesture recognizer which can detect an object grab, shake, swing, thrust, tilt, and circle motions(cf. Figure 2.2). When the object is grabbed or any of the other gestures are performed, the accompanying sound is played. The OnObject system is envisioned to be used with any kind of object and in theory is only coupled to the RFID tag. The system does not need to know anything about the object itself except that the object should be held in the hand with the OnObject device. This allows the system to be location agnostic. However, the authors see its use in the areas of language learning and in interactive story telling.

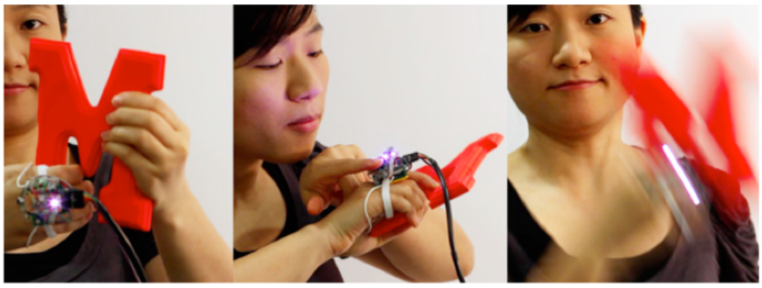

Figure 2.1: The OnObject Device. The user grabs an object attached with a tag and demonstrates a shake gesture. She then presses a button to record a custom sound to be tagged to the object. Shaking the object produces the recording.

Figure 2.2: The in-built gestures supported by OnObject.

2.1.2 iCon: *Cheng et al. [2010]*

The iCon project proposed the use of everyday objects as instant desktop controllers (Cheng et al. [2010]). The recognition of objects and associated gestures was accomplished using two prototypes. Both of the solutions were camera based with the objects augmented with fiducial markers. The first prototype had the camera looking down on the workspace with a bird's eye view of the area. This was used for translation gestures and rotations while the second prototype was for detecting *clicks* with the camera placed under the table looking up.

iCon used objects augmented with fiducial markers.

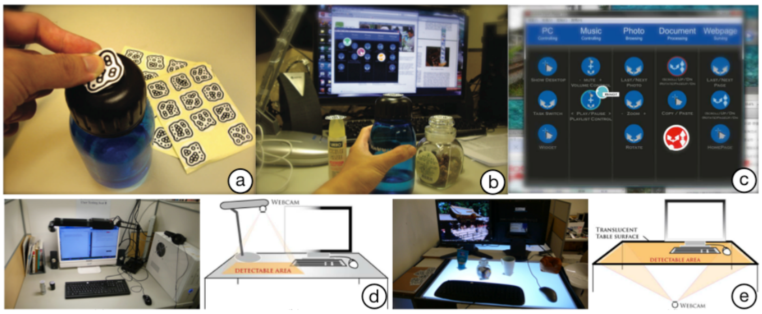

Figure 2.3: The complete iCon system. (a), (b) Everyday objects augmented with fiducial markers are utilized as controllers; (c) The user interface of iCon which lists the possible gestures; (d) Prototype with camera looking down on the desk; (e) Prototype with the camera placed under the table, looking up.

They were focused on cylindrical objects.

The implementation was focused on cylindrical shaped objects like bottles. This was because they could be easily grabbed and moved around. The implemented gestures included rotations, dragging and clicking (cf. Figure 2.3).

The system targeted computer tasks.

The system concentrated on the objects found around a work space at the office. Consequently the target tasks included computer tasks such as web browsing, desktop control and document processing.

2.1.3 Around-Device Devices: *Pohl and Rohs [2014]*

Around-Device Devices looked at objects around a mobile phone as alternate controllers.

Pohl and Rohs [2014] discuss the appropriation of everyday objects found around a phone in day-to-day life as a means of interacting with the phone itself. The paper focuses on identifying everyday objects that could be present around a phone under envisioned scenarios, mainly around the house. They also identify a set of tasks and perform a gesture elicitation study to identify interactions that users find useful instead using the application on the phone itself (cf. Figure 2.4).

The objects were generally categorized as shape primitives such as cubes or cylinders as possible. There were no ex-

Figure 2.4: A conceptual visualization of the around-device scenario where objects around the phone are detected as object primitives.

plicit instructions on how to use the objects or exploit affordances. Most of the actions included rotations or translations or touch gestures.

Objects were categorized as shape primitives.

Scenarios were envisioned generally as spaces around the house such as living room, dining area and workspace and a list of common objects around was gathered from a diary study. The tasks however were confined in general to interacting with a smartphone.

Spaces around the house were the envisioned scenarios.

2.1.4 Chair as an Ubiquitous Computing Device: *Probst et al. [2014]*

A chair is one of the most readily available objects available around our work place. Probst et al. [2014] explored the concept of using the office chair as an ubiquitous computing device. A motion sensing smart phone was attached to the chair and was used as the sensor that communicated with the computer.

The chair is used as an ubiquitous computing device

Figure 2.5: The interactions possible with the chair as an ubiquitous computing device: rotate, tilt, bounce up and down.

Chair movements controlled by the body as a whole left the hands free.

An office chair that could tilt, rotate and bounce up and down was used for the study. The chair movements were controlled by the body as a whole, so the users hands were free (cf. Figure 2.5). Thus, the users could use it as an additional interaction instead of substituting the keyboard or mouse.

An office was the study location.

The location for the study was the office workspace and the tasks were confined to controlling applications on the computer. They included web-browsing and music player control.

2.1.5 Instant User Interfaces: *Corsten et al. [2013]*

IUIs allow mapping any device within reach as a controller.

The Instant User Interface (IUI) paradigm introduced by Corsten et al. [2013] allows users to instantiate any object within reach as a controller by mapping a gesture to a given task. The authors proposed a vision based implementation without any augmenting markers or modifications to everyday objects. It relies on the presence of an initial model of the object being tracked. Depth thresholding is used to then detect touches on the objects.

Gestures include rotation of mug and clicking of the pen.

A pen and a mug were used for the evaluation study. They were sourced from a set of everyday objects that are found in work environments. The gestures included rotation of the mug and clicking of the pen button.

Objects found in work and home environments.

The evaluation tasks included controlling a presentation and lights. Though the object set was drawn only from work environments, the objects can be usually found in most home locations as well.

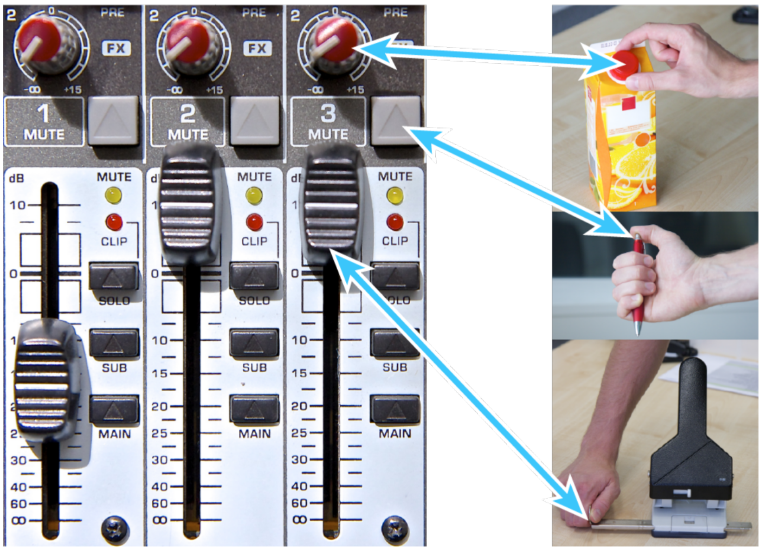

Figure 2.6: Everyday objects on the right such as the juice carton, pen, and hole puncher mimic controllers on a mixer shown on the left.

2.1.6 Overview: *The Contrast to Our Work*

In general the focus of research into everyday objects revolves around interaction with rigid bodies. This could be because most interaction paradigms rely on translations and rotations such as turning a knob to vary volume or directional navigation for music players. The interactions used in 2.1.4 "Chair as an Ubiquitous Computing Device: *Probst et al. [2014]*" and the use of a pen in 2.1.5 "Instant User Interfaces: *Corsten et al. [2013]*" include gestures that arise from relative displacements of the object parts. These actions are however built into the objects for that specific purpose. For example, the button that is used to activate the pen for writing is not a physical property of the material itself. Our study into deformable everyday objects does not cover such interactions. We look into deformation as an additional physical property of the object. IUIs point towards a future where every object around us is a potential means of interacting with technology. This is a vision that motivates our work to understand how people perceive deformable everyday objects.

In contrast to the work presented that deals with rigid bodies, we look into deformation as an additional physical property of the objects.

2.2 Deformable Interfaces

Deformation as an interaction paradigm has undergone extensive research in the past decade. We examine the related work based on:

We analyze literature in deformable objects mainly based on addressed deformation types.

1. The kind of deformations involved. These could be bending, twisting, squeezing or folding among others.
2. The device dimensions and a brief foray into the technical details of the artifact.
3. Evaluation of user studies and/or application area.

2.2.1 Gummi: *Schwesig et al. [2004]*

Gummi was a concept device based on the bending deformation.

Schwesig et al. [2004] introduced deformation as an interaction paradigm for hand held digital devices. They envisioned a separate interaction metaphor for deformation based interfaces to move away from WIMP(Windows, Icons, Mouse, and Pointer). The Gummi device was based around bending. Users could bend the device either up or down (cf. Figure 2.7).

The device could be held in the hand.

As stated the device could be held in the hand and included only a single size. The evaluation unit consisted of an LCD panel placed on a panel that included a bend and 2D position sensors.

The presented a set of GUI principles for such devices.

They presented a set of GUI principles for such devices such as mapping the bending to accomplish tasks including complicated ones like text input. Nonetheless, they do recognize that it is more suited for tasks such as map navigation and scrolling.

2.2.2 Nokia Kinetic Device: *Kildal et al. [2012]*

Kildal et al. [2012] described the concept of the Kinetic De-

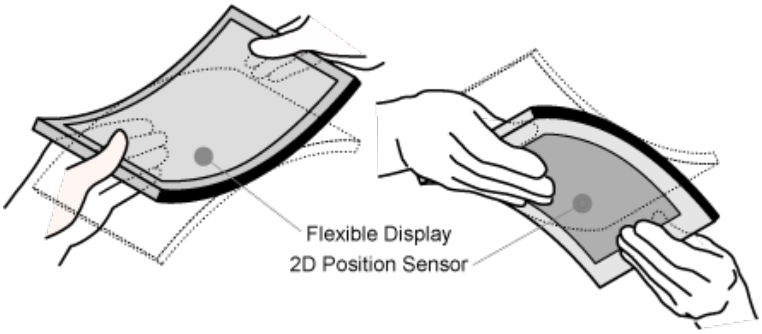

Figure 2.7: An illustration of the Gummi interaction device and the bending interaction. It would have flexible display and a 2D position sensor.

formable UI Research Prototype (DUI-RP) that was similar to the dimensions of a smart phone. The device supported bending in both directions as well as twisting (cf. Figure 2.8).

The Kinetic DUI device had the dimensions of a smartphone. It could be bent, and twisted.

The device had rigid ends connected via a deformable material. They could build devices with different stiffness of the deformable materials and also included variable deformation sensitivity.

They used the devices to study:

What is device bending good for?: They concluded that bending is more beneficial for continuous tasks such as zooming in maps, images etc., though discrete gestures are good for quick reactions. (Kildal and Wilson [2012]).

Device bending is suitable for continuous input.

Effect of stiffness on deformable UI: The study suggested that stiffness generally does not affect the ability to deform. Soft devices make it easier to bend but harder to maintain constant angle and vice versa. Soft devices feel nicer as far as user experience is concerned (Ahmaniemi et al. [2014]).

Stiffness does not affect deformation ability but soft devices feel nicer.

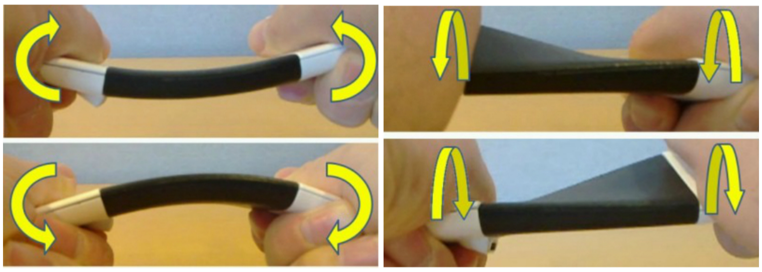

Figure 2.8: Possible interactions with a Kinetic prototype. It could be bent up, bent down, and twisted both ways.

2.2.3 Towards More Paper-like Input: *Gallant et al. [2008]*

FID allows us to interact with it like paper.

Gallant et al. [2008] demonstrated a foldable input device (FID) which allows us to interact with it similar to paper. Apart from sliding the finger on the FID like a touch, one could scoop, fold, bend or squeeze the device (cf. Figure 2.9).

The FID device uses a pattern of reflective markers that are tracked for gesture recognition.

The FID device has multiple reflective markers placed on them uniformly that are tracked with a camera. The change in pattern of the markers is used to recognize the various gestures. They could be used to control the flexible user interface (FUI) which are flexible displays simulated on a LCD screen. They present a set of interactions such as desktop navigation, document browsing etc. in the FUI.

2.2.4 Exploring the Effects of Size on Deformable User Interfaces: *Lee et al. [2012]*

The usability of large vs small DUIs was explored through mock-ups.

Lee et al. [2012] studied the effect of size on deformable interfaces. They used two plastic mock up devices to perform a gesture elicitation study to explore the usability of large vs small DUIs. The sizes used corresponded to an A4 paper and an iPhone 4 (cf. Figure 2.10). They presented the users with a set of eleven tasks such as zooming, open, close etc. and asked them to generate deformation gestures.

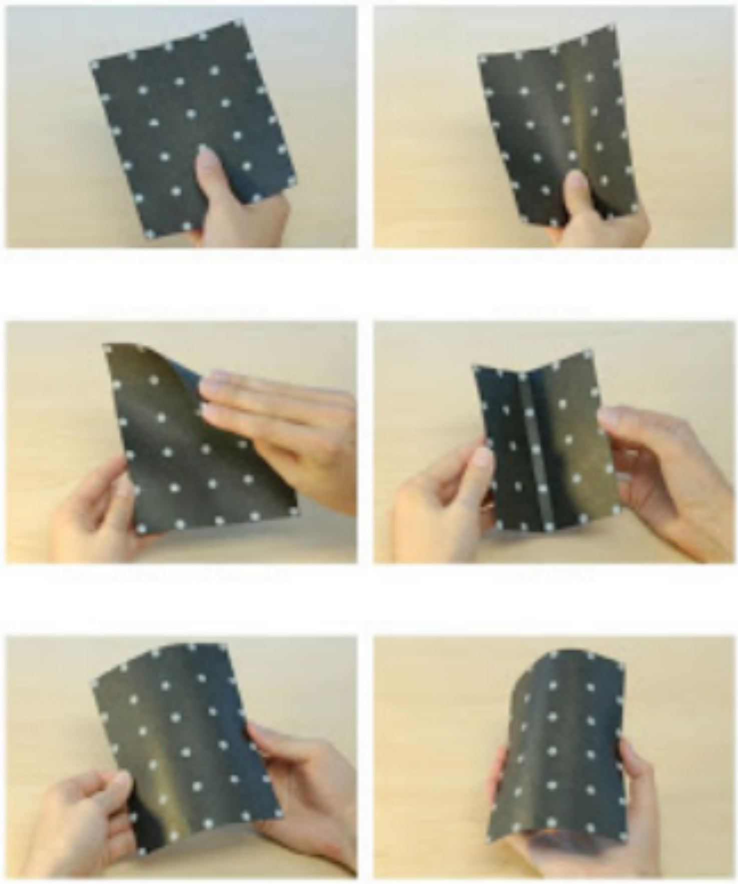

Figure 2.9: The supported deformation interactions of the FID device include thumb slide, scoop, top corner bend, fold, leafing, and squeeze.

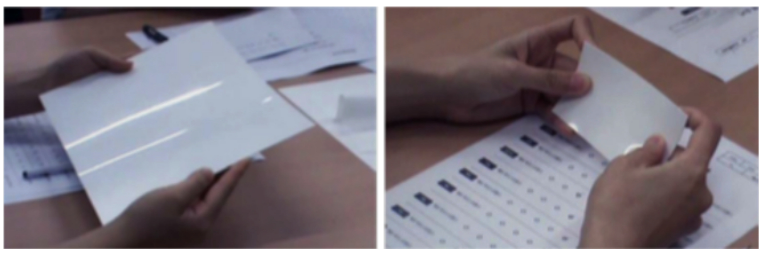

Figure 2.10: The prototypes used for studying effect of size on DUIs. They used paper mock-ups of A4 size (left), and iPhone 4 size (right).

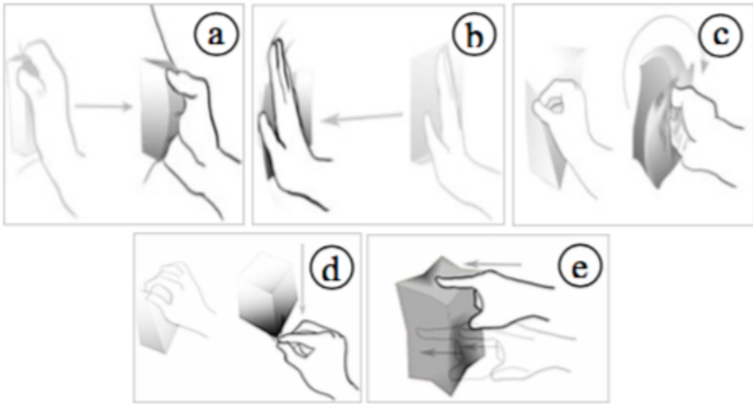

Figure 2.11: The deformable gestures elicited on a deformable display: (a) grab and pull, (b) push with flat hand, (c) grab and twist, (d) pinch and drag, (e) push with index finger.

The smaller device had a higher gesture agreement.

Users had a higher agreement on the smaller device and mainly used a single hand for the gestures. They indicated preference for combining touch for the larger display.

2.2.5 User-Defined Gestures for Elastic, Deformable Displays: *Troiano et al. [2014]*

A gesture elicitation study was conducted for deformable displays.

Troiano et al. [2014] conducted a gesture elicitation study for large elastic, deformable displays. Compared to the works presented before, the large size of the device under study enables new gestures such as pinching, grabbing and twisting, pulling and pushing (cf. Figure 2.11).

Visual content was projected on a fabric based display.

A piece of fabric measuring 76x47 cm was attached a wooden frame to act as the display. Visual content for tasks was projected on the fabric using a camera placed at the back (cf. Figure 2.12).

Push and drag were the major actions

Push followed by a drag were the two major type of actions. They presented a consensus of twenty seven gestures. Twenty of the gestures were uni-manual gestures and indicated users preferred uni-manual over bi-manual interaction.

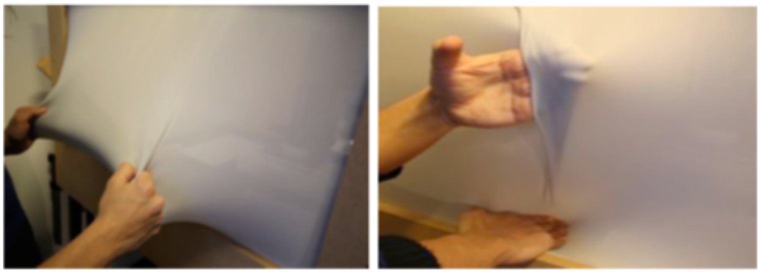

Figure 2.12: The deformable display prototype used for the gesture elicitation study. A user deforms a fabric attached on a wooden frame.

2.2.6 Overview: *The Contrast to Our Work*

The study of deformation gesture relevance is an important area that is being explored. It is a necessary pre-requisite to build user interfaces that can actually take an advantage of deformation to provide a richer user experience. The related work however, relies on building deformations into current devices such as mobile phones that usually come with regular shapes and sizes. On the other hand, the varied nature of everyday objects could mean users have a completely difference experience from one object to the other. We try to narrow down this wide range of objects based on user preferences. Understanding what possible gestures exist could also help shape future natural gesture based interfaces.

The related work relies on building custom devices that usually come in regular shapes and sizes.

2.3 Automatic recognition of object size and shape: *Vatavu and Zaicti [2013]*

We had outlined our interest in focusing on hand based deformation gestures when dealing with deformable everyday objects. Vatavu and Zaicti [2013] presented an in-depth work on using the posture of the hand while grasping objects to recognize object size and shape. They mainly draw their inspiration from the field of psychology to study changes in hand posture around while exploring and trans-

Hand postures are captured and analyzed on a set of rigid objects to infer size and shape.

lating a set of objects. They used rigid objects based on basic shapes (cube, parallelepiped, pyramid, sphere, cylinder and surface) to collect hand posture data. They then use various machine learning techniques to recognize size and shape of the objects. They mention citing current research, multiple benefits of exploring hands as a tool for the field of HCI:

They used rigid objects based on basic shapes.

- It helps further the design and development of natural gesture-based interfaces.
- Recognizing hand posture aids improving context-aware interaction.
- It could also be used as a way to retrieve information from ambient objects without the need to embed identification technology into them.

Multiple benefits of exploring hands as a tool for HCI research were listed.

We see our work in-line with this contribution to HCI as we attempt to understand deformable everyday objects around using hand based deformation gestures.

Chapter 3

Identifying Deformable Everyday Objects: A Survey

The first requirement is to identify deformable everyday objects. The range of deformable everyday objects that one interacts with in daily life varies depending on their location. For example, in a home environment, it is more likely that one finds a dishwasher bottle in the kitchen while in the bedroom one would find a moisturizer bottle (cf. Figure 3.1). They however share similar shape and affordances.

Similar shaped deformable objects maybe found in different locations.

To find a common set of deformable everyday objects, we decided to conduct a house-to-house survey with personal interviews. Apart from giving us a concrete list of everyday deformable objects, this also allowed us to have a glimpse into how people perceived these objects. Personal discussions on individual deformable objects helped us to group them and bring down their immense number into more representative study friendly groups.

We conducted a house-to-house survey to make a list of common everyday deformable objects.

Figure 3.1: A dishwasher bottle usually found in a kitchen and moisturizer bottle found in the bedroom. They share a similar cuboid like shape.

3.1 Goal and Procedure

The goal was to have a representative subset of deformable everyday objects from a home environment.

The goal of the survey was to have a representative set of deformable everyday objects from a home environment that could be used for an in-depth study.

A home environment could refer to multiple locations in the house, so the following four living spaces or scenes were considered in each home:

Multiple locations in a house were considered in the survey.

A: Living Room.

B: Bedroom.

C: In-House Office/Study.

D: Kitchen.

We divided houses surveyed in the study as far as possible.

Depending on the house, few or all of the above may result in a common scene. We tried to demarcate the spaces as far as possible depending on location of certain furniture, or how the user used a given space. If there was no separate in-house office room, area around the user's desk where one may have their laptop and/or stationary would be considered the in-house office. The participants were provided with the definition of what we considered a valid deformable everyday object for the study and explained with the use of a sponge (cf. Section 1.1.1 "Towards Deformable Everyday Objects"). They were then asked to point out at least ten objects in each scene that would fit the definition. The number ten was decided on the basis of a pilot study. The participants were encouraged to perform a deformable interaction with each object that caught their attention even if they would initially dismiss an object as similar to a preceding one.

Users were asked to point out at least ten deformable objects per scene.

The users were further asked to rate each object on two properties that we termed *Ease of Deformation* and *Volatility*. The rating was given on a 5-point Likert scale. The sponge was used as reference that was to be rated 5 for both properties. We defined the properties as:

Ease of Deformation rates how comfortable it is deform a certain objects with the user's hands. This could be user and material dependent.

We asked the users to rate the objects on two properties.

Volatility was defined as the property of a deformable object that allowed it to return to its original shape when the deformation force is released. This is just material dependent.

During the survey participants were asked to comment on objects that they would consider as similar deformable objects. This categorization was also done across the different scenes. This was to enable grouping objects together that may be located in different scenes but may have similar affordances (cf. Figure 3.1). Certain objects that occurred at higher frequency like bottles and books were given their own category. Depending on the survey they were further sub-categorized based mainly on the user responses on the additional properties.

An informal interview was conducted to help us create object groups.

3.2 Participants

A total of twelve houses were surveyed. They were chosen such that they covered various living situations such as

We surveyed twelve houses in total with participants aged 20–39.

family, student dormitory, and shared apartments. Only one participant per house was interviewed in the study. The participants were aged 20–39 (M = 25.58, SD = 5.23, three females). Seven of the participants were students with a Computer Science background and the rest were studying Mechanical Engineering and Economics. There was also an IT consultant with a family plus a participant who was employed as a restaurant cook.

3.3 Results

The observations and analysis helped place the objects into common groups.

The final list consisted of a grand total of 460 objects across the survey with around 260–270 unique objects[1].The analysis of the results focused on grouping objects under limited categories. Table 3.1 shows the categories, sub-categories, and some objects from the list. It only contains objects that had an observed frequency of occurrence higher than two. We subsequently present the rationale for the categorization.

3.3.1 Observations and Analysis

Participants would break down objects with different affordances into known simpler objects.

In general, all participants would initially dismiss objects having similar shapes, for example bottles. However, as soon as they hold and start to interact with the object they would start assigning them as a different category of objects due to with difference in quality of plastic (difference in the thinkness of the plastic). Objects that had different sub-parts with different affordances would be compared to the objects that offer the base affordance. For instance, many participants would liken a bag strap to a belt, and the flap similar to part of a jacket. Clothes on hangers would be compared to a hanging curtain. Participants would indicate that certain objects like cardboard boxes could be squeezed similar to a bottle but were shaped similar to a book, thus indicating the presence of a third category based on their interaction.

[1] A semi filtered list of objects per location can be found in the appendix (cf. Appendix A "Appendix for the Preliminary Survey")

Bottle	Cable	Books	Pillow	Couch(Fixed)	Foldables	Boxes/Utensils	Bendables*	Misc
Medium [b,k] (Detergent/ Syrup/ Dishwasher/ Moisturiser)	Thick* (Extension/ LAN)	Thin [l,o] (~1cm/ Folders)	All Pillows [b]	Couch [l]	Thin Blanket [l,b]	Plastic/Tin Cup/Can [k,o]	Plastic Ladle/ Spoon/ Scraper/ Knife [k]	Sponge [k]
Thick Plastic [k,o] (Cola)	Thin [l,b,o] (Phone Charger)	Medium [o] (~1-3)	Cushion [l,o]	Fixed Seat Cushion [l,o]	Curtain [l,b,o]	Plastic/ Cardboard/Tin Box*	Plastic/Steel Hanger [b]	Pocket Tissue Box [b,o]
Thin Plastic [l,k,o] (Oil/ Water)	Lamp Neck [b,o]	Thick [o] (>~3cm)		Mattress [b]	Carpet/Mat [l,b,o]	Plastic Bowl/ Steel Strainer [k]	Fly Swatter [k]	Tissue Roll [k]
	Vaccum Hose [l,b]			Bean Bags [l]	Cleaning Cloth/Towel/ Cloth Bag [k]	Plastic Tub/ Dustbin [l,o]	Wallet [o]	Spectacle/ Pen Case [b,o]
	Belt/ Watch Strap [b,o]			Thick Blanket [b]	Mouse Pad [o]	Basket [b,k,o]		Soft Toy [b,o]
					Paper/ Tissue/ Packaging [o,k]	Suitcase [b]		Whisk [k]
					Rucksack [l,b,o]	Lamp Shade [l,b,o]		
					Jacket [b]			

l - Living Room b - Bed Room k - Kitchen o - In-House Office * - All Locations

2-4	5-8	8+

Table 3.1: Major categories and sub-categories identified in the house-to-house survey. The colour coding reflects the frequency of the objects in the groups. The colours correspond to the frequency of occurrences of the listed objects whose legend is shown in the last row.

Objects were grouped together based on shape and frequency of occurrence noted.

Users also pointed out items like fruits, vegetables and potted plants as deformables, if they would be in the vicinity. Such perishable items were discarded during the initial screening phase of the object lists. The naming of the objects was based on the how the user described them, for example an *oil bottle* or *cola bottle*. The list of objects for a given scene was then analyzed to group similar objects together based on shape, and frequency of occurrence calculated.

Nine broad categories including a miscellaneous category were created to group all the objects.

Observing the various shapes and sizes of the objects, we created nine broad categories. As mentioned high frequency objects like bottles, books, cables, and pillows were given their own category. Objects such as a couch or a mattress that would usually have a fixed place in the house, were placed together in one category. All utensils and cuboid shaped objects like boxes were placed together. The objects in this category would have similar affordances as bottles. Objects that generally offered a bending affordance such as ladles or hangers were clubbed together. Certain objects that had occurred in significant frequency but could not be placed into the basic categories such as a tissue roll or a soft toy were placed under miscellaneous. Table 3.2 lists the created categories and the shape/size criterion they are based on.

Category	**Shape: [Size]**
Bottle	Cylindrical
Cable	Cylindrical: small diameter
Book	Cuboid: small
Pillow	Cuboid: medium
Couch	Cuboid: large
Foldable	Cuboid: thin
Box	Cuboid: hollow
Bendable	Cuboid: thin/narrow
Miscellaneous	Irregular shape/ Other objects which could not be classified

Table 3.2: Table lists the object categories and their corresponding shapes and sizes.

We created further sub-categories based on size or material to account for within-category differences in objects. For

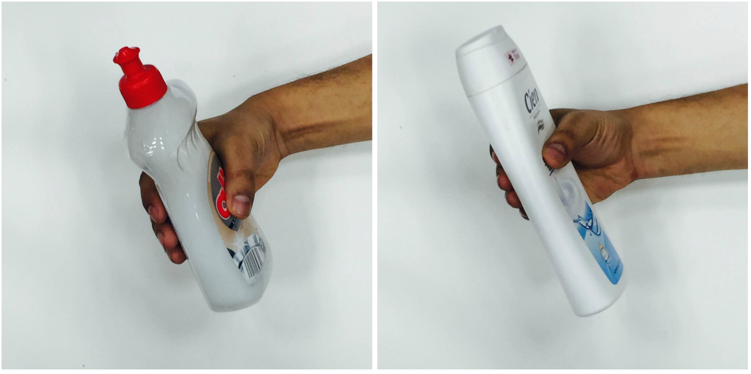

Figure 3.2: A dishwasher bottle and moisturizer bottle that are to be squeezed for their regular use.

example, bottles were further divided into three groups. One group included bottles with a thicker plastic shell such as cola bottles (All bottles in the study were made of plastic). Bottles that are squeezed for their regular use such as dishwashers and moisturizers were grouped together (cf. Figure 3.2). The third category included bottles that were in similar shape to the ones with the thick plastic body but were made of thinner material. They included mainly the cooking oil bottles or water bottles found in kitchens. Similarly books were divided into three groups based on the thickness: those roughly less than 1 cm, those between 2–3 cm, and greater than 3 cm (These are approximate dimensions). These were based on the survey where the users would identify magazines as different from other books. The sub-categories for pillows included pillows, and cushions. However the sofa and fixed seat cushions which are similar in terms of material to pillows were placed in a separate category as objects that users seldom move around the house.

Sub-groups were created based on size or material. Bottles had an additional category for squeezable objects such as dishwasher bottles.

Books were split into sub-categories based on their approximate thickness.

3.4 Summary and Representative Object Set

Observing the participants interaction with deformable everyday gave us a birds eye view of the possible gestures.

The survey gave us a birds-eye view of interactions with everyday deformable objects.

Thus helping us arrive at a broad categorization of everyday objects. However due to the non-uniform shapes of these everyday objects, it opens up the possibility of secondary and tertiary gestures on any given object. The follow up study was intended to give us an in-depth insight into possible gestures as perceived by users when presented with an everyday object.

We select a representative sub-set for further study.

To have a manageable number of objects for the study, we had to choose a representative subset of the objects identified in the survey. The objects were chosen such that they covered all locations under each category. Figures 3.3 and 3.4 show the objects selected for the next study.

Figure 3.3: Representative set of objects chosen for the qualitative study . -(1/2) A1: Thick Bottle, A2: Thin Bottle, A3: Medium Bottle, B1: Thick Cable, B2: Earphone Cable, B3: Lamp Neck, B4: Watch Strap, C1: Thick Book, C2: Medium Book, C3: Thin Book, D: Pillow, E: Couch, F2: Curtain.

Figure 3.4: Representative set of objects chosen for the qualitative study. -(2/2)
F1: Hand Towel, G1: Plastic Cup, G2: Cardboard Box, G3: Plastic Bowl, H1: Ladle, H2: Clothes Hanger, H3: Wallet, I1: Pocket Tissue Pack, I2: Tissue Roll, I3: Soft Toy, I4:Whisk.

Chapter 4

Gesture Elicitation on Deformable Everyday Objects: A Qualitative Study

The preliminary survey allowed us to identify deformable objects around us and offered a glimpse into how users exploit deformable affordances. The non-uniform shape and materials provide an opportunity for multiple deformable interactions. Therefore, we followed up with a study that allowed the users to explore the various affordances offered by everyday objects. We sought to answer the following research questions:

The goal of the study is to understand deformation gestures on the objects identified in the house to house survey.

1. What are the different affordances offered by everyday objects that enable deformations?
2. What are the common perceived affordances across users on a given deformable everyday object?

We decided to conduct a gesture elicitation study on a subset of objects obtained from the previous study (cf. Chapter 3 "Identifying Deformable Everyday Objects: A Survey"). Gesture elicitation studies have been used to define gesture sets for various tasks such as free hand TV control (Vatavu

We conducted a gesture elicitation study.

Gesture elicitation studies in current research usually have a defined gesture goal.

[2012]). Troiano et al. [2014] used gesture elicitation on deformable displays for tasks such as map navigation and 3D modeling. Such studies begin with a defined goal for each gesture. In contrast we are more interested in finding the kind of gesture suggested by an object independent of any desired goal. We are also trying to identify the parameters that define a gesture for a given object.

4.1 Study Protocol

We recorded gestures elicited from the users on deformable everyday objects.

Since this study involved a closer observation and analysis of gestures on a given object, we decided to record the actions of the users on camera to help us conduct the analysis. The objects used in this study were a representative subset of results from the previous study (cf. Chapter 3 "Identifying Deformable Everyday Objects: A Survey"). The objects chosen, covered all basic shapes and location occurrences as observed in the survey.

4.1.1 Setup

We simulated a living room environment with a couch and a coffee table.

We wanted to keep the test environment as close to what one could expect in a home and had fixed camera angles across users. We set up the test area in a part of our laboratory with a couch and coffee table. We ensured that there was a curtain hanging within arms length of the user as well (The couch and the curtain were also a part of the study object list). The participants were filmed using two GoPro™wide-angle cameras (cf. Figure 4.1). The objects themselves were hidden from the user and presented one at a time according to a Latin square.

Independent and Dependent Variables

The independent variables for the study were the 24 representative objects from the house-to-house survey (cf. Figures 3.3 and 3.4).

Figure 4.1: The study setup for the exploratory study. We simulated a part of a living room environment with a couch, a coffee table, and a curtain at arms length of the user. The gestures were recorded using two GoPro™ cameras.

The dependent variables were the elicited deformation gestures. These represent the perceived affordances that suggest deformation.

4.1.2 Procedure

The participant was presented with the definition of a deformable everyday object and also what we considered as a deformation gesture (cf. Section 1.1.1 “Towards Deformable Everyday Objects”).

The objects were presented to the user one at a time. The

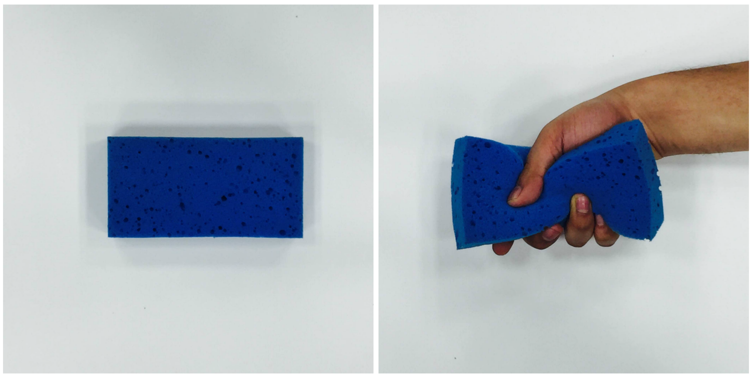

Figure 4.2: A sponge used as a reference of an object that had a Likert scale rating of 5 for ease of deformation.

Users had to perform 3–4 gestures on each object. They were free to hold the object as they liked.

user was instructed to perform the first gesture that came to mind and was discouraged from handling the object before beginning the gestures. The user was asked to perform around 3–4 distinct deformable gestures. Use of different number of hands or fingers, but with similar actions were to be considered separate gestures. The users were free to keep the object on the table or lift it up in case the gesture was not dependent on its location. The user was asked to signify the approximate beginning and end of a gesture by saying *OK* and *Done* or another predetermined indicator words. In case a gesture could be performed independent of its location *on* or *off* the table, the user was asked to repeat the gesture under both conditions. This was asked in order to observe the role played by such a support structure. For gestures such as bending a ladle with one hand using the table as support or bending use both hands in mid-air this was exempted.

Users rated ease of deformation on a Likert Scale.

The user was asked to rate the ease of deformation for each gesture on a 5 point Likert scale. To give the user a frame of reference, squeezing the sponge (Figure 4.2) was given as an example of 5 (maximum) while a peanut can (Figure 4.3) was given as an example of 1.

At the end of the study the users were asked to rate the order of priority on what influences their gesture type before interacting with the object:

Figure 4.3: A peanut can used as a reference of an object that had a Likert scale rating of 1 for ease of deformation.

- Shape
- Size
- Material

User perception on what may have influenced their gestures was ranked.

The users were also asked to comment on their experience with individual objects, and if any conflicted with their initial perception before interacting with the object. Users seemed apprehensive while using certain objects for deformation. They were nevertheless encouraged to perform the gesture. This was then discussed during the feedback and used as a basis for filtering the object list.

An informal discussion was conducted at the end of the study.

4.1.3 Method of Analysis

A coding system was set for analyzing the gestures in the videos. Each gesture was coded as a combination of various primitives that gave information on how many hands and fingers were involved in the gesture as well the resultant action/deformation. It also would indicate, if during a bimanual interaction, both hands were actively involved or if one was purely used for support. The following syntax was used for representing a deformation gesture:

A coding system to analyze gestures was fixed based on observations during the study.

$[Hands][ActiveFingers][Deformation/Action][Side(optional)]$

Explanations of the syntax parts are given below along with used codes in brackets along with an example at the end.

- Hands Indicator
  - Uni manual interaction (1)
  - Uni manual interaction with the table used as support (1′)
  - Bi manual interaction (2)
  - Bi manual interaction with one hand purely used as support (2′)
- Active Fingers Indicator
  - 1–X: (where X indicates all 10 fingers)
  - Palm (P)
- Deformation/Action
  - Squeeze (S): Thumb and rest of the fingers in a hand moving towards each other.
  - Pinch (Pi): Special form of squeeze with thumb almost in contact with the index finger.
  - Push (P): Pressing against the object, away from the user.
  - Pull (Pu): Opposite of a push.
  - Bend (B)[f: forward b: backward]: The object being shaped into a curve.
  - Fold (F): Special form of bend with one part of the object covering the other.
  - Twist (T): Deform the object by turning hands in opposite directions to each other.
  - Roll: Special form of twist and bend with a flat object deformed into a cylindrical shape.
- Side: Indicates alternate sides of an object that may not have a regular shape.

The coding system indicates the number of hands, fingers as well the deformation/action performed by the user.

A SAMPLE CODE: 2′5BF:
This code indicates that the gesture was performed using two hands where one hand was used for support. All the five fingers of the active hand were on the object and the deformation was a bend in the forward direction.

Example: *A sample code: 2'5Bf*

The coding system presented here is the refined and final version based on the observations during the study and an initial pass of the study videos. It helped us group gestures by various users and helped disregard discrepancies in the orientations of the object under study or minor differences in the extent of a resulting deformation. The gesture codes observed in the study could be categorized into three different groups:

The codes help group gestures into comprehensible groups.

Gesture Group A (GG A): This includes gestures like the one handed squeeze (1$\bar{5}$S) where the thumb and the other fingers of each hand move relative to each other. In a two handed gesture like a 2XS, both hands could be considered independent of each other.

Gesture Group B (GG B): The two handed bend (2XB) would fall into this category where the left and right hand move relative to each other after the object has been grasped. Two handed interactions where one hand is used for support could also be placed in this category.

The observed gesture set could be divided into 3 gesture groups based on how the hands and fingers move relative to each other.

Gesture Group C (GG C): Gestures that use the table such as 1′5B as support are included in this category. The hand movement relative to an external frame of reference distinguishes this group from the others.

4.2 Participants

A total of twelve participants were a part of the study aged 24–30 (*M* = 26.67, *SD* = 1.83, four females). All of the participants were University students with everyone with Computer Science background, except one. To account for the discrepancy in the Latin square due to half the number of participants as treatments (in our case the number of objects under study), we used alternate combinations in the Latin square. Only two of the twelve participants were left handed. Three of the participants were also a part of the house-to-house survey but we did not think that would influence any decision as there was no specific gesture related activity in the previous task.

Twelve participants took part in the study. Alternate combinations in Latin square were used to compensate for discrepancy.

4.3 Results

Over 288 gestures were generated for 24 objects with around 4 gestures per object.

Close to 480 minutes of video footage was recorded during the study. The gestures in the videos were encoded using the scheme described under Section 4.1.3 "Method of Analysis". The users generated over 288 gestures for the 24 objects. It was observed that very rarely users would generate more than four gestures for any given object, so gestures apart from the top four were discarded for the final analysis. Table 4.1 lists the top two gestures in terms of frequency for each object with a third gesture mentioned in case the top two gestures had the same deformation/action type.

Users tend to start off with single hand gestures and attempt to pick the object off the table.

It was observed that users would tend to use the whole hand while interacting with the objects. Only after the users attempted most of the gestures involving the whole hand, would they go exploring using combinations for fingers. This was however not applicable to smaller size objects such as the watch strap. Users would usually attempt to begin with a single handed gesture for most objects and attempt to lift the object *off* the table first. However, for objects like a ladle that was hard to bend using a single hand in mid-air, the users would use both hands. The use of a table for support was only explored as a secondary gesture.

Users had similar ratings for same gesture on or off the table.

Figure 4.4 shows the cumulative *ease of deformation* ratings across the objects for gestures performed with the object *OFF* and *ON* the table surface (M_{OFF} = 3.92, SD_{OFF} = 1.65, M_{ON} = 3.99, SD_{ON} = 1.12). We see that the means and the confidence intervals of the two conditions are quite similar. Participants however, preferred to pick the object *off* the table as far as possible.

Users rated material having a slight influence on their initial gesture.

Friedman test showed a statistically significant difference in the users ranking of what influenced their gesture , (χ^2(2,N=12) = 7.17, p <.05, ϕ=.54). Post-hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at p <.017. There were no significant differences between *Shape* and *Material* (Z = -1.098, p >.017) or between *Shape* and *Size* (Z = -1.647, p >.017). Though there was no sig-

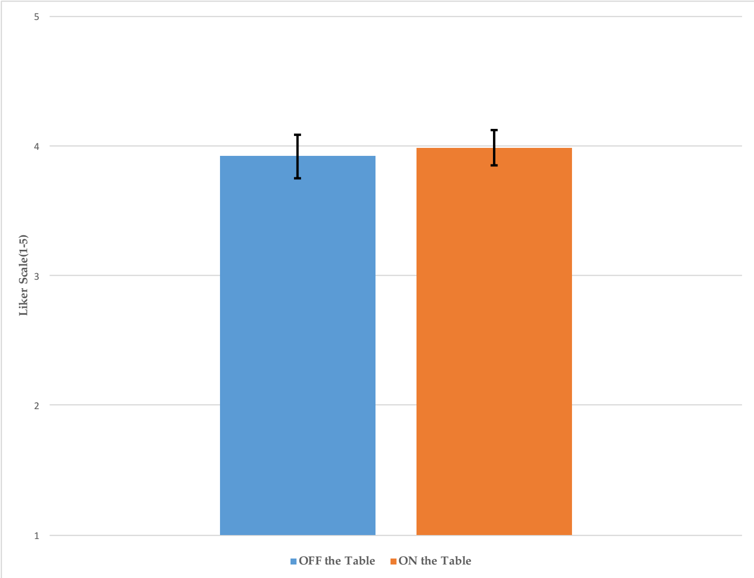

Figure 4.4: Mean Likert scale rating for *ease of deformation* performing a gesture on the same object with *OFF vs ON the table* conditions. The bars show 95% confidence intervals. We see that the two conditions are very close in how users rate the *ease of deformation*.

nificant difference between *Size* vs *Material*, we could see a trend towards *Material* having certain influence (Z = -2.75, p = 0.02).

The discussion at the end of the study along with the observations provided us with some insight into the user's mental model. Users who were presented with the thick bottle after the thin bottle said they were surprised by how harder it was to perform the same interaction. A user who had the thin bottle second, mentioned that she was more tentative with deforming bottle due to the experience with the thicker bottle. Users would perform similar gestures on the tissue roll and the bottles though they are made of different materials (cf. Figure 4.5). This indicated a trend towards the shape playing an important role relative to the material. The participants mentioned they would have to think for a moment before performing a second or subsequent gestures on the objects. The first gesture would in general be a gesture they may have been performing in their daily

The observation however suggested users may be influenced more by the shape of the object.

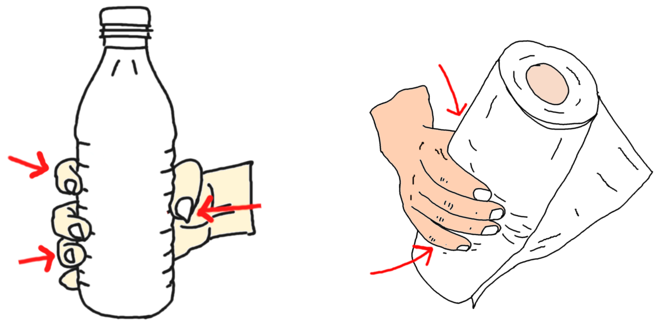

Figure 4.5: Similar gestures on objects with similar shape but different materials. Users squeeze the bottle with one hand (Left). A similar one handed squeeze was also observed on the tissue roll (Right)

lives.

4.3.1 Object and Gesture List for Quantitative Study

We select objects that have the most common gestures among the twelve participants.

To answer the second research question of having a set of objects that consistently elicit a set of common gestures across users, we selected the objects with the highest number of common gestures (cf. RQ2). Since we had twelve participants, the best case scenario would be a gesture that had a frequency of twelve. In case the highest frequency for a gesture on a given object was less than twelve, we would look for the gesture with the second highest frequency. Only objects with at most two different gestures, were filtered from the study.

Participants did not wish to deform personal items at all.

We also took into account the discussion with the users on their overall experience with the objects. Users were wary about deforming leather objects such as the watch strap and the wallet, along with the soft toy wich they considered too personal. The whisk was also thought to be too delicate to be deformed. The thick book was considered too unwieldy as an object for deformation. The cardboard box was quickly damaged after a couple of gestures.

Object ID	Object	Gesture I	Count	Gesture II	Count	Gesture. Alt	Count
Bottle							
A1	Thick Bottle	15S	12	2XS	9	2XB	3
A2	Thin Bottle	15S	12	2XS	9	1PP	4
A3	Medium Bottle	15S	11	2XS	10	2XB	5
Cables							
B1	Thick Cable	2′5F	7	2XF	6	15S	5
B2	Thin Cable	2′5F	10	2XF	8		
B3	Lamp Neck	2′5T	11	2′5Bf	7		
B4	Watch Strap	14′1B	9	2XB	3		
Books							
C1	Thick Book	2XB	12	15Flip	6		
C2	Thin Book	2XRoll	10	15F	7		
C3	Medium Book	2XB	9	1′5B	4		
Pillow							
D	Pillow	2XF	8	2XS	6	2PP	5
Couch – (Location Specific)							
E	Couch	15Sa	8	2XS	7	2PP	7
Foldables							
F1	Hand Towel	2XF	8	2XT	9		
F2	Curtain	15S	10	2XT	6		
Box							
G1	Cup	15S	9	1′PP	7		
G2	Cardboard Box	2XS	12	15S	8	1′PP	4
G3	Bowl	2XS	12	1′PP	4		
Bendables							
H1	Ladle	1′5B	12	2XB	10		
H2	Clothes Hanger	15S	11	2XB	11		
H3	Wallet	2XB	11	15S	6		
Miscellaneous							
I1	Pocket Tissue Pack	15S	9	2XS	5	2XB	6
I2	Tissue Roll	15S	11	2XS	10	1PP	6
I3	Soft Toy	2XS	10	15S	9	2XT	6
I4	Whisk	15S	9	2′5S	4	2′PP	4

Table 4.1: All the objects used in the study and their associated gestures are listed here. Gesture I and II were the top two gestures in terms of frequency across the elicited gesture set. In case both gestures had a same action/deformation code, we list Gesture Alt. as the next gesture with the highest frequency for that object. This has a different action/deformation. The coding system is based on the description in Section4.1.3. The *a* in the gesture for couch represents the armrest. **Note: The Pillow class is referred as Sponge.**

The final list had eleven objects with two gestures per object.

We wanted a representative object and gesture combination for each category. Objects with similar shapes were replaced with one object (Size was not taken into account). The final list of objects and the associated gestures are listed in Table 4.2. Figures 3.3 and 3.4 illustrate the observed deformable gestures on the objects.

Category	Object	Gesture	Gesture Group
A: Bottles	Thin Bottle	15S	GG A
		2XS	GG A
	Medium Bottle	15S	GG A
		2XS	GG A
B: Cables	Thin Cable	2′5F	GG B
		2XF	GG B
	Lamp Neck	2′5T	GG B
		2′5Bf	GG B
C: Book	Medium Book	2XB	GG B
		1′5B	GG C
	Thin Book	2XRoll	GG B
		15F	GG A
D: Pillow	Pillow	2XF	GG B
		2XS	GG A
F: Foldables	Hand Towel	2XF	GG B
		2XT	GG B
G: Box	Plastic Bowl	2XS	GG A
		1PP	GG C
H: Bendables	Ladle	1′5B	GG C
		2XB	GG B
	Clothes Hanger	15S	GG A
		2XB	GG B

Table 4.2: This table lists the objects and the gestures filtered from the qualitative analysis. The gesture groups mentioned in section 4.1.3 “Method of Analysis” are noted alongside. **Note: The Pillow class is referred as Sponge.**

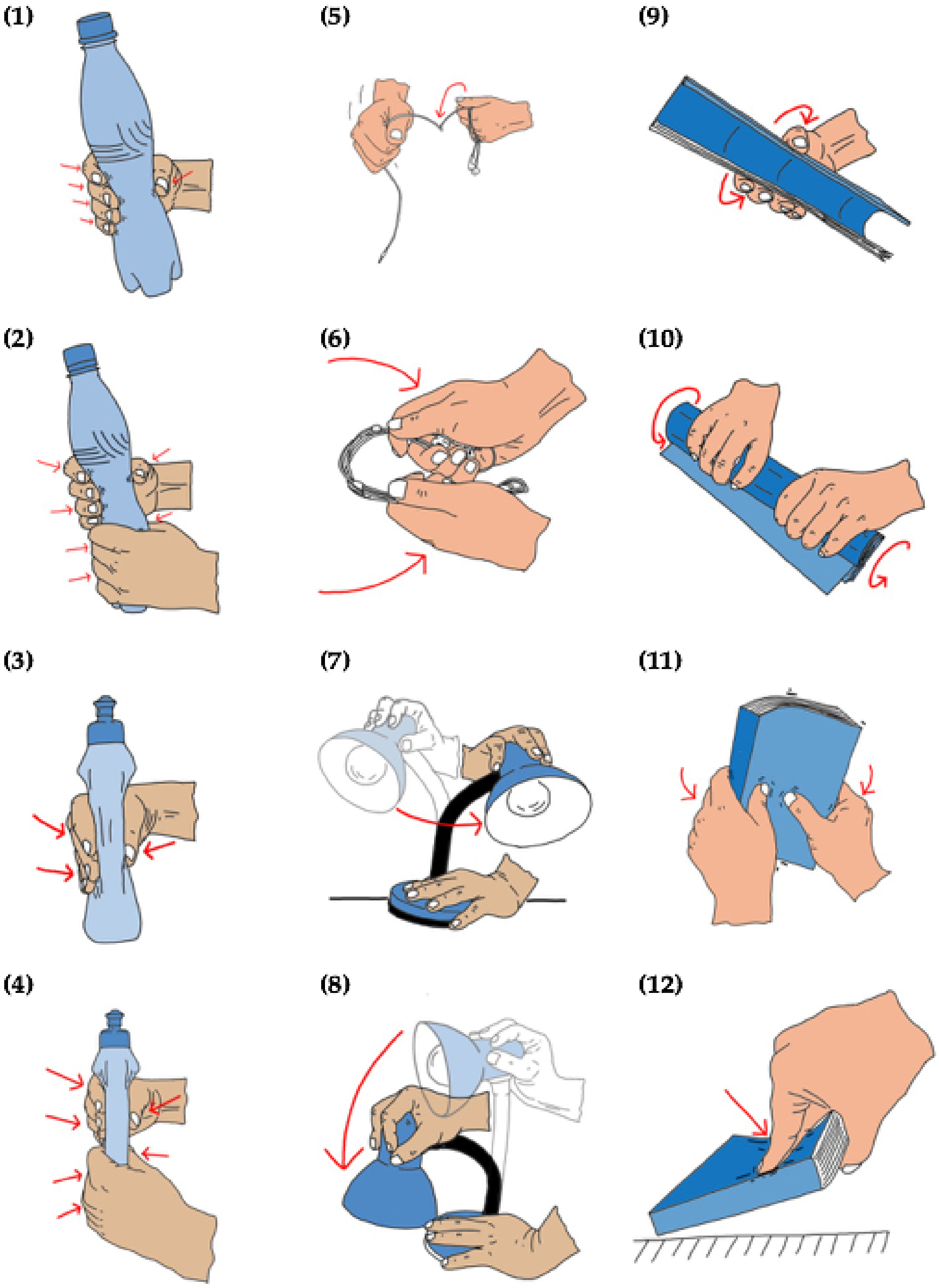

Figure 4.6: Deformable gestures on the list of objects chosen for the quantitative study -(1/2). (cf. Appendix B.3 and B.4 for descriptions). (1)TB15S, (2)TB2XS, (3)MB15S, (4)MB2XS, (5)TC2'5F, (6)TC2XF, (7)LN2'5T, (8)LN2'5Bf, (9)TBk15F, (10)TBk2XRoll, (11)MBk2XB, (12)MBk1'5B.

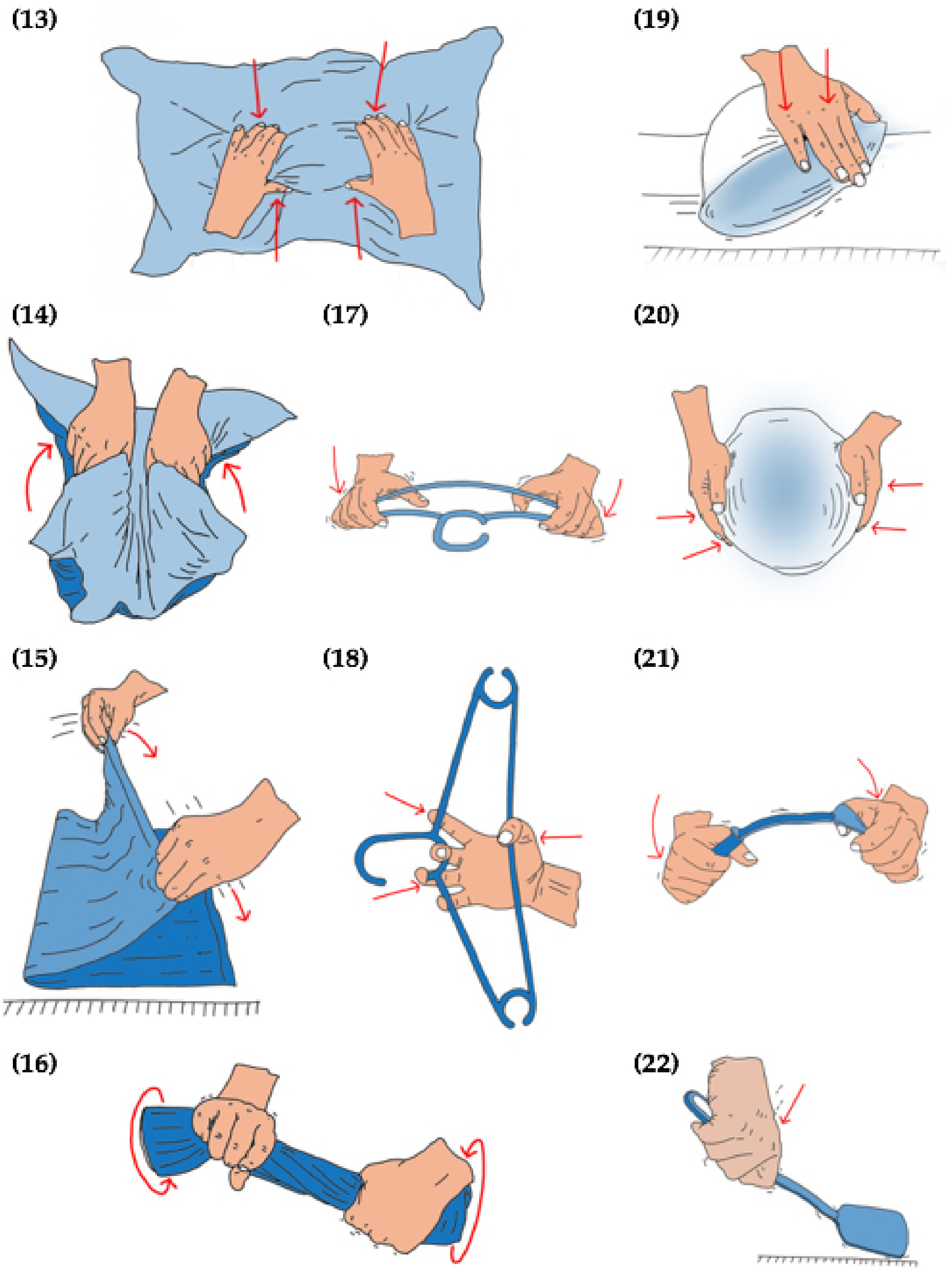

Figure 4.7: Deformable gestures on the list of objects chosen for the quantitative study -(2/2). (cf. Appendix B.3 and B.4 for descriptions). (13)P2XS, (14)P2XF, (15)HT2XF, (16)HT2XT, (17)CH2XB, (18)CH15S, (19)PB1PP, (20)PB2XS, (21)L2XB, (22)L1′5B.

Chapter 5

An Overview of the Tools Used for the Quantitative Study: An Excursus

The qualitative video data based user study allowed us to have a detailed insight into how users perceived individual deformable everyday objects. It also helped us to further group together certain objects based on how users interacted with them even though they might have differed in material or texture.

The video data based showed how users perceive affordances on deformable everyday objects.

As seen in the previous chapter (cf. Chapter 4 “Gesture Elicitation on Deformable Everyday Objects: A Qualitative Study”), we could divide the whole scene into the users actions and its effect on the object. The effect on the object would involve trying to measure the extent of deformation or quantifying the change with respect to the object at rest. For certain gestures such as squeezing, this is hidden under the users hand and differs from person to person depending on the force they would have applied. The users actions on the other hand, can be more generalized across users as denoted by the coding scheme we adopted.

We observe from the previous study that users actions could be used to code deformation gestures.

We wanted to quantify the users actions as depicted by the

A camera based motion capture system was used to capture millimeter precision trajectory measurements.

coding scheme from the previous study. To get accurate trajectory measurements of the relative movements of the fingers and hands, we decided to use a camera-based motion capture system. This allowed us to capture millimeter precision data of the movements as three dimensional co-ordinates. We then wanted to see if it was possible to classify the object categories or the objects given the gesture data. For this analysis we used the Random Forest machine learning technique (Breiman [2001]).

5.1 The Vicon Motion Capture System

Retro-reflective markers are tracked with a multiple camera motion tracking system.

We used the infrared (IR) light emitting Bonita camera system from Vicon[1] for our motion capture needs. This is a passive optical motion tracking system that relies on the use of retro-reflective markers placed on the subject to be tracked. A multiple camera setup is placed around a tracking area (cf. Figure 5.1). Each camera only sees a two dimensional image with the marker showing up as pixelated blobs against a dark background. The data from the cameras are fed to the Nexus software also provided by Vicon for further processing. The cameras are calibrated using a calibration stick with fixed markers and set a defined origin in the capture volume. This step helps the Nexus obtain the positions and orientation of the cameras with respect to the capture volume or tracking area. When at least two cameras see a marker, the three dimensional position of the marker can be determined by the system.

The Nexus software can automatically use an algorithm or copy trajectories from other visible markers to fill gaps.

The Nexus software allows us to label the markers and define a fixed subject consisting of multiple markers. The software then stores the trajectories as 3D co-ordinate data over time. Since complex movements mean that not all the markers may be visible at all times, there are gaps created in the captured data. The Nexus software also includes some basic gap filling tools that can be done automatically using an algorithm or a more tedious technique that copies trajectories from other visible markers[2]. We can also obtain

[1] http://www.vicon.com/bonita
[2] Automatic technique was found unreliable for gap size >10

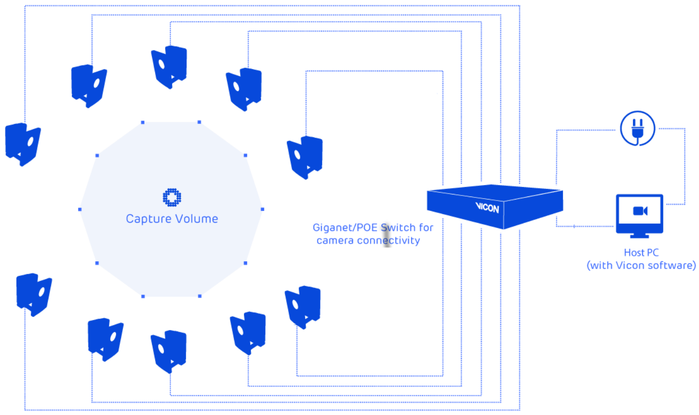

Figure 5.1: At typical Vicon camera Setup. The Vicon cameras are arranged around a capture volume (Left of image). The object to be tracked should be in this area. The cameras send their data to the Nexus software on the connected computer.

voltage and acceleration data of the markers that can be exported together with the position data as comma separated value (csv) files.

5.2 The Random Forest

Random forest is an ensemble machine learning technique based on decision trees. A decision tree tries to partition the given data by inherently asking yes/no questions depending on the target labels. This splitting goes on recursively till the leaf nodes have values identifying the targets. Given a test input, the decision tree is then traversed based on the test attributes till it reaches a target and the test sample is classified accordingly.

A Random Forest is a machine learning technique based on decision trees.

A random forest is a collection of such decision trees that are created by a random sampling of the training data. Each tree votes for a classification label based on the test input. The random forest classifier then assigns the class that receives the highest number of votes (cf. Figure 5.2). The

Multiple decision are created from samples of the same data.

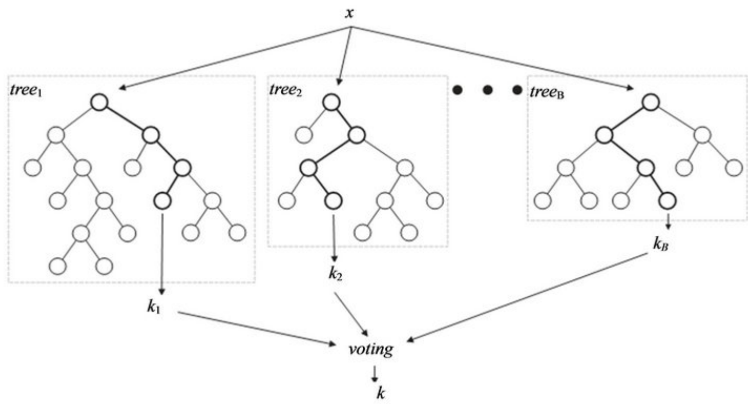

Figure 5.2: Combination of multiple decision tree voting results. Figure adapted from (Nguyen et al. [2013])

building of multiple multiple trees helps compensate for variances in different samples of the training data. We used the random forest implementation in sci-kit learn python package (Pedregosa et al. [2011]).

We use a splitting criterion called *Gini Impurity* as a parameter.

The nodes of the decision trees are split on measure called *impurity* that in our use will be the *Gini Impurity* (Stat.berkeley.edu [2016]). Since a random forest consists of a multitude of decision trees built using randomly drawn features from the complete training set, it is not easy to examine each tree separately and interpret the results. The random forest classifier in Sci-kit Learn exposes the feature importance measure. It is an averaged value of the splitting criterion for each feature in the nodes of the decision trees. When one of two co-related features is picked by the random forest, the importance of the other could be reduced for the given model. Thus, a redundant feature may be filtered from the feature importance list.

Chapter 6

Capturing Deformation Gestures on Everyday Objects: A Quantitative Study

The qualitative study outlined in Chapter 4 "Gesture Elicitation on Deformable Everyday Objects: A Qualitative Study" gave us a list of the most common gestures performed on the set of deformable everyday objects. We were also able to use the video analysis to identify how deformations on objects could be expressed using fingers, hands and resulting actions. The study also helped us gain a better insight into how certain objects could be grouped together based on the gestures. Our goal is now to see if quantified representation of the gesture could be used to validate our observations.

Our goal now is to see if quantified representation of the gesture could be used to validate our observations.

We used the Vicon system (cf. Section5.1 "The Vicon Motion Capture System") to generate gesture data sets. The data set includes trajectory information on hand and finger movements. We also stored the corresponding velocity of the acceleration values of the tracked markers. We then used the Random Forest technique to classify objects based on the gestures.

Deformable gestures were quantified using a motion capture system.

6.1 Study Protocol

This study focused on capturing pre-defined gestures.

An overview of the tools used for the study has already been described in Chapter 5 "An Overview of the Tools Used for the Quantitative Study: An Excursus".This study focused only capturing data of the final set of pre-defined gestures identified in Chapter 4 "Gesture Elicitation on Deformable Everyday Objects: A Qualitative Study".

6.1.1 Setup

The objects were placed on a table with the user facing a display.

The gestures under study constitute a varied set of hand positions and multiple motion paths. The goal of the setup was to have a detailed capture of the hands and as well as avoid as many gaps in the capture data as possible. The objects for the study were placed on a table that occupied the space of the calibrated tracking volume (cf. Figure 6.1). The chair for the user was not height adjustable but the user could move it forward or backward as required. The origin for the capture volume was located at the left bottom of the table. The user was sitting facing a computer display which would be playing the gesture required for the user to perform (cf. Figure 6.1).

The Vicon system had seven cameras in total.

The Vicon camera system used for the study had seven cameras in total. Six cameras were arranged close to a circle formation focusing on the tracking volume. Three of the cameras were placed on the right side of the tracking area with two on the left. One camera was placed opposite where the user was sitting and another was placed looking down at the tracking area from the ceiling. The Nexus software captured the data at 100 frames per second.

In prior studies users would not constantly indicate start and end of gestures.

In the prior study (cf. Chapter 4 "Gesture Elicitation on Deformable Everyday Objects: A Qualitative Study"), users were asked to signify the start and end of a gesture by saying *OK* and *DONE*. However, it was generally observed that the users would sometimes forget to mention the signifiers or use alternative ones such as *START* or *END*. To avoid this ambiguity for the data capture study, the control

Figure 6.1: The setup for the quantitative study (Left). The user sat facing a display that played a video of the gesture to be performed (Right).

Figure 6.2: The users hands were to be in the area marked with the black tape at the beginning and the end of the gesture (Left). The table had dimensions of 70x60 cm while the marked areas were approximately 15x20 cm. The origin is at the bottom left corner. The user had to tap the button with the feet to start or stop the capture process (Right).

of the capture phase was given to the user directly. This was done through the use of a button placed under the table, that could be operated by the users foot (cf. Figure 6.2). The press of a button simulated a mouse click on the Nexus software machine. The user had to press the button to start the capture and then once again when the gesture ended.

The users had to press a button with their feet to start, and end the motion capture.

To ensure that all the capture data would include the com-

Users had to place their hands in marked areas at the start and end of every gesture.

plete gesture information and avoid ambiguity across users on what constitutes the start and end of a gesture, the users were to start and end their gestures from a common position. This was done by marking two areas on the table where the users had to place their hands at the beginning and the end of the gesture.

Eleven objects filtered from the qualitative study were used in this study.

The final object list of eleven objects from the previous study was a part of this study. Certain objects like the bottles were painted to decrease interference due to unintended reflections that could be captured by the Vicon cameras. The clothes hanger used in the study was also sandpapered as the original texture was reflecting ambient light at times. The Table 4.2 lists the objects and the associated gestures used in the study. The gesture codes were prefixed with an identifier for the objects. A detailed definition of each gesture encoding is provided in the appendix (cf. Figures B.3 and B.4).

Study Groups

The pilot user showed signs of fatigue due to the duration of the study. So we split the study int two parts.

The pilot study conducted had a duration of over one and half hours. The user showed signs of fatigue and disinterest over the end. Since the goal of the study was to gather data and objects were mutually exclusive, the study was split up into two parts. The objects were divided into two groups of six and five. Thus we had twelve gestures for the study and ten for the second. The grouping was done so as to get similar gestures across objects grouped together.

Group I: Thin Bottle, Medium Bottle, Thin Cable, Medium Book, Ladle, Clothes Hanger

Group II: Lamp Neck, Thin Book, Pillow, Hand Towel, Plastic Bowl

Marker Setup

Since we were only capturing the hand and finger movement, the Vicon system did not need to have any information about the object itself. The retro-reflective markers had to capture the precise hand and finger movements. There were two options to accomplish this, either place the markers directly on the users using double sided tape/adhesive or use a glove augmented with the markers. We decided to opt for the glove solution as it had the following advantages:

We placed the retro-reflective markers on a glove that was same for all the users.

1. It would enable us to have a constant marker location across users.
2. It would enable a quick setup as the users would have to just slip the gloves on instead us attaching each individual marker for different users.
3. The users would not be subject to discomfort of the adhesive tap removal after the study.

To make sure the glove fits varied hand sizes, the glove chosen was of stretchable material. To account for the minor changes in distances between markers across users, the Nexus system has a calibration facility. The Vicon system enables us to create a *template* for a given set of markers which includes information on all the markers and subjects. This template can then be *fit* to different sized hands and the software creates an user specific skeleton.

The glove chosen was made of stretchable material.

The complicated and varied nature of the gestures under study means the Vicon system may not be able to track all markers at all times. In certain conditions, the markers get swapped as well. To compensate for this, we used multiple markers to track each part of the hand under study. This creates redundancy in case markers go missing, and also help us during the manual reconstruction process post-study. One of the assumptions from the qualitative video study was that for most of the gestures, the four opposing fingers move together. So, for this study, we considered the four fingers as a unit. This helped us reduce the markers

We used multiple markers to account for redundancy as the complex movements may lead to missing markers in Vicon data.

Figure 6.3: Gloves used for the study with ten markers on each hand. The marker layout was varied to aid during the manual gap filling post study.

and consequently reduce the manual post-processing work that would have otherwise added considerable overhead. Figure 6.3 shows the gloves with the markers used for the user study. Each hand is divided into four main segments with multiple markers placed on each segment:

We use markers on on certain fingers and the dorsal area of the hand.

1. Thumb
2. Index finger
3. Ring finger
4. Dorsal area of the hand (Back of the hand)

The layout for left and right gloves was slightly different.

The left and the right hand gloves differ slightly in the number of markers placed on the back and the thumb. This was to help us for a better identification during the post-study manual corrections.

6.1.2 Procedure

As already mentioned, the study was conducted in two parts. However, the procedure was identical. Only the objects and the corresponding gestures differed across both studies.

Procedure for both groups was identical.

The user was presented with the definition of deformable objects and deformable gestures (cf. Section 1.1.1 "Towards Deformable Everyday Objects"). Before the study began, it was explained that the user had to rate each gesture on the *Ease of Deformation*. This had to be rated on a five point Likert scale ranging from *Strongly Agree* to *Strongly Disagree*. The same sponge and a can of peanuts from the previous study were used as examples for both extremes (Figures 4.2 and 4.3). The user was also asked to answer in *Yes* or *No*, whether the given gesture was also one of the first few gestures that the user would have done. This was to help us get a validation of the resultant data from the previous study. In case the user answered *No* for both gestures, it was followed up with an informal interview on what the user would prefer as gesture on the object.

The user was presented with the definiton of deformable objects and gestures.

To familiarize the user with the use of the button, the user was given a small task to perform a deformation gesture on the sponge. The user had to press the button, squeeze the sponge or do any other gesture and then press the button again. The user was then presented with the object and shown a video of the required gesture. The video included the whole sequence of the gesture from the hand at rest position to the handling of the object and then placing the hand back to the start position. As the study was to capture quantified data on the pre-determined gestures, the users were asked to refrain from deviating too much from the video. There were no restrictions on the choice of hand for single handed gestures, how the hands were positioned, the orientation of the object, the placement of the object *on* or *off* the table, if applicable or where the object was held. The object was however required to be in the general tracking area above the table.

The user was explained the use of the button and allowed a trial gesture to get acquainted with the system.

The first attempt was considered a trial, even though the

The user had to repeat the trail six times including the trial.

data was recorded. In case the user had some confusion or did something different from what we were expecting entirely, we deleted the trial and the process was repeated. The user then had to repeat the gesture another five times. This meant we had at least six capture files for each gesture. Each gesture file comprised of the entire sequence as described in the video. After the first attempt, the user was asked to answer the two questions, on ease of deformation and gesture preference.

The user was asked to fill a small feedback form at the end to rate study experience.

At the end of the study, the user was asked to fill up a small feedback form which rated the users experience with the study. We conducted our study in a controlled environment with all objects placed on the table in front of the user. This was necessary for the Vicon setup and made an in-house study unfeasible. This may differ from how the objects would usually be found in a home environment. The users also had to wear the marker gloves and interact with the button placed at the foot. The questions asked were to aid us in analyzing the users perception on how the setup may have affected their gesturing interaction. The feedback forms can be found in the appendix (cf. Figures. B.5 and B.6).

6.1.3 Method of Analysis

Gap filling was done using the Nexus software.

The data obtained from the Nexus software had position, velocity and acceleration information on all the markers. As described in section 5.1 "The Vicon Motion Capture System" gap filling was done using the software. Bigger gaps meant more manual reconstruction effor was required. Gaps which still existed were replaced with null values. The multi-level header for each file was replaced with an indicator for each column. This is the final data set which would be provided at the end. The files were also named accordingly describing the user number, object and gesture combination.

We divide each file into equal chunks.

Different trials could have different lengths due to the time taken for each gesture. This means each file has unequal number of frames or rows. We divided the data in all files into equal number of slices and took the median of each slice as the feature set for the classifier. This allowed us to have uniform length features for all trials. We called this the chunk size. For our analysis we chose to focus only on the position data of the trajectories. All the rows of each file are then concatenated with each row representing the complete gesture file. The following example describes chunk size 2.

We only focus on position data for our analysis.

CHUNK SIZE 2:
The position data for each gesture is made up of 60 features (X,Y, and Z co-ordinates for the 20 markers). Each file will have 100 rows for each second of the gesture duration. A chunk size of 2 means that the data in each file corresponding to a gesture is divided into two halves. In case the number of rows is not even, the first half will contain an extra row. The median values of each half are then taken as representative features. We now have a representation of the gesture in two rows. These two rows are then concatenated to form one row. Thus we represent the gesture with a feature vector of size 120.

Example: *Chunk Size 2*

Each row is then accordingly labeled in the following ways:

1. An unique label for each object and deformation gesture combination according to the *ID* and *Gesture* coding shown in Lists B.3 and B.4 (e.g., TB15S).
2. A label for the corresponding deformation gesture (e.g., 15S).

We labeled trials using two ways.

An input matrix where each row corresponds to a label is passed on to the classifier. We try to understand if the gesture sequences can be used to classify individual objects using label method 1. Label method 2 would help analyze the possibility of classifying the group of objects that afford similar gestures.

A matrix of the labeled data was passed to the classifier.

The data is then partitioned into training and test data using the stratified cross validation technique (Kohavi [1995]). Stratified cross validation helps ensure that each label in the multi-class sample is represented in the partitioned data.

We use stratified k-fold cross validation.

The partitioned data is passed to the Random Forest classifier which had the following fixed parameters[1]:

n_estimators: 300 (This is the number of decision trees that are grown.)

We fix Random Forest parameters for our analysis.

[max_features:] 30 (This is number of features that are picked for every split.

We use 10-fold cross validation and present the accuracy as the mean of ten results.

6.1.4 Participants

22 participants took part in the study across two groups.

A total of 22 participants took part in the study divided across both studies. The number of participants was equal to the number of conditions which in our case was the number of gestures. So Group I had twelve participants while Group II had ten. All but one of the participants, were university students between aged 22–37 (M = 26.68, SD = 3.58, seven females). 17 of the 22 participants had a background in computer science. There was one student each with a background in mathematics, communication engineering, simulation science, electrical engineering and biomedicine. One of the participants was employed as a software engineer.

6.2 Evaluation

We had 1451 useful trial files.

At the end of the study we had a total of 1451 trials which were used for our analysis. There were an unequal number of trials across gestures as some files had to be discarded due to incomplete data. We tested against various chunk sizes from 1 to 124 (Ref. Section 6.1.3). The number was

[1]The parameters were picked from test runs. It was generally noted that growing around 300–400 trees gave slightly higher accuracy. There is additional information in the appendix explaining this rationale (cf. AppendixB).

chosen as the chunk size could not be higher than the data file with the lowest number of rows (The lowest file size was of 135 rows). We present the classifier accuracy from the 10-fold stratified cross validation on our data for each chunk.

We tested chunk size vs accuracy as well.

Classification of Individual Objects (Label Method 1)

The following result shows the ability to distinguish individual objects based on the given gesture. Thus, this takes into account all 22 gestures and the eleven objects associated with them separately. Figure 6.4 shows the overall accuracy of the classifier for all the objects against the chunk sizes. We can see that accuracy increases till around chunk size 10 then remains around 86%.

Accuracy for individual objects stays around 85% at chunk sizes greater than 10.

Figure 6.6 shows the cumulative confusion matrix of the 10-fold cross-validation for chunk size 20 (Accuracy M = 0.88, SD = 0.08)[2]. The diagonal elements show the percentage of correctly classified instances. The values in each horizontal row, apart from the diagonal cell show the miss-classified samples.

Accuracy mean for chunk size 20 is 88%

Classification of Objects Based on Similar Gestures (Label Method 2)

The following result shows the ability to distinguish only the gestures across the objects. For example, we look at all the objects which share the 15S gesture as a whole. Figure 6.5 shows the overall accuracy of the classifier for all the objects against the chunk sizes. We can see that accuracy increases till around chunk size 10 then remains around 86%.

Accuracy for gestures across objects stays around 86% after chunk size 10.

Figure 6.7 shows the cumulative confusion matrix of the 10-fold test for chunk size 20 (Accuracy M = 0.86, SD = 0.05)[3]2. The diagonal elements show the percentage of cor-

Accuracy mean for chunk size 20 is 86%.

[2]These values can be reproduced on our data set using the parameters given in section 6.1.3 "Method of Analysis" and random seed as 8.

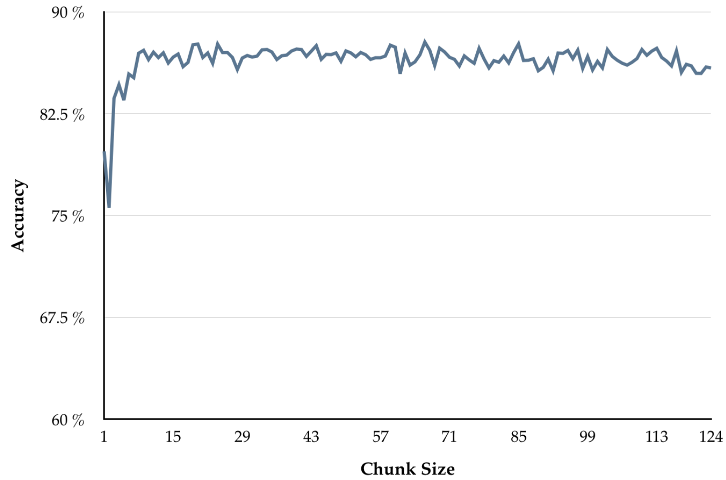

Figure 6.4: The classification accuracy for the identification of individual objects based on the gesture plotted against different chunk sizes used for pre-processing the data.

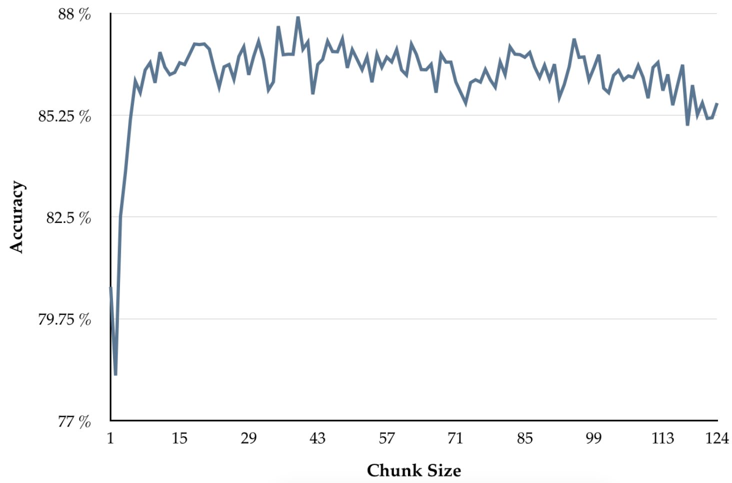

Figure 6.5: The classification accuracy for the identification of the gestures only plotted against different chunk sizes used for pre-processing the data.

	MBk2XB	HT2XF	L1'5B	TB2XS	P2XF	LN2'5BF	MB15S	TC2'5F	CH2XB	TB15S	HT2XT	PB1'PP	MB2XS	L2XB	PB2XS	P2XS	LN2'5T	TC2XF	MBk1'5B	TBk15F	CH15S	TBk2XRoll
MBk2XB	97.14%																			1.43%		1.43%
HT2XF		100.00%																				
L1'5B			85.92%						1.41%									1.41%	5.63%	1.41%	2.82%	1.41%
TB2XS			1.41%	83.10%				5.63%					8.45%							1.41%		
P2XF		1.85%			98.15%																	
LN2'5BF				6.67%		88.33%											3.33%			1.67%		
MB15S			4.17%				79.17%			15.28%											1.39%	
TC2'5F								98.57%														1.43%
CH2XB	1.39%							6.94%	77.78%					5.56%				8.33%				
TB15S							4.17%			88.89%		6.94%										
HT2XT						1.67%					81.67%		1.67%	10.00%	1.67%			3.33%				
PB1'PP						1.64%					1.64%	86.89%			3.28%		4.92%			1.64%		
MB2XS	1.33%			13.33%		1.33%		2.67%					81.33%									
L2XB								1.37%	2.74%					90.41%				4.11%				1.37%
PB2XS															93.33%	6.67%						
P2XS															3.45%	94.83%				1.72%		
LN2'5T						13.33%		1.67%					6.67%				78.33%					
TC2XF	2.82%							1.41%										95.77%				
MBk1'5B			1.35%							1.35%								6.76%	79.73%	8.11%	1.35%	1.35%
TBk15F	3.45%	1.72%										1.72%							1.72%	84.48%	5.17%	1.72%
CH15S			7.04%							1.41%									4.23%	7.04%	78.87%	1.41%
TBk2XRoll								8.62%					1.72%									89.66%

Figure 6.6: The cumulative confusion matrix for the classified individual objects based on the associated gestures across the 10-fold cross validation. The diagonal values show the percentage of correctly classified samples.

	1'PP	2XRoll	15S	2-5BF	2XB	2-5F	2XF	1'5B	15F	2XS	2'5T	2XT
1'PP	86.89%			1.64%					1.64%	4.92%	4.92%	
2XRoll		86.21%			3.45%	8.62%				1.72%		
15S	0.93%		90.70%		1.86%		0.47%	3.26%	2.79%			
2'5BF				76.67%			1.67%			16.67%	5.00%	
2XB		0.47%	0.47%		94.88%		2.79%		0.47%	0.93%		
2'5F		1.43%				98.57%						
2XF					10.27%		83.24%			6.49%		
1'5B			6.90%		4.14%		1.38%	83.45%	4.14%			
15F			8.62%		12.07%		5.17%	3.45%	70.69%			
2XS			0.76%		1.52%	0.38%	1.89%		0.38%	95.08%		
2'5T				8.33%		1.67%				15.00%	75.00%	
2XT					16.67%		10.00%			15.00%		58.33%

Figure 6.7: The cumulative confusion matrix for the classified gestures over the 10-fold cross-validation. The diagonal values show the percentage of correctly classified samples.

rectly classified instances. The values in each horizontal row apart from the diagonal cell show the miss-classified samples.

6.2.1 Observational and Feedback Analysis

All the users used their dominant hand for the single handed gestures and as the primary hand for gestures which needed one hand for support. For the gesture of folding the thin cable (2′5F), three of the twelve users preferred to rotate the cable over their dominant hand. Similar to observations from the study in Chapter 4 "Gesture Elicitation on Deformable Everyday Objects: A Qualitative Study", users would in general prefer to lift the object *off* the table. The plastic bowl however, was always kept on the table, apart from the particular gestures where the table was explicitly needed for support.

The dominant hand was used in all single handed gestures.

Feedback Analysis

The graphs in Figure 6.8 show the response for the question: "Would this be one of the first few gestures that you would attempt when presented with such an object?".

Group I: Users found it harder to grip and pick up the clothes hanger due to the gloves and the shape of the hanger itself. They also found the TC2XF on the thin cable to be an unfamiliar gesture. Two users mentioned they would prefer two handed gestures on both the bottles used in the study. The gloves also hindered the TC2′5F gesture, but eleven of the twelve users said it was a very familiar gesture. The 1′5B gesture on the book and the ladle was also not preferred by the users. However, they accepted it as a second gesture as they could not think of a better alternative. One user categorically mentioned that he would not prefer to deform books.

The glove seemed to have hindered certain gestures such as those with the hanger.

Group II: Users found it very inconvenient to perform the TBk15F gesture. Eight of the ten users used the second

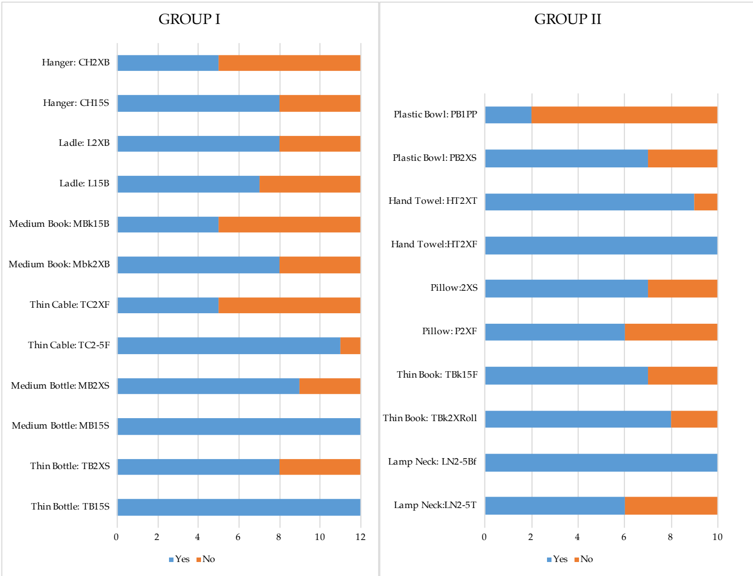

Figure 6.8: Users response for the question:"Would this be one of the first few gestures that you would attempt when presented with such an object?". In general, users agreed with the presented gesture set.

TBk15F was inconvenient for most users and the gloves slipped on the plastic surface.

hand as minor support for this gesture. This was said to be primarily due to the gloves slipping on the texture of the book. The gloves also hindered the gestures on the plastic bowl, especially the PB1PP. Four of the ten users did not want to deform the plastic bowl in any way. One user mentioned that the P2XS on the pillow felt different each time. This could be due to the small change in overall shape of the pillow with repeated use.

Study Experience Feedback

Users rated the study experience through a small feedback form.

The users rated their experience with the study on a 5-point Likert scale with 1 meaning *strongly disagree* and 5 meaning *strongly agree*. The forms used in the study are provided in the appendix. Figure 6.9 shows the mean ratings of the responses by both groups.

Q1: The homing part of the gesture (movement from the

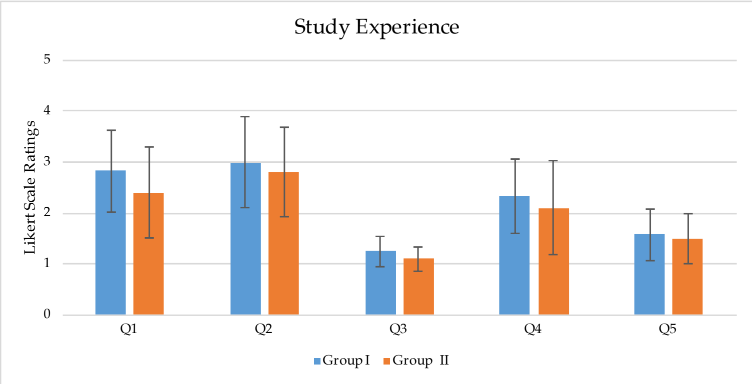

Figure 6.9: User responses for questions on study experience. The bars indicate 95% confidence intervals. The user experience between two groups yielded similar responses to all five questions.

start to the object and back) affected my gesturing.
Group I: $M = 2.83$, $SD = 1.27$; Group II: $M = 2.4$, $SD = 1.26$
Participants pointed out that certain objects such as the clothes hanger and the hand towel would not be lying on the table. The way one would *home* onto everyday objects might not be similar to the way, users had to do it in our controlled study environment.

Overall both groups had similar responses for the study experience questionnaire. The gloves hindered certain gestures due to slipping.

Q2: Wearing the glove affected my gestures.
Group I: $M = 3$, $SD = 1.41$; Group II: $M = 2.8$, $SD = 1.23$
Certain objects such as the thin book, plastic bowl, and the clothes hanger, were hard to grip according to the users. They also had to be careful with the markers while rolling the book or folding the thin cable over their hands.

Q3: The camera system distracted me while doing the gestures.
Group I: $M = 1.25$, $SD = 0.45$; Group II: $M = 1.1$, $SD = 0.32$
Overall users did not seem to have a problem with all the cameras around them.

Q4: Repeating the gestures 5 times was tiring.
Group I: $M = 2.33$, $SD = 1.15$; Group II: $M = 2.1$, $SD = 1.29$

Users mentioned that gestures on the bottles or the lamp neck were not tiring but they had harder time repeating TBk15F (Thin Book) and the PB1′PP (Plastic Bowl).

Q5: The tapping of the button by the feet affected my gesturing.
Group I: M = 1.58, SD = 0.79; Group II: M = 1.5, SD = 0.7
The trial task at the beginning was helpful for the users to understand the button. Though very rarely, one would forget to tap the button to start or stop the capture process.

6.3 Summary

The classifier results point towards the possibility of using deformation gestures as identifiers.

The results of the quantitative study show that the deformation gestures could be an indicator to identify and classify deformable everyday objects. We had used a 10-fold cross validation to verify our data. This meant that ten percent of our collected data was used as the test set. Figure 6.10 shows that the accuracy is also dependent on the amount of training data. The marker system we used for our gloves included multiple markers for each tracked location. This also adds a lot of redundant information to our data set. We discuss these limitations and suggest scope for future work in Chapter7 "Summary and Future Work".

Certain gesture-object combinations such as the PB1PP did not seem natural to the users.

The gesture set on the representative object set that we derived from the qualitative study was validated in general by the users. However there are certain gestures and object combinations that are not preferred by the users. Partly this was caused by the gloves which slipped on certain objects as well as the presence of markers on the gloves. The orientation of certain objects like the clothes hanger was not what users expect in their day-to-day life. Although the users agreed that certain gestures like the 1PP on the plastic bowl was a valid deformation gesture, they indicated that they could not understand the use of such a gesture.

These objects play a completely different role in daily life

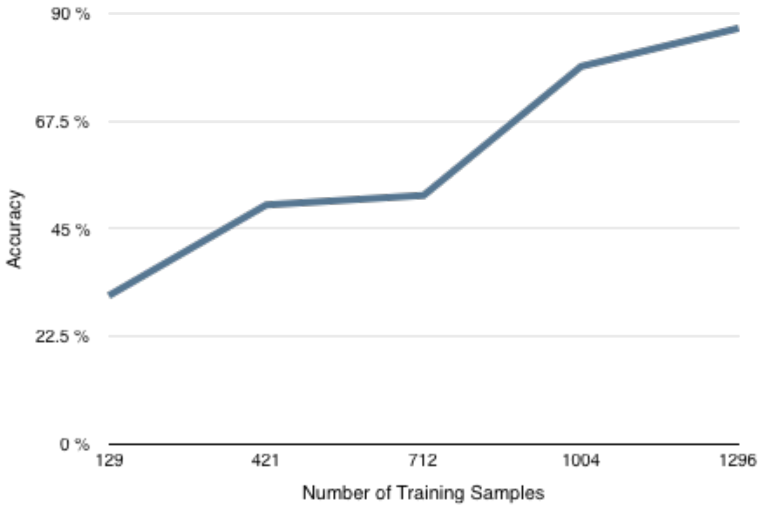

Figure 6.10: The graph shows the change in accuracy for classifying individual objects against number of training samples for 10-cross validation. We see that accuracy increases with increases in amount of training data.

and thus sometimes users did not readily map certain gestures to an object. Although identifying relevant tasks for deformation gestures is an important aspect of HCI research, it is beyond the current scope of this work. We shall attempt to address this concern as a part of our future work.

Task relevance may help users have context.

Chapter 7

Summary and Future Work

We now conclude with summarizing the thesis, stating the limitations of our studies, the dataset, and briefly discuss scope of future work.

7.1 Summary and Contributions

Inspired by the concept of re-purposing everyday objects as digital controls, we presented an exploratory study on how users exploit deformable affordances of such objects.

We presented an exploratory study.

To make an informed decision on how users could interact with deformable everyday objects, we compiled a list of everyday objects in a typical home environment. We conducted a house-to-house survey in twelve houses and had an informal interview to understand what objects are actually perceived as deformable objects. Observing these everyday objects along with how users briefly interacted with them helped us create object categories. These categories mainly depended on the generic shapes of the observed objects. Certain objects such as bottles and books that were found in comparatively larger quantities were given their own exclusive categories. The final categories consisted of

We conducted a house-house survey to compile a list of everyday objects.

bottles, books, cables, pillows (including portable sponge-like objects such as cushions), couches (including larger sponge-like objects such as mattresses), foldables (Thin objects that could be regularly folded like towels), boxes (including plastic containers that may be closer to hemisphere in shape but are hollow), bendables (objects that are usually bent but not to an extend that they can be considered foldable), and miscellaneous (including objects that may share similar shapes to bottles or books). The broad categorization of objects along with insight gained from the personal discussions with the users in a home environment ensured we could pick a representative set of objects for further studies.

We categorized the objects into shape and size based groups.

In order to understand the possible deformation affordances that are perceived by users, we conducted a gesture elicitation study. We created a simulated home environment resembling a part of a typical living room. We recorded the users actions though GoPro™cameras. Users were asked to perform deformation gestures on the selected subset of objects identified from the survey. They were encouraged to explore as much of the deformation affordance offered by each object as possible. It was seen that generating more than three to four gestures on an object was taxing for the user. They were then asked to rate the ease of deformation using each gesture on a five point Likert scale. The exploratory study recordings helped us identify how users used their hands to perform deformations on the objects. We used this information to come up with a generic coding scheme to analyze the gestures. Using this coding scheme we grouped similar gestures across users and objects. Thus, we arrived at a deformation gesture based classification of everyday objects used in the study. This also allowed us to identify objects that would elicit a common a set of gestures across users. It was observed that users would prefer to pick an object *off* the table as far as possible rather than use it *on* the table. There was no significant difference in the ease of deformation ratings given by users when they did gestures with the object *on* or *off* the table surface. In general, people preferred gestures which involved a complete hand grasp.

We conducted a gesture elicitation study on a representative subset of objects compiled from the house-house survey.

We created a coding scheme to group gestures based on how users used their hands and fingers.

We had observed in the previous study that our hands are expressive enough to convey information about the object under use. We used the Vicon motion capture system quantify hand based gestures as they deformed everyday objects (a subset of objects from the previous study). Then we trained a Random Forest based classifier with our data and used cross-validation to verify our assumptions. A 10-fold cross-validation on the gesture data showed 87% accuracy while trying to identify individual objects. The classification of object groups which share similar gestures also shows around 87% accuracy. We also employed this study to validate the gesture set compiled from the previous study. Users found the proposed gestures to be in line with their preferences overall. In certain cases where they did not agree with a gesture, it was noticed that the same gesture had a low frequency of occurrence in the qualitative study as well. The follow up feedback on the study also revealed that artificial constraints imposed by the motion capture system could also have affected certain gestures.

We quantified gesture data on deformable everyday objects via motion capture and used machine learning on the gesture data to classify the deformable objects.

Our findings point towards a broad canvas of possible deformable interactions on everyday objects. The identification of these deformable interactions would help designers to make informed decisions while building new interfaces. The results of the motion capture study show that the identification of deformable everyday objects is possible through tracking the deformation gesture without explicit knowledge of the objects themselves. The full recorded data is available and can used for future analysis. (cf Appendix B.2 "The Gesture Dataset").

The work revealed new interaction possibilities via deformation gestures on everyday objects.

7.2 Limitations

The number of everyday objects around us is quite enormous. Thus making it a daunting task to investigate all possible deformation gestures across these varied objects. Although the studies were designed and analyzed to encompass as many everyday objects and natural interactions with them as possible, we had certain limitations. The exploratory study had twelve users who were sitting on a couch during the course of the study. They interacted with all the presented objects placed on a coffee table in front of

The users sat on a couch during the exploratory study.

them. From our preliminary survey (cf. Chapter 3 "Identifying Deformable Everyday Objects: A Survey") we had observed that objects around the house are found in certain locations and orientations. For example, a clothes hanger is usually hanging on a rod while pillows maybe found lying on the bed. This may influence how the user approaches the object and the user's primary interaction with the object. Users were instructed to produce multiple gestures on the objects and thus may have generated a gesture that might not be readily natural at first glance. This thought was echoed in user responses in the second study (cf. Chapter 6 "Capturing Deformation Gestures on Everyday Objects: A Quantitative Study") during user feedback.

The location and orientation of an everyday object plays a role in how our hands approach them.

The motion capture gesture set included the whole sequence of the users hands homing on the object from a predetermined area on the table and back. So the actual deformation gesture phase as envisioned in the coding is only a part of the data set. In real life users could be approaching objects from different positions and orientations. We used markers only on the edges of the dorsal part of the palm, thumbs, index, and ring fingers. We also do not take into account the joints of the fingers that may provide additional discerning information across gestures. This was based on the complexity and varied differences in the gesture sets under study. The complex trajectories of different kinds of gestures meant, markers may not be tracked at all position due a fixed camera system. This led to missing and swapped markers which require extensive manual reconstruction. We feel this could be further explored by investigating only a narrow set of similar deformable objects (cf. Section 7.3 "Future Work").

We use markers only on the dorsal part of the palm, thumb, index, and ring finger.

Our analysis does not explicitly take into consideration the temporal nature of the gesture data. We compensate for the variation in gesture by diving all gesture sequences into equal parts and considering the median as a representative value. This may lead to a loss of lot of information that may be useful in discerning gestures on objects that differ slightly from each other. The concatenation of the time series rows to generate a single data point meant each marker information was split across multiple features. The multiple markers while compensating for redundancy also add

We use only median information in our machine learning feature set.

a significant amount of data to the gesture set. This also leads to a lot of noise in the large data set which affects performance analysis. We believe other methods for preprocessing the data to include only relevant marker data would help reduce analysis time and performance.

Data set may have lot of redundant features.

7.3 Future Work

As has been the theme through this entire thesis, we once again reiterate the vast scope of deformable everyday objects. This opens up multiple possible avenues that could be taken to have a deeper understanding of deformable everyday objects. We can begin with looking back at certain limitations of the work presented in this thesis (cf. Section 7.2 "Limitations").

We look at our limitations to expand on future work.

Everyday objects in a home environment are found in specific locations and orientations. We used the Vicon tracking system that confined the user to to a specific tracking area and the objects were always placed on the table. We envision a better designed setup that takes natural object location, and orientation into consideration would yield more concrete gesture-object mappings. We could also capture the natural approach movement to the object as one would do in their daily lives.

A study which takes into account their natural orientation may give better object-gesture mapping.

We only tracked certain fingers in our study based on our observation that in most gestures under study, the four fingers moved together. A complex marker system combined with the varied set of gestures would have added a lot of swapped and missing data due to the limitations of our camera setup. But, how well can we discern objects which share similar shapes and only vary slightly in dimensions? Our motion capture study had two different bottles which elicited similar gestures but they did differ in their shape (cf. Figure 4.6). A study which includes similar shaped bottles (e.g., cylindrical only) but with varying dimensions could be designed to collect gesture data. A focused study which only takes into account a single gesture type (e.g., 15S) may need a more extensive marker system. However, the camera setup would be easier to track markers which

A study focused on very similar object shapes but varying slightly in dimensions.

we know would not move in complex trajectories.

Using machine learning techniques like Dynamic Time Warping.

Machine learning tools like Dynamic Time Warping (Müller [2007]) could help understand the relationships between different gestures. One interesting area is locating common sub sequences between multiple gestures. This may show how the different gestures branch off from each other as represented in the coding scheme we employ. For example, how the finger and movement differs between a one handed bend and a squeeze on deformable objects. This knowledge could be beneficial in aiding the design of new deformation gesture based interfaces.

We could refer current literature in deformable interactions to design task relevance based studies.

Task relevance for deformation gestures on everyday objects was not in scope of our work. Nonetheless, it is a very crucial aspect to the study of deformable everyday objects. Participants in our motion capture study mentioned their difficulty to understand why one would do certain gestures. We shall draw from the existing body of work in deformable user interfaces and design experiments to study task relevance (Ahmaniemi et al. [2014], Troiano et al. [2014]). We would seek to answer the following research question:

- What are the preferred deformation gestures in everyday objects for a given task?

Mapping gestures to certain tasks would greatly aid in an improved understanding the how users would exploit perceived affordances that aid deformation in everyday objects.

Appendix A

Appendix for the Preliminary Survey

Appendix A contains the object lists collected from the survey. They are divided on the basis of locations. It is a semi-filtered list with some initial object groupings.

The following lists have been presented as figures.

1. Figure A.1 lists the objects found in the living room.
2. Figure A.2 lists the objects found in the in-house office.
3. Figure A.3 lists the objects found in the kitchen.
4. Figure A.4 lists the objects found in the bedroom.

A.1 Ease of Deformation and Volatility Ratings for Final Set

The comments of the participants was noted during the study and the grouping was done at the lab. The ratings given by the users helped us judge object groupings when we could not judge the pictures collected alone. Figure A.5

Object	Total		Object	Total
Bean Bag	1	L	Laptop Bag	1
Blinds	1	I	Magazine	1
Medium Book	1	V	Mat	1
Carpet	2	I	Mop Cloth	1
Couch	4	N	Orange	1
Curtain	3	G	Pillow	3
Cushion	5		Plastic Box	1
Thin Bottle	1	R	Rucksack	2
Juice Packaging	1	O	Sleeping Bag	1
Lamp	1	O	Thick Cable	5
Lamp Shade	2	M	Thin Blanket	2
Vaccum Hose	2		Thin Cable	3

Figure A.1: This table lists the count for objects found in the living room.

shows the user ratings for *Ease of Deformation* and *volatility* for the final set of objects. It should be noted, not all objects were explicitly named so some ratings may include objects grouped together. For example, thick bottle may include cola ratings as well other thick bottles.

Object	Total	IN HOUSE OFFICE	Object	Total
Bag	2		Pencil Bag	1
Basket	1		Phone Charger	2
Cable Organiser	1		Plant	1
Cardboard	1		Plastic Bag	1
Cardboard Box	2		Plastic Container Lid	1
Chair	1		Plastic Dustbin	1
Charger Joint	1		Plastic Folder	2
Couch	1		Plastic Package	1
Credit Card	2		Plastic Pen Case	1
Curtain	4		Plastic Plant Tub	1
Cushion	6		Pocket Tissue Pack	3
Exercise Band	1		Rubber Duck	1
Flag	1		Rucksack	6
Frisbee	2		Small Football	1
HDD Cover	1		Soft Toy	3
Key Ring	1		Tablet Packaging	1
Lamp Neck	2		Tape Roll	1
Lamp Shade	2		TeaBag Rope	1
LAN Cable	2		Thick Book	9
Laptop Cable	1		Thick Bottle	3
Laptop Sleeve	1		Thick Cable	6
Mat	1		Thick Folder	1
Medium Book	6		Thin Book	10
Mouse Pad	2		Thin Cable	8
Paper	10		Thin Folder	2
Paper Organiser	1		Tin Can	2
Paper Package	2		Tin Cup	1
Paper Puncher Backs	1		Wallet	3
Thin Bottle	4		Watch Strap	2

Figure A.2: This table lists the count for objects found in the in-house office.

Object	Total		Object	Total
Aluminium Foil	1		Plastic Box	5
Basket	1		Plastic Cover	1
Bottle Lids	1		Plastic Cup	2
Boxes	1		Plastic Scraper	1
Brush	1		Plastic Spoon	2
Cable	1		Plastic Ladle	7
Cardboard Box	9	K	Salt Container	1
Cardboard Box Lid	1		Sauce Bottle	3
Carton	1	I	Shelf steel	1
Cleaning Cloth	6	T	Sponge	12
Cloth Bag	1		Steel Wool	1
Cloth Towel	1	C	Tea Strainer	2
Curtain	1		Tea/ Coffee Tin	2
Dishwasher Bottle	10	H	Thick Book	1
Egg Carton	1	E	Thick Bottle	4
Fly Swatter	1		Thick cable	3
Knife	2	N	Thin Bottle	9
Lamp Neck	1		Thin Plastic Box	2
Oven Mit	3		Tin Can	1
Paper	1		Tissue Roll	7
Paper Bag	3		Tube	1
Plastic Bag	6		TupperWare Bag	1
Plastic Beaker	1		Water Bottle	2
Window Décor	1		Whisk	2

Figure A.3: This table lists the count for objects found in the kitchen.

Object	Total	BEDROOM	Object	Total
Yoga Mat	1		Moisturiser Bottle	5
Water Bottle	5		Mirror	1
Velvet	1		Mattress	11
Vaccum Hose	1		Mat	1
Tube	3		Lens Solution Bottle	1
Tissue Box	2		Laundry Bag	2
Tissue	1		Laptop Charger	1
Tin Lid	1		Lamp Shade	1
Tin Can	2		Lamp Neck	4
Tin Box	1		Jacket	4
Thick Blanket	11		Hanger	4
Thin Blanket	2		Hairband	1
Thin Cable	6		Gloves	1
Thick Cable	5		Eyedrop Dispenser	1
Thick Bottle	1		Detergent Bottle	1
Suitcase	5		Deoderant Bottle	1
Specs Box	1		Cushion	6
Soft Toy	3		Curtain	2
Soft Book	1		Cream Tube	1
Slippers	2		Carpet	1
Shoes	1		Cardboard Box Lid	1
Rucksack	4		Cardboard Box	3
Pocket Tissue Pack	3		Cable	1
Plastic Plant Tub	1		Blinds	1
Plastic Box	2		Blanket	1
Plant	1		Big Cardboard Box	1
Pillow	13		Belt	1
Paper Packaging	2		Bedsheet	1
Paper	1		Basket Lid	1
Necklace	1		Balloon	1

Figure A.4: This table lists the count for objects found in the bedroom.

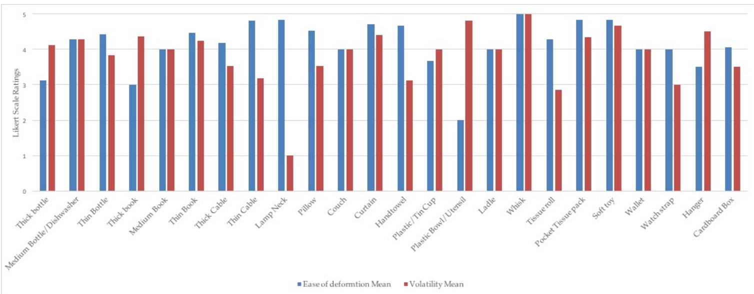

Figure A.5: *Ease of Deformation* and *volatility* ratings for the final object list chosen for qualitative study. Objects like the lamp neck which maintains its deformation have very low volatility.

Appendix B

Appendix for the Motion Capture Study

Appendix B contains additional material for the motion capture study. It includes information about the data set. text description of the predefined gesture labels used in the analysis, some graphs from the test runs for determining the Random Forest parameters, and the feed back forms handed out to the users.

B.1 Random Forest Parameters

While the random forest allows one to tune multiple parameters from tree size to impurity functions, we mainly looked at these basic parameters:

1. n_estimators
2. max_features

One thing we looked at is out of bag error rate (oob). Breiman [2001] states that this metric could be used as a measure of the classifier accuracy itself. We checked multiple *max_features* values against *n_estimators*. We present

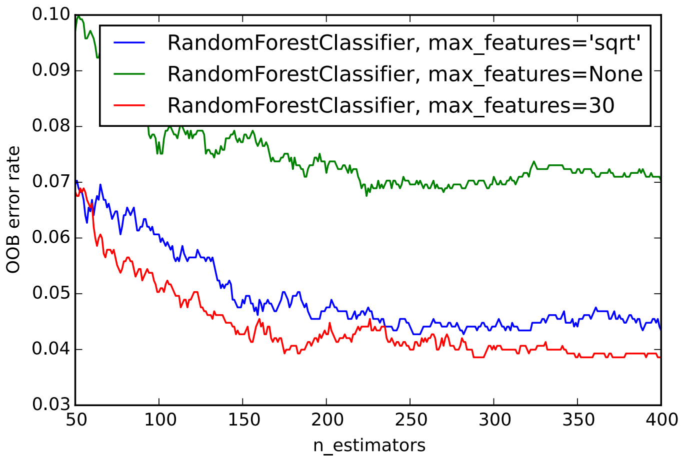

Figure B.1: The graph lists out of bag error with respect to number of estimators for different max_feature values on the motion capture data. (The lower the better)

one of the graphs on our data adapted from the sci-kit user-guide. We notice the graph tapers around 300 for *n_estimators* and *max_features* shows a better score with value 30.

B.2 The Gesture Dataset

The gesture set collected during the motion capture study will be provided as an attachment. It includes 1451 files in total saved as comma separated values (csv). Each file corresponds to one gesture trial.

The file names are coded as:
$UserXX\_GestureCodeYY.csv$: where XX refers to the unique user ID and YY is the trial number.

USER04_LN2-5BF02.CSV:
This name indicates the gesture represented in the file is the bending forward of the lamp neck using one hand as support. This was the second trial for the gesture and came from user ID 4

Example: *User04_LN2-5Bf02.csv*

Note: Gesture codes have - instead of ′ in the file names.

The data in the files is structured as:
Frame,Sub Frame, [Label ID][Position co-ordinates],[Label ID][Velocity],[Label ID][Acceleration]

1. The *Frame* column can be discarded.
2. *Sub Frame* refers to the actual frame number. (The data was collected at 100 frames per second)
3. Figure B.2 shows the markers and their label IDs.
4. Position, velocity and acceleration have values from X, Y and Z dimensions each.
5. Any gaps in the data have been filled with zeroes.

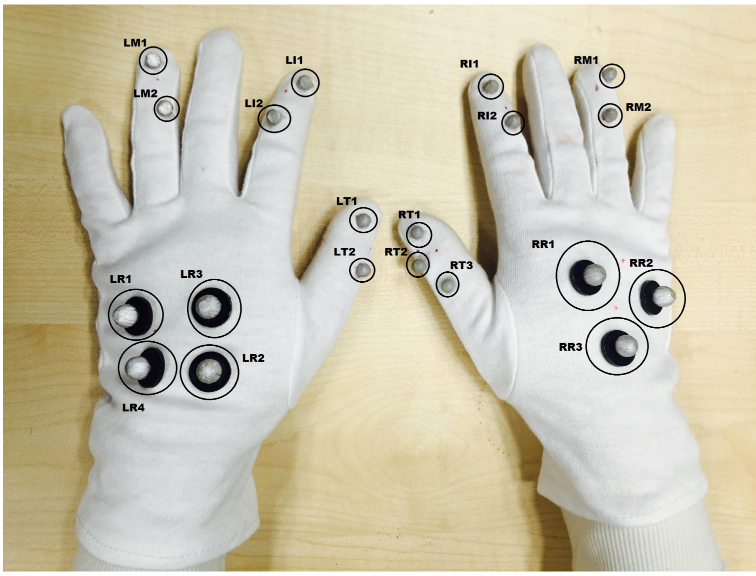

Figure B.2: The glove used for the motion capture study annotated with the label ids for the markers.

GESTURE DESCRIPTIONS

Thin Bottle: TB15S
A one handed squeeze

Thin Bottle: TB2XS
A 2 handed squeeze with one hand under the other vertically

Medium Bottle: MB15S
A one handed squeeze

Medium Bottle: MB2XS
A 2 handed squeeze with one hand under the other vertically

Thin Cable: TC2'5F
Hold an end of cable in one hand and roll the cable in loops (create a roll)

Thin Cable: TC2XF
Use both hands to hold a bunched cable and fold it

Lamp Neck: LN2'5T
Hold base with one hand and twist the neck around

Lamp Neck: LN2'5Bf
Hold base with one hand and bend neck forward

Medium Book: MBk2XB
Hold book on both sides with two hands and bend (one hand on each side, book held vertically)

Medium Book: MBk1'5B
Use the table as a support and bend the book with one hand

Thin Book: TBk2XRoll
Roll the book using both hands

Figure B.3: This lists gesture descriptions along with assigned labels for the analysis of motion capture data. - 1

GESTURE DESCRIPTIONS

Thin Book: TBk15F
Use one hand to fold the book (in half)

Pillow: P2XF
Use two hands and old pillow in two (hold and fold from both sides)

Pillow: 2XS
Use 2 hands and squeeze the pillow (from the top)

Hand Towel: HT2XF
Use two hands fold the towel (Pinch two edges and fold in half)

Hand Towel: HT2XT
Use two hands roll and twist the towel

Plastic Bowl: PB2XS
Use two ends to squeeze the bowl (hands on opposite sides)

Plastic Bowl: PB1'PP
Lay bowl on its side on the table and press down with palm

Ladle: L15B
Use the table as support and bend the ladle with one hand

Ladle: L2XB
Use both hands and bend the ladle (hands on two ends)

Hanger: CH15S
Do a one hand squeeze (hold it in the middle)

Hanger: CH2XB
Use both hands to bend (hands on two ends)

Figure B.4: This lists gesture descriptions along with assigned labels for the analysis of motion capture data. - 2

QUESTIONNAIRE

1. The homing part of the gesture (movement from the start to the object and back) affected my gesturing

   ☐ *Strongly Disagree* ☐ *Disagree* ☐ *Neither /Nor* ☐ *Agree* ☐ *Strongly Agree*

2. Wearing the glove affected my gestures.

   ☐ *Strongly Disagree* ☐ *Disagree* ☐ *Neither /Nor* ☐ *Agree* ☐ *Strongly Agree*

3. The camera system distracted me while doing the gestures.

   ☐ *Strongly Disagree* ☐ *Disagree* ☐ *Neither /Nor* ☐ *Agree* ☐ *Strongly Agree*

4. Repeating the gestures 5 times was tiring.

   ☐ *Strongly Disagree* ☐ *Disagree* ☐ *Neither /Nor* ☐ *Agree* ☐ *Strongly Agree*

5. The tapping of the button by the feet affected my gesturing.

   ☐ *Strongly Disagree* ☐ *Disagree* ☐ *Neither /Nor* ☐ *Agree* ☐ *Strongly Agree*

6. Please place a cross on the objects from the study that you would not come across in your home surroundings, **if any**. (If there are similar objects do name them or give a short description)

Figure B.5: This was the feedback form to rate the study experience for the motion capture study. This was handed out to the users in Group I

QUESTIONNAIRE

1. The homing part of the gesture (movement from the start to the object and back) affected my gesturing

☐ *Strongly Disagree* ☐*Disagree* ☐*Neither /Nor* ☐*Agree* ☐*Strongly Agree*

2. Wearing the glove affected my gestures.

☐ *Strongly Disagree* ☐*Disagree* ☐*Neither /Nor* ☐*Agree* ☐*Strongly Agree*

3. The camera system distracted me while doing the gestures.

☐ *Strongly Disagree* ☐*Disagree* ☐*Neither /Nor* ☐*Agree* ☐*Strongly Agree*

4. Repeating the gestures 5 times was tiring.

☐ *Strongly Disagree* ☐*Disagree* ☐*Neither /Nor* ☐*Agree* ☐*Strongly Agree*

5. The tapping of the button by the feet affected my gesturing.

☐ *Strongly Disagree* ☐*Disagree* ☐*Neither /Nor* ☐*Agree* ☐*Strongly Agree*

6. Please place a cross on the objects from the study that you would not come across in your home surroundings, **if any**. (If there are similar objects do name them or give a short description)

Figure B.6: This was the feedback form to rate the study experience for the motion capture study. This was handed out to the users in Group II

Bibliography

Teemu T. Ahmaniemi, Johan Kildal, and Merja Haveri. What is a Device Bend Gesture Really Good For? In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 3503–3512. ACM, 2014.

Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

Kai-Yin . Y. Cheng, Rong-Hao . H. Liang, Bing-Yu . Y. Chen, Rung-Huei . H. Laing, and Sy-Yen . Y. Kuo. iCon: Utilizing Everyday Objects as Additional, Auxiliary and Instant Tabletop Controllers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1155–1164. ACM, 2010.

Keywon Chung, Michael Shilman, Chris Merrill, and Hiroshi Ishii. OnObject: Gestural Play with Tagged Everyday Objects. In *Adjunct proceedings of the 23nd annual ACM symposium on User interface software and technology*, pages 379–380. ACM, 2010.

Christian Corsten. Co-Optjects: Instant User Interfaces Through Everyday Objects. Master's thesis, RWTH Aachen University, Aachen, January 2012.

Christian Corsten, Ignacio Avellino, Max Möllers, and Jan Borchers. Instant User Interfaces: Repurposing Everyday Objects As Input Devices. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '13, pages 71–80, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-2271-3. doi: 10.1145/2512349.2512799. URL `http://doi.acm.org/10.1145/2512349.2512799`.

David T. Gallant, Andrew G. Seniuk, and Roel Vertegaal. Towards more Paper-like Input: Flexible Input Devices for Foldable Interaction Styles. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 283–286. ACM, 2008.

Johan Kildal and Graham Wilson. Feeling It: The Roles of Stiffness, Deformation Range and Feedback in the Control of Deformable UI. In *Proceedings of the 14th ACM international conference on Multimodal interaction*, pages 393–400. ACM, 2012.

Johan Kildal, Susanna Paasovaara, and Viljakaisa Aaltonen. Kinetic Device: Designing Interactions with a Deformable Mobile Interface. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 1871–1876. ACM, 2012.

Ron Kohavi. A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'95, pages 1137–1143, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc. ISBN 1-55860-363-8. URL `http://dl.acm.org/citation.cfm?id=1643031.1643047`.

Sang-su . S. Lee, Youn-kyung . K. Lim, and Kun-Pyo . P. Lee. Exploring the Effects of Size on Deformable User Interfaces. In *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile Devices and Services Companion*, MobileHCI '12, pages 89–94, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1443-5. doi: 10.1145/2371664.2371682. URL `http://doi.acm.org/10.1145/2371664.2371682`.

C.L. MacKenzie and T. Iberall. *The Grasping Hand*. Advances in Psychology. Elsevier Science, 1994. ISBN 9780080867588. URL `https://books.google.de/books?id=V9G5Yd46VlEC`.

Meinard Müller. *Information Retrieval for Music and Motion*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007. ISBN 3540740473.

Cuong Nguyen, Yong Wang, and Ha Nam Nguyen. Random Forest Classifier Combined with Feature Selection for Breast Cancer Diagnosis and Prognostic. *Journal of Biomedical Science and Engineering*, Vol.06No.05: 10, 2013. doi: 10.4236/jbise.2013.65070. URL http://www.scirp.org/journal/PaperInformation.aspx?PaperID=31887%%5D%2031887.

D. A. Norman. *The Design of Everyday Things: Revised and Expanded Edition*. Basic Books, 2013. ISBN 9780465072996. URL https://books.google.de/books?id=nVQPAAAAQBAJ.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Henning Pohl and Michael Rohs. Around-device Devices: My Coffee Mug is a Volume Dial. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*, pages 81–90. ACM, 2014.

Kathrin Probst, David Lindlbauer, Michael Haller, Bernhard Schwartz, and Andreas Schrempf. A Chair as Ubiquitous Input Device: Exploring Semaphoric Chair Gestures for Focused and Peripheral Interaction. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 4097–4106. ACM, 2014.

Carsten Schwesig, Ivan Poupyrev, and Eijiro Mori. Gummi: A Bendable Computer. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 263–270, New York, NY, USA, 2004. ACM. ISBN 1-58113-702-8. doi: 10.1145/985692.985726. URL http://doi.acm.org/10.1145/985692.985726.

Stat.berkeley.edu. Random Forests - Classification Description, 2016. URL http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm.

Giovanni Maria Troiano, Esben Warming Pedersen, and Kasper Hornbaek. User-defined Gestures for Elastic, Deformable Displays. *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*, pages 1–8, 2014.

Radu-Daniel . D. Vatavu. User-defined Gestures for Freehand TV Control. In *Proceedings of the 10th European conference on Interactive tv and video*, pages 45–48. ACM, 2012.

Radu-Daniel . D. Vatavu and Ionuct Alexandru Zaicti. Automatic Recognition of Object Size and Shape via User-dependent Measurements of the Grasping Hand. *International Journal of Human-Computer Studies*, 71(5):590–607, 2013.

Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. User-defined Gestures for Surface Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1083–1092. ACM, 2009.

Index

Typeset March 17, 2016