Extracting Hand Grasp & Motion for Intent Expression in Mid-Air Shape Deformation : A Concrete & Iterative Exploration through a Virtual Pottery Application

Vinayak[a], Karthik Ramani[a,b]

[a]School of Mechanical Engineering
Purdue University
[b]School of Electrical and Computer Engineering
Purdue University (by courtesy)

Abstract

We describe the iterative design and evaluation of a geometric interaction technique for bare-hand mid-air virtual pottery. We model the shaping of a pot as a gradual and progressive convergence of the pot-profile to the shape of the user's hand represented as a point-cloud (PCL). Our pottery-inspired application served as a platform for systematically revealing how users use their hands to express the intent of deformation during a pot shaping process. Our approach involved three stages: (a) clutching by proximal-attraction, (b) shaping by proximal-attraction, and (c) shaping by grasp+motion. The design and implementation of each stage was informed by user evaluations of the previous stage. Our work evidently demonstrates that it is possible to enable users to express their intent for shape deformation without the need for a fixed set of gestures for clutching and deforming a shape. We found that the expressive capability of hand articulation can be effectively harnessed for controllable shaping by organizing the deformation process in broad classes of intended operations such as pulling, pushing, and fairing. After minimal practice with the pottery application, users could *figure out* their own strategy for reaching, grasping, and deforming the pot. Users particularly enjoyed using day-to-day physical objects as tools for shaping pots.

Keywords: Mid-air gestures, depth sensor, virtual pottery, shape deformation, hand grasp.

1. Introduction

Mid-air gestures have been widely used as the *symbolic* means for expressing user's intent in 3D shape modeling [1, 2, 3, 4, 5, 6]. Gesture-based interactions enable the user to focus on the design task rather than dedicating significant time towards learning the usage of the tool itself [7]. With the recent commercialization of depth cameras, gesture-based interactions have become accessible to the common user; creative applications for free-form shape modeling [8] in mid-air have gained significant popularity. The user input in these applications is represented as a combination of some special hand posture (such as pointing with a finger), and the motion of a representative point (such as the palm or finger-tip) on the hand.

Hand and finger movements in real-world shaping processes (such as pottery or clay sculpting) are complex, iterative, and gradual. Such processes are essentially governed by the physics and geometry of contact between the hand and clay. Thus, the true expressive potential of finger movements remains underutilized despite advances in hand pose and skeletal estimation [9, 10]. This is what drives our research wherein, our intention is to bridge the gap between the user's expression of intent and the corresponding deformation of a virtual shape.

In this paper, we give an comprehensive account of our recent works [11, 12] by describing the iterative design and evaluation of a geometric interaction technique for bare-hand mid-air virtual pottery. Our broader goals are to (a) identify aspects of real-world interactions that can be emulated in free-form 3D shape deformation, (b) understand the expression of design intent in shape deformation in terms of the user's hand grasp and motion, and (c) design an interaction that integrates the geometric information in user's actions with shaping operations in virtual space.

1.1. Contributions

This paper is an extension of our recent work [12], where we modeled the shaping of a pot as a gradual and progressive convergence of the pot's profile to the shape of the user's hand represented as a point-cloud (PCL). We presented a method that uses the kernel-density estimate (KDE) of the hand's PCL to extract the grasp and motion for deforming the shape of a pot in 3D space. This feature of our method directly allows a user to shape pots by using physical artifacts as tools without the need for computing any finite set of gestures or hand skeleton. In doing so, we demonstrate that it is possible to achieve controllability in bare-hand mid-air shape deformation using raw PCL data of the user's hand.

There are two differences between this paper and our prior works [11, 12]. First, we present the complete evolution of our algorithm in three stages of iterative design (section 3.3). At the end of each stage, we describe a user evaluation that informs the algorithm development of the subsequent stage. Second, we evaluate our KDE based approach in comparison to our prior work [11]. Our evaluations help reveal two core aspects of mid-air interactions for shape deformation, namely, intent & controllability. We characterize user behavior in pottery design in terms of (a) common hand & finger movement patterns for creating common geometric features, (b) user perception of intent, and (c) engagement, utility, and ease of learning provided by our approach.

2. Related Work

2.1. Mid-air Gestures

Gestures can be designed effectively for pointing, selection [13, 14], and navigation, since they define an unambiguous mapping between actions and response. Such tasks are implemented using deictic gestures [15] and can usually be segmented into discrete phases, with each phase triggering an *event* or a *command* [16]. Pointing in the direction of a virtual object creates the association between the user and the object. A recent study [17] shows dwell-time to be an effective method of pointing and selecting objects without hint to the users. In manipulative tasks such as ours, a direct spatial mapping is required between the user's input and the virtual object [18, 15]. Particularly in our case, such an association would be in terms of the proximity of the user's virtual hand to the shape being deformed.

2.2. Gestures for 3D Modeling

Let us consider a mid-air interaction scenario of selecting and displacing a mesh vertex for deforming a 3D mesh. Since the user's hands are interacting in the air, there is no physical or natural mechanism for triggering events. Here, gestures could serve two fundamental purposes. First, they help define a beginning (e.g. reaching and clutching some region of interest) and end (e.g. de-clutching the region after required deformation) of an interaction [16, 19]. Secondly, they help define the exact operation from a set of operations defined in the context of the application. For example, the type of deformation could be selected by using different gestures (e.g. fist to pull, point to push, open palm to flatten).

On these lines, most existing bare-hand interaction techniques for 3D shape conceptualization, use gestures combined with arm and full-body motions. Segen and Kumar [1] showed examples of computer-aided design (CAD) with their *Gesture VR* system, using computer vision for general virtual reality (VR) applications. Wang et al. [2] presented *6D Hands* to demonstrate CAD using marker-less hand tracking. The modeling of sweep surfaces using hand gestures and body motion was demonstrated by Vinayak et al. [4, 5]. Han and Han [3] demonstrated an interesting surface-based approach with particular focus on audiovisual interfaces for creating 3D sound sculptures. Holz and Wilson proposed *Data miming* [7] as an approach towards descriptive shape modeling wherein voxel representation of a user's hand motion is used to deduce the shape which the user is describing. This approach uses hands without the explicit determination of gestures for recognizing the user's description of an existing shape.

2.3. Hand Grasp

Prehension is a common phenomenon in real-world interactions. Jeannerod [20] notes two functional requirements of finger grip during the action of grasping, (a) adaptation of the grip to the size, shape, and use of the object to be grasped and (b) the coordination between the relative timing of the finger movements with hand transportation (i.e. whole hand movements). Intended actions strongly influence motion planning of hand and finger movements [21]. This suggests that the intent for deformation can be recognized before the user makes contact with the surface being deformed. Grasp classification [22] and patterns of usage and frequency [23] have been integral to robotics research. Literature in virtual reality [24, 25] has studied and implemented grasping in the context of object manipulation (pick-and-place). Kry et al. [26] implemented a novel hardware system to emulate grasping for desktop VR applications such as digital sculpting. It is worth noting that the primary methodology for investigating grasp taxonomies is mostly derived from the geometry of the hand in relation to a physical object that is held or manipulated by the hand. What we aim to do is to understand what is the minimal and sufficient characterization of the user's hand and finger movements, that could be used for mid-air deformation. Our goal is not to explicitly detect the hand grasp, but to design a deformation approach where the grasp is automatically and implicitly taken into consideration.

3. Overview

3.1. Intent & Controllability

The general term *intent* is literally defined as "*the thing that you plan to do or achieve : an aim or purpose*". In our case, intent (*what one wants to achieve*) can be described in terms of the context of shape deformation (*what operations one can perform on the shape*). Based on Leyton's perceptual theory of shapes [27], Delamé et al. [28] proposed a process grammar for deformation by introducing structuring and posturing operators. Here, structuring operators involve adding/removing material to the shape, while posturing operators allow for modifications such as bending or twisting some portion of the shape. Since our context is that of deformation, we define the intent in terms of two basic operations: pulling and pushing. These are analogous to structuring operators.

We see controllability as the quality of intent recognition and disambiguation *as perceived by the user*. Specifically, in our context, controllability is defined as a function of two factors: (a) the disparity between what a user intends for the shape to be and what the shape actually becomes after the deformation and (b) the responsiveness of the deformation. The goal is to minimize the disparity and optimize the responsiveness.

3.2. Rationale for Pottery

We have two goals in this paper. First, we seek a concrete geometric method that takes a general representation of the user's hand (PCL) and allows the user to deform 3D geometry. Second, we want to investigate this geometric method in light of intent and controllability. Thus, our focus here is not to build a comprehensive and feature-rich 3D modeling system. Instead, we intend to investigate spatial interactions for 3D shape deformation with an unprocessed representation of the hand.

In a general shape deformation scenario, an arbitrary triangle mesh is the ideal and generic shape representation. However, a controlled study is prohibitively challenging in such a case, for two reasons. First, the hand PCL data obtained from a single depth sensor is partial and noisy. Second, dynamic and complex finger motions add further complexity to the occlusions and noise. Subsequently, designing interaction tasks for a quantitative evaluation is difficult, particularly for users that have no prior experience with mid-air interactions for free-form 3D modeling. Hence, it is essential to constrain the geometric representation of the object being modified.

Our broader motivation in this work is to cater to the creative needs of individuals that are inclined towards 3D modeling and design and but do not have the expertise require for working with design tools. With this in view, we use pottery as our application context for two reasons. First, it offers a well-defined and intuitive relationship between the use of hands and the shaping of pots to a user. This allows us to concretely construct a geometric relationship between the shape of the hand PCL and the corresponding user intent. Secondly, the simplicity of the geometric representation and deformation lends itself to quantitative measurement of the user's response to our system.

3.3. Approach

Given the context of pottery, our approach involved the following three stages:

Stage 1: Using hand as one-point manipulator, we implemented *proximal-attraction*, an interaction technique for clutching and de-clutching without hand gestures. Our technique (section 4) generalizes the notion of *dwell-time* in the context of mid-air shape deformation. We conducted a preliminary study to evaluate the feasibility and effectiveness of this technique.

Stage 2: We extended the *proximal-attraction* method to the whole shape of the hand (section 5) [11]. Here,the hand was represented as a collection of multiple points (i.e PCL) obtained via a depth sensor. Each point in the PCL deformed a small local region on the pot using the proximal-attraction approach. On the whole this amounted to a gradual and progressive convergence of the pot-profile to the shape of the user's hands. Through experimentation, we found that users had significant difficulty in creating convex (pulling) and flat (fairing) features on the pot. This method was also found to be agnostic to the user's grasp and hand movements.

Stage 3: Based on our experiments, we implemented our final technique for pot deformation using hand PCL (section 6). We used kernel-density estimation to characterize the contact between the hand and the pot. This allowed us to classify the users' intent to push, pull or fair the surface of the pot depending on the hand grasp, finger movements, and motion of the hand on the pot's surface. We conducted a final user evaluation to investigate the efficacy of this approach.

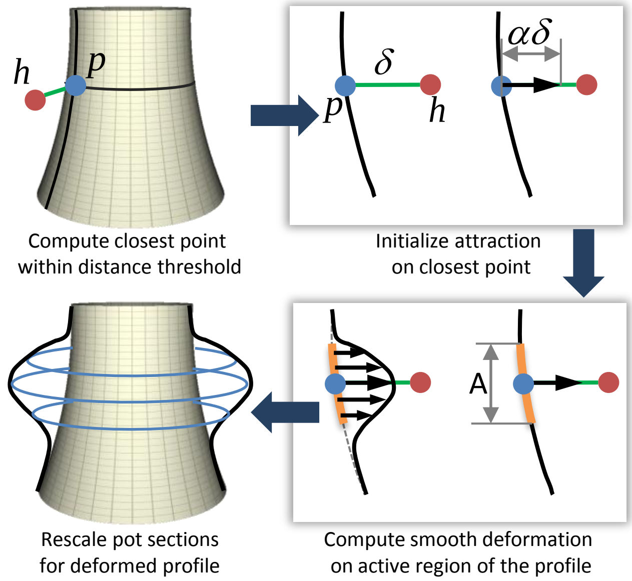

Figure 1: Algorithm for one-point pot deformation is illustrated for proximal-attraction. The pot is gradually deformed by attracting the profile towards the hand (represented by a point). Subsequently, each section is re-scaled to obtain the deformed pot surface.

3.4. Pot Representation & Deformation

The deformation algorithm for the pot evolved through iterative implementation and evaluation. Here we describe the basic geometric representation of a pot and the general computational setup of deforming the pot.

We represent a pot as a simple homogeneous generalized cylinder. The surface of the pot is defined as a vertical stack of circular sections. Each section is a polygonal approximation of a circle, i.e. a closed regular polygon. Note that a sequenced list of pairs (radius, height) is the *profile curve* of the pot. The deformation of a pot is achieved by deforming the profile curve, i.e by modifying the radii of each section. For a 3D pot, this essentially corresponds re-scaling each section by the corresponding amount of deformation.

4. Hand as a Point: Clutching by Proximal Attraction

In the first stage, we developed a method wherein the hand is represented as a single point manipulator, as is the case with many gesture-based methods. The main goal was to allow users to deform the surface the of pot without using hand gestures for clutching and de-clutching the pot.

4.1. Technique

Let h be the location of the hand in 3D space and p be the point on the pot that is closest to h. The main idea of proximal-attraction is to deform the pot gradually by attracting p towards

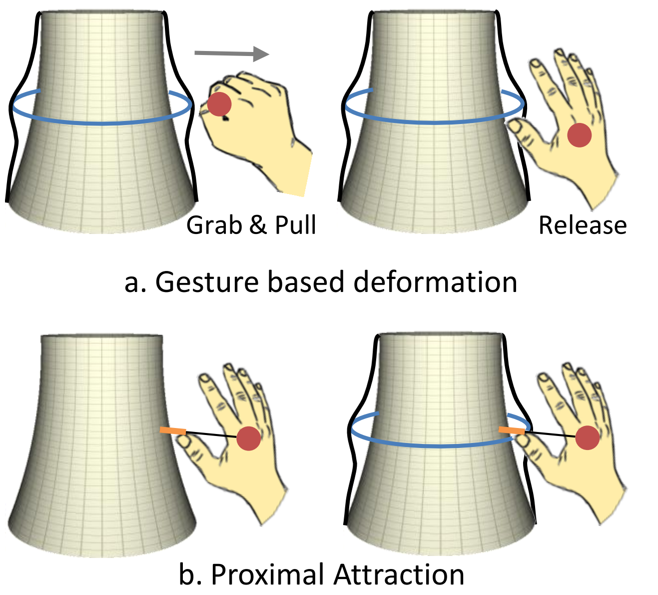

Figure 2: Two strategies are shown for clutching and deforming a pot using hand as a single point. In the first approach (a) grab and release gestures. The second (b) is the proximal-attraction approach

h in the *horizontal* plane. The condition of proximity is that the distance between h and p should be less than a pre-defined threshold (say ε). We implement the approach in the following steps:

1. Given h and A, compute p
2. if($||h - p|| < \varepsilon$)
   (a) Set δ to horizontal distance between h and p
   (b) Set attraction at p to $\alpha\delta$
   (c) Compute smooth deformed profile using Laplacian smoothing ($\nabla^2\delta = 0$ for all points in A)
3. Rescale pot sections

Here, $\alpha \in [0, 1]$ is the rate of attraction where $\alpha = 0$ implies no attraction and $\alpha = 1$ implies maximum attraction. Our idea is inspired by exponential smoothing [29]. The main step was to determine the right balance between the rate of attraction and the distance threshold. The responsiveness of deformation is directly proportional to both, attraction rates and distance threshold. From our pilot studies, we found $\alpha = 0.3$ and $\varepsilon = 0.05$ to be the optimal values. Here, the distances are in the normalized device coordinates. In our current implementation, we pre-defined the active region A to be 50% of the total profile length.

4.2. Preliminary Evaluation

Our main goal was to examine the feasibility and effectiveness of the proximal-attraction approach for pot shaping in terms of user performance and behavior. We also wanted to determine the differences between our method and a typical gesture-based approach. Additionally, we wanted to understand the reception of a creative application such as pottery for a wide variety of participants - particularly those without prior knowledge of CAD tools. For this, we conducted a two-day field study [1] in an exhibition setting.

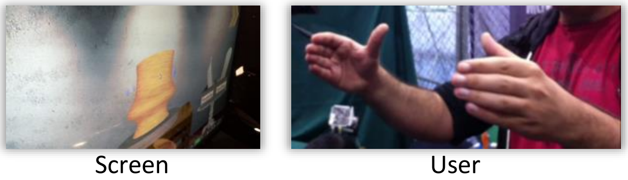

Figure 3: An example of common behavior is shown wherein users shaped their hands to express their intent for deformation.

Apparatus. Our hardware setup consisted of a ThinkPad T530 laptop, a 60" display, and the Microsoft Kinect camera. The Kinect camera was placed on a tripod below the display facing a user standing at a distance of around 1.5 – 2.0 meters from the display. Our pottery prototype was developed in C++ and openGL.

Implementation. We implemented two versions of our pottery application, one using mid-air gestures and the other based on the proximal-attraction approach. We first obtained the position of the hand using the skeletal tracking algorithm provided by the *openNI* API. Owing to the nature of the venue, the study was not conducted in a controlled environment leading to disturbances in skeletal tracking, posture recognition, and ambient noise. Thus, appropriate measures were taken to isolate the user from the audience.

The gesture-based prototype uses two simple hand postures, *grab* and *release*, which correspond to closed and open palms respectively (Figure 2(a)). We used the random forest algorithm for posture recognition as detailed in [5]. The grab and release postures allowed the user to *clutch* and *de-clutch* a certain region of interest on the pot. The user could create concave and convex profiles of the pot by *grab-and-push* and *grab-and-pull* actions at the desired location of the pot surface in 3D space. In the second prototype, we implemented our proximal-attraction technique (Figure 2(b)).

Participants & Procedure. Participants within a wide age range (5-60 years) were invited to use our pottery prototype wherein, the task for each participant was to create a pot as per the participant's liking. Although we did not carry out a formal demographic survey, we found that the participants were from a variety of backgrounds including non-technical users, engineers, designers, artists, and professional potters. Our evaluation was mainly informal and observational wherein we recorded videos of sessions subject to the participant's permission and the time taken to complete the creation of a pot. Due to the nature of our venue, we constrained the maximum time for each participant to about 8-10 minutes.

[1] MakerFaire, Bay Area (2013)

Table 1: Behavioral observations in our preliminary evaluation

Age	Value	Behavior
5-10	Fun, Play	Excitement, Random hand movements
11-15	Entertainment, Education	Controlled movements, Explored tool features
16-30	Entertainment, Art, Education	Controlled movements, Investigated pot behavior
30-60	Entertainment, Meditative	Controlled movements, Expected real-world like response

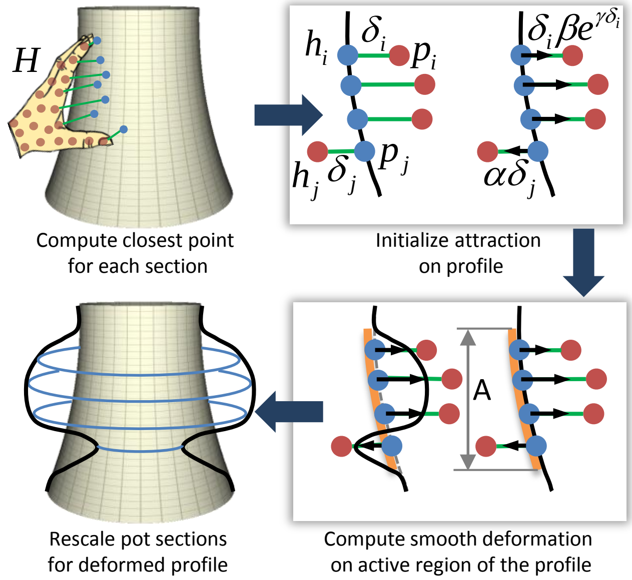

Figure 4: Algorithm for pot deformation is illustrated for proximal-attraction. The profile is deformed based on the proximity of the points on a given hand PCL. Subsequently, each section is re-scaled to obtain the deformed pot surface.

A total of 360 participants responded to our invitation and used our prototype to create pots. In the first session (day 1), 180 participants used the prototype implemented using the grab and release gestures. In the second session (day 2), 180 participants used the proximal-attraction technique for pot deformation. There were participants that were either completely unable to create any meaningful shape of the pot or did not find the resulting shape as the intended one. These attempts we removed from our database leaving us with the recorded times for 113 participants per session (i.e. 226 participants in total).

4.3. Results

We categorized the perceived value and user behavior during the use of the pottery applications on the basis of age. Young participants (5-10 years) were mostly interested in simply playing around with the application and usually applied arbitrary hand movements during the deformation of the pot's profile. Participants in the age range of 11-15 years provided more controlled movements of the hands during pot shaping with slower and more careful hand movements and accurate hand gestures. They also adopted a more exploratory approach towards the applications in that they were primarily interested in the various software features rather than the realism in the pot's deformation.

However, in case of participants above the age of 15, we observed that they instinctively shaped their hands according to geometry of the pot on the screen. Specifically, users within 16 and 30 years of age were mainly interested in investigating how the gesture and motion of the hand was related to the deformation of the pot. They would frequently expect the pot to deform according to how they shaped and moved their hands on the pot's surface. This strongly suggested that the internal learning of physical interactions, combined with some prior expectation of the pot's response, increased with the participants' age. In case of the gesture-based approach, this was also a cause for intermittent gesture misclassification, resulting in user frustration. Despite their simplicity, the *grab* and *release* gestures were tedious to use while using virtual tools. This was mainly the case with participants who were completely new to interfaces developed for RGBD cameras.

On the other hand, users found the proximal-attraction approach easier to learn and use. The participants could immediately start deforming the pot, and at the same time they could shape their hands as they saw fit. A common mental model that the users seemed to create was that of a surface which "*sticks*" to their hands upon coming close. Thus, the users were invariably slower while approaching the pot (so as to reach the right location) and retreated faster when they wanted to release contact with the pot. For some users, fast retreat also caused accidental deformation leading to frustration.

4.4. Takeaways

The two main insights we gained were: (a) the intent for deformation directly translates to how users shape their hand and (b) the rate of attraction for pulling and pushing must be determined separately so as to make them consistent. We found that full-body interactions caused significant fatigue and difficulty in controlling deformation. Thus, our subsequent stages, we implemented interactions at close range wherein a user could perform pottery sitting in front of a desktop or a laptop computer.

5. Hand as a PCL: Shaping by Proximal Attraction

Our main objective in this stage was to adapt the proximal-attraction method that could use the shape of the whole hand to deform the pot. Thus, we used a representation of the hand as a collection of multiple points (i.e PCL) obtained via a depth sensor.

5.1. Technique

Consider the hand H as a set of points $\{h_i\}$ in 3D space. Each point in the PCL deforms a small local region on the pot using the proximal-attraction approach. On the whole this amounts to a gradual and progressive convergence of the pot-profile to the shape of the user's hands (Figure 4).

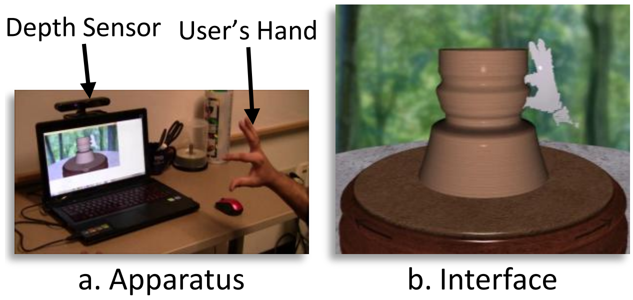

Figure 5: The apparatus (a) consists of the user, a computer and a depth camera. The user sees a PCL of their hand deforming a rotating pot (b).

Pushing vs. Pulling. A push is characterized by an inward displacement ($\delta < 0$). This is the simplest case wherein a user would typically approach the pot and subsequently recede away once the desired deformation has occurred. A pull is characterized by an outward displacement ($\delta > 0$). This is a non-trivial intent to recognize since a user would invariably approach the surface first and then recede to pull. The overall motion of the hand is similar to that of a push. In order to distinguish pulling and pushing, we used two different rates of attraction. For pulling, we defined the attraction rate as a smooth function of the distance between the hand point and pot. The function is given by $\beta e^{\gamma\delta_i}$. For pushing, we defined the rate of attraction as α. This essentially allows the user to first approach the pot without deforming it during the process of approach. The algorithm is as follows:

1. For each section i

    Compute unique h_i such that $\|h_i - p_i\| < \varepsilon$ is minimum.

    Set δ_i to horizontal distance between h_i and p_i

2. Set δ_r to $\delta_{max} - \delta_{min}$

3. Set γ to $\frac{0.1}{\delta_r}$

4. For each i on profile

    if($\delta_i < 0$): Set attraction at p_i to $\alpha\delta_i$
    else: Set attraction at p_i to $\beta e^{\gamma\delta_i}\delta_i$

5. Compute Active region A

6. Smooth deformation ($\nabla^2\delta = 0$ for all points in A)

7. Compute deformed profile

8. Rescale pot sections

Initialization Time. In order to avoid accidental or unintended deformation of the pot, we implemented an that allows for the pot to deform only when contact with the pot is maintained for a sufficient amount of time. We achieved this in two steps. First, we reset α and β to 0 at every new contact that the hand made with the pot. Subsequently, we linearly increase them to their maximum values within a stipulated amount of time T. We call this the initialization time. Intuitively, T is the time taken by the pot to gradually initiate the response to the user's hand after a contact is made.

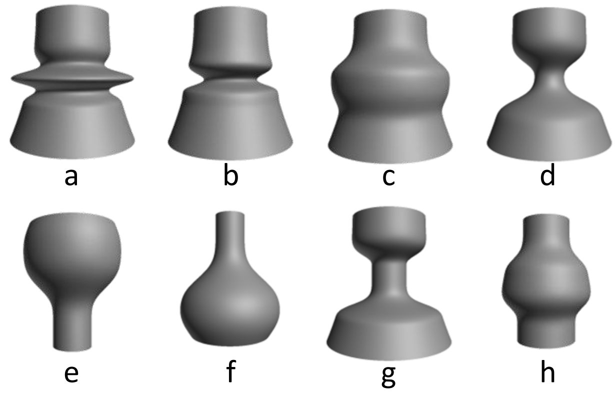

Figure 6: Eight pre-defined pots were shown to participants in the quiz. These are: (a, b) thin convex and thin concave, (c,d) fat convex and concave, (e, f) round and flat, and (g, h) flat at center and ends. (from Vinayak et el. [11])

5.2. Experiment

We conducted a lab experiment to evaluate the proximal-attraction approach. The results of this experiment led us to develop the final approach in this work. In the paragraphs below, we will describe selective details of our prior work for the sake of completeness. For a comprehensive analysis of this experiment, the reader can refer to our prior published work [11].

Apparatus. Our setup consisted of a Lenovo ideaPad Y500 laptop computer with an intel i7 processor and 8GB RAM, running 64-bit Windows 8 operating system with a NVIDIA GeForce GT 750M graphics card, and the SoftKinetic DS325 depth sensor (Figure 5(a)). SoftKinetic DS325 is a close range (0.1m-1.5m) time-of-flight depth sensor that provides a live video stream of the color and depth image of the scene. Every pixel on a given depth image can be converted to a 3D point using the camera parameters.

Implementation & Interface. After segmenting the hand from the scene, we use the SoftKinetic iisu API for tracking the hand PCL. However, the tracking method provided in this API does not work with hand-held objects - a feature that we required in order to allow users to utilize physical objects for deformation. Thus, we used a pre-defined a volumetric workspace as the active region in front of the computer screen. Our interface comprises of a 3D scene with a rotating pottery wheel on natural outdoor background (Figure 5(b)). The user sees the potter's wheel and the PCL of their hands, or the tools held in their hands. We designed this interface based on the guidelines provided by Stuerzlinger and Wingrave [30]. Finally, we provided keyboard shortcuts to the allow the participants to undo and redo a particular deformation at any time. Additionally, we also made provisions for the participants to reset the current shape to the blank pot.

Participants. The participants of this evaluation comprised of 15 (13 male, 2 female) science and engineering graduate students within the age range of 20 – 27 years. Out of the 15 participants, 5 participants self-reported familiarity with mid-air gestures and full body interactions through games (Kinect, Wii). Due to engineering background, most participants (12 of 15) reported familiarity with 3D modeling and computer-aided design. Incidentally, we also had 3 participants who had prior experience with physical ceramics and pottery.

Procedure. The total time taken during the experiment varied between 45 and 90 minutes. We began the study with a demographic surface where we recorded participants' background regarding their familiarity with depth cameras, full-body games, and pottery. Subsequently, we provided a verbal description of the setup, the purpose of the study, and the features of the pottery application. This was followed by a practical demonstration of the pottery application by the test administrator. The participants were then asked to perform the following tasks:

- **P** *Practice:* To get an overall familiarity with the interaction of their hands with the pot surface, each participant was allowed to practice with our interface for a a maximum time of three minutes. The participants were allowed to ask questions and were provided guidance when required.
- **T1** *Quiz:* A pre-defined target shape was displayed on the screen and the participant was asked to shape a "blank" pot so as to roughly match the most noticeable feature of the target shape. We showed a total of eight target shapes in a randomized sequence (Figure 6). The participants were allowed to undo, redo, and reset the pot at any given time and for as many times as they required.
- **Q1** *Questionnaire 1:* Each participant answered a series of questions regarding the association of the deformation to the shape of the hand, responsiveness of the deformation, and consistency of pushing and pulling.
- **T2** *Composition:* The participants were asked to think of (and verbally describe) a set of *intended* pot shapes and subsequently create those shapes using their hands. Although the maximum duration of time for each shape was fixed to five minutes, we allowed the participants to complete their last composition that was started before the end of the specified duration.
- **Q2** *Questionnaire 2:* Finally, each participant answered a series of questions regarding enjoyability, ease of use and learning. The participants also commented on what they liked and disliked about the application, interface and interaction.

5.3. Results

The following paragraphs briefly summarize the observations that we have detailed in our prior work [11].

Reaching, Grasping, & Deformation Strategies. Each user had a different perception of the process necessary to achieve the profile of a given target shape. Most users attempted the quiz problems in multiple trials, wherein they would refine their strategy to deform the profile in every trial. However, we observed that these strategies of reaching, grasping, and deforming the profile converged to patterns common across users (Figure 8). Typically, users would first estimate the size and shape of the grasp according to the geometric feature of the profile and then move the whole hand in the intended grasp to deform the profile [21]. The most common usage pattern observed across users was the recursive smoothing and refining of the pot after deforming the profile reasonably close to the target shape. This was typically done by moving the hand vertically along the surface of the pot (Figure 8). This was the cause of frustration for two reasons. First, the accidental contact of the hand with the pot's surface resulted in unintended deformations. Second, the proximal attractions did not allow for an explicit way to smooth or straighten a region of the pot. Despite being reminded of the undo, redo, and reset functionalities, most users preferred using their hands for reversing an accidental deformation. For the thin-convex profile, most users first created a convex feature in the center followed by pushing the top and bottom portions inward. For concave features, users first pulled the top and the bottom portions of the pot and subsequently pushed the central region of the pot (Figures 9(a)). This was an interesting common pattern since we had assumed that users will create concave features in a single inward action. This was also the case with flat-round features (Figures 9(b)) wherein many users first pulled out the round feature followed by straightening the

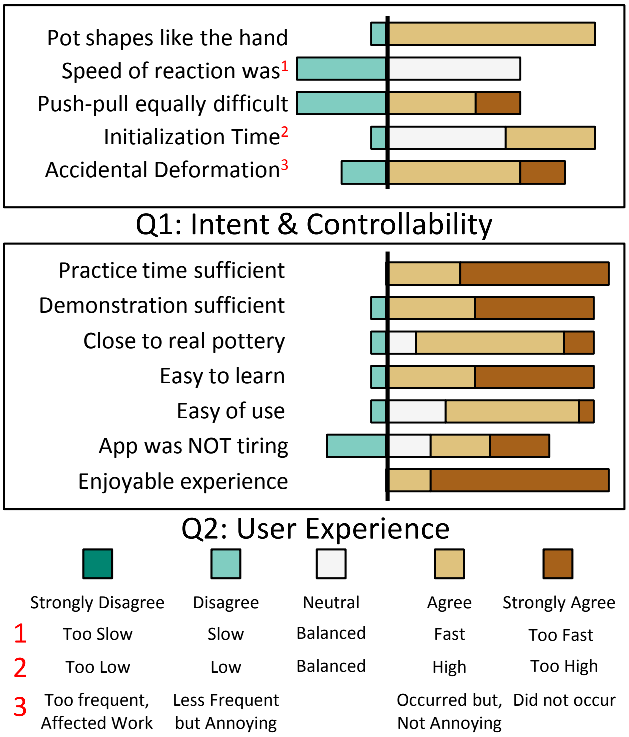

Figure 7: User response to are shown for proximal-attraction. The main issue in terms of controllability (a) was the slow response and difficulty of pushing in comparison to pulling. (from Vinayak et el. [11])

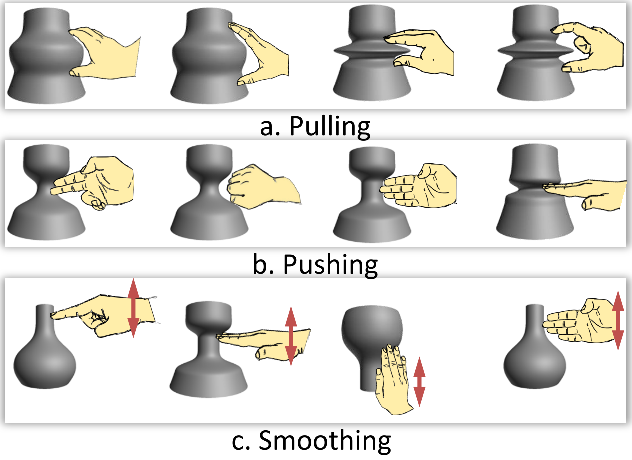

Figure 8: Common user patterns are shown in terms of grasp and motion performed by users for each target shape (in decreasing order of occurrence along columns). The hand images represent the grasp and the arrows (red) show the motion of the hand. The most successful strategies are indicated by blue boxes for each target shape.

flat regions of the pot. The *pointing* posture of the hand was commonly observed during the creation of thin concave features. However, in subsequent trials, most users resorted to using an open palm. This was because the *pointing* pose limited the depth to which the users could push the surface inwards, owing to the interference of the fingers other than the index finger. The cupping of the hands in conjunction with vertical movement of the hands was a common approach for round features.

The use of two hands was particularly prevalent for round-flat combinations. Due to arm fatigue, some users also changed from their dominant hand to the non-dominant hand. This was a cause for frustration due to the limited volume of the workspace and unintended deformations caused by the asynchronous motions of two hands. Most users commonly approached the pot from the sides. The reason, as stated by a user, was: " “*my own hand blocks the view of the pot*”. Difficulty in depth perception caused many users to inadvertently reach behind the pot's surface. This caused further unintended deformations when the user did not expect one, or the lack of response when it was expected.

Intent & Controllability. In general, users agreed that the shape of the profile behaved in correspondence to shape of the hands (Figure 7(a)). However, only 50% of the users agreed that the response speed of the deformation was balanced. There was a common agreement on the initialization time and robustness to accidental deformation. There were two common and expected difficulties that the users faced. These were: (a) pulling specific regions of the pot and (b) creating straight and flat features on the top portion of the pot. As a user stated: “*Pushing seems easier than pulling. Part of the reason I suspect is the visual feedback. It is easier to determine if my hand starts to touch the pot, while it's not as easy to determine if my hand is still attached with the pottery or leaving it.*”. This indicated that

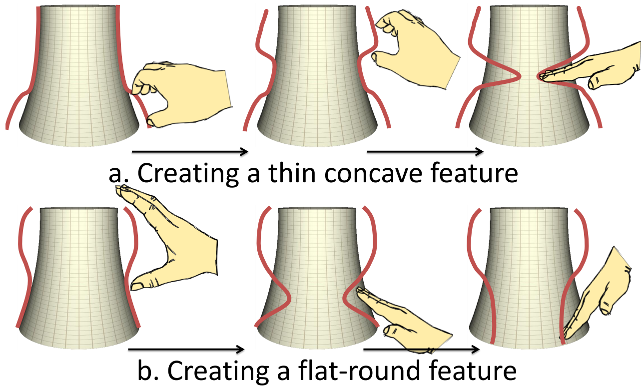

Figure 9: Two examples are shown of common deformation strategies are shown through which users created (a) thin concave and (b) flat-round features. (from Vinayak et el. [11])

perceiving the depth difference between the hand and the pot was difficult for the users.

5.4. Takeaways

There were two main issues with the proximal-attraction approach. First, pulling was clearly more difficult since the rate of attraction was *designed* to be lower than that of pushing. Secondly, the users clearly distinguished between several operations of fairing, straightening, carving, pulling and pushing. However, the proximal-attraction approach, was not designed to explicitly identify or classify the type of operation the user intended to perform. Our main goal in our third and final stage was to resolve these two issues. Our first step was to identify the main characteristics of users' preferences towards grasping to pull and motion patterns for smoothing the pot. Subsequently, the aim was to design a geometric approach that could recognize these identified characteristics and broadly classify the intended actions from the hand PCL.

6. Hand as a PCL: Grasp + Motion

Our observations strongly indicated that users distinguished their intent in three broad categories: pulling, pushing, and smoothing. In our final stage, we implemented a grasp and motion based approach to identify these three classes of intent.

6.1. Technique

The basic idea of the *grasp+motion* approach is to *summarize* the grasp of the hand *in relation to* the surface of the pot and subsequently classify the user's action (Figure 10). We achieve this by using kernel-density estimation of the point cloud on the axis of the pot. In our context, this kernel-density estimate (KDE) is essentially a smoothed histogram of the distribution of the hand's PCL on the pot's. We use the exponential function to determine the KDE. For a given section i, the KDE is given by:

$$\phi_{i,j} = \sum_{j=1}^{j=|H|} e^{a\|\delta_{i,j}^2\|} \tag{1}$$

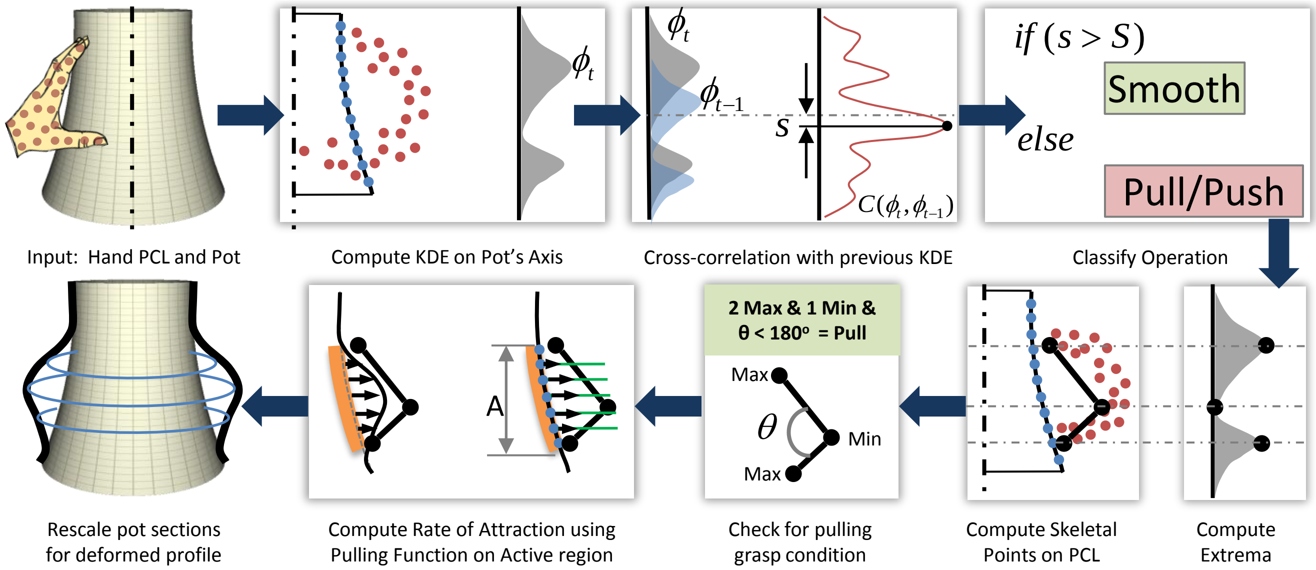

Figure 10: Algorithm for grasp+motion technique is illustrated. The main steps involve computation of axial KDE for hand PCL, detection of intent for smoothing, differentiation between pulling and pushing, and deformation of the pot. In this example, we show the details of the pulling deformation (row 2).

There were three main observations (Figure 8) that helped us use the KDE to classify the user's intent. First, users moved their hands in a fixed pose along the surface of the pot to express their intent for smoothing. This corresponds to detecting the vertical shift of the KDE. We used normalized cross-correlation [31] between the two consecutive KDE signals to determine the shift. Secondly, for pulling the pot, we observed that users used specific grasps. In this case, we note that the KDE has two maxima and one minima (Figure 11). Here, each maxima corresponds to the fingers making contact with the pot and the minima corresponds to the center of the grasp. This essentially allows us to track a basic skeletal representation of the hand. We then define the attraction rate using a based on the angle of grasp (ϕ) (Figure 12). Finally, all actions that do not correspond to either smoothing or pulling, are assigned as pushing. For pushing, we use the proximal-attraction approach for deformation. The steps of the algorithm are:

1. Compute the KDE ϕ_t at time t
2. Compute normalized cross-correlation $C(\phi_t, \phi_{t-1})$
3. Compute Active region A
4. Set s to the shift of correlation
5. if($s < S$): Smooth pot profile in A
6. else:

    Compute extrema

    Detect skeleton

    Compute θ

    if(#maxima = 2 & $\theta < 2\pi$): Apply pulling in A

    else: Apply proximal-attraction in A

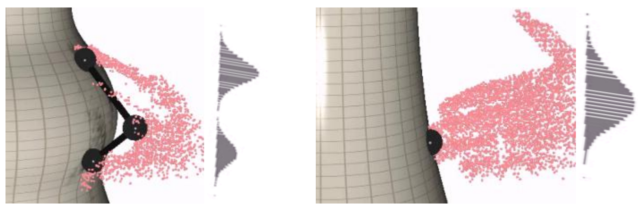

Figure 11: KDE functions are shown for a pulling (left) and pushing (right) intents.

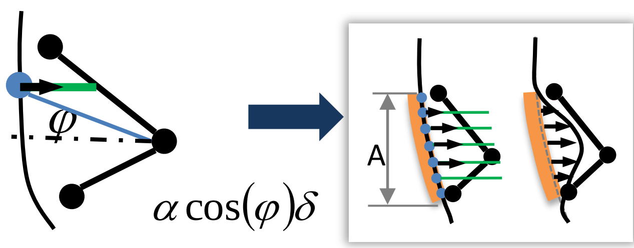

Figure 12: Computation of attraction rate using the angle of grasp.

7. Smooth deformation ($\nabla^2\delta = 0$ for all points in A)
8. Compute deformed profile
9. Rescale pot sections

6.2. *Experiment*

We used identical apparatus and interface to evaluate our final stage. Additionally, we made two important modifications to the interface. First, we added a shadow of the hand on the surface of the pot. The goal was to enable users to estimate their

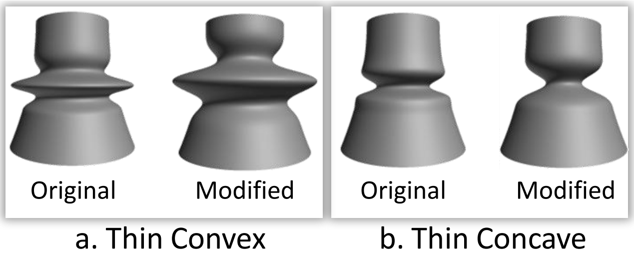

Figure 13: The thin convex and concave features were modified according to the capability provided by the *grasp+motion* technique.

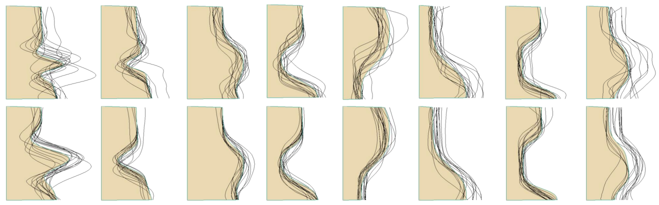

Figure 14: User created pot profiles (black curves) are shown relative to the target shapes (light brown cross sections). The top and bottom rows shows the results for proximal-attraction and grasp+motion approaches respectively. Visual inspection evidently shows improvements in the creation of flat, round and smooth features. More significant improvements were observed in the creation of fat convex features in comparison to proximal-attraction.

proximity to the surface. Secondly, we clamped the hand PCL so as not to allow points on the hand to reach behind the surface of the pot.

Participants. We recruited 15 (11 male, 4 female) participants within the age range of 19 – 30 years. None of these participants had prior knowledge of mid-air interactions or had participated in any of our previous studies with pottery interface. All participants were from science and engineering background wherein 10 participants had familiarity with mid-air gestures and full body interactions, and 11 participants reported familiarity with 3D modeling and computer-aided design. 5 participants reported that they had practical familiarity with real ceramics via informal workshop sessions but did not pursue pottery as a regular activity or professional practice.

Procedure. Our overall experimental procedure was identical to the one that we used for evaluating the proximal-attraction approach (Section 5.2, *Procedure*). However, we made three modifications to the evaluation procedure as listed below:

1. One of the main goals of our work was to enable users to invoke their tacit knowledge of deforming physical objects. To this end, we designed the grasp+motion approach such that it is geometrically-driven and can potentially be used even for user inputs that used other physical objects as tools in addition to the use of hands. In order to verify the generality of our approach with respect to user input, we added another composition task (**T3**) wherein participants were given a duration of five minutes to create pots using a set of physical artifacts as tools. Our "tools" comprised of day-to-day objects (e.g. white-board marker, pair of scissors, ruler) and also some special objects such a Shapescapes^{TM}[2].

2. In order to understand user experience with physical objects tools, we also added questions to the questionnaire **Q2** regarding the utility, ease of use, and preference of tools over hands.

3. We modified the target shapes for the thin convex and concave features (Figure 13). The rationale behind this modification was that the graph+motion technique is sensitive to the size of the hands, finger thickness. Thus, the detection of single-point pulling intent is not possible, as in the case of proximal-attraction.

For each participant and task (**T1**, **T2**, and **T3**), we recorded the completion time and the profiles of the pots shaped by the users. Even though we designed **T1** towards statistical analysis, we observed that each user perceived the target shapes differently and consequently the measured data did not provide sufficient insights regarding the strengths and weaknesses of our approach. With this in view, we present a visual comparison of the numerical data recorded during the evaluation of proximal-attraction and grasp+motion techniques.

6.3. Results

User Performance (T1). Visual similarity with respect to the target shapes evidently increased in comparison to the proximal-attraction approach (Figure 14). This was primarily due to the explicit smoothing. Overall, the completion time (Figure 15(a)) was reduced as expected. Surprisingly, the maximum completion time across all users and all target shapes was recorded for the thin-concave feature (14.4 minutes) followed by the thin-convex feature (13.2 minutes). The mean completion time was highest for the thin-convex feature (3.4 minutes) followed by the central-flat feature (3.3 minutes). The main aspect that we sought from **T1** was the quality of the final outcome across participants for a given quiz problem. We used curvature cross-correlation (*CCC*) as a measure of the quality of user created profiles (see [11] for details). As expected, the smoothness of the results was notably superior in comparison to the proximal-attraction (Figure 15(b)). We also recorded the number of trials per user per target shape (Figure 15(c)). The global maximum number of trials were 7 and 5 for proximal-attraction and grasp+motion techniques respectively. In case of grasp+motion, most users required only one trial for fat-convex, central-flat, and top-bottom-flat features. On the other hand, thin-concave and thin-convex features required more iterations.

Each user perceived and approached a given target shape in different ways. Consequently, there was no evident correlation between the time taken by each user and the quality (*CCC*) of the final pot created by the user for any of the target shape. To

[2] www.shapescapes.com

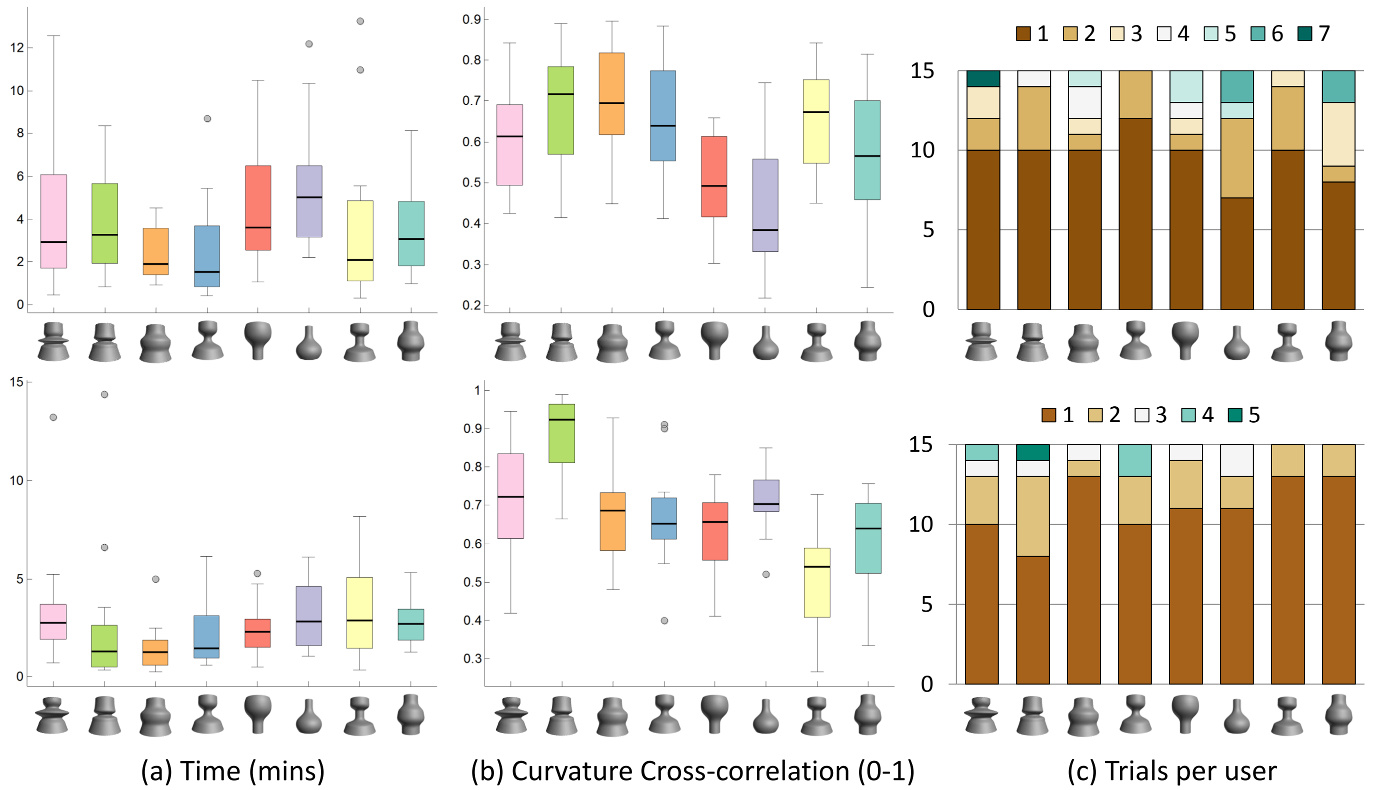

Figure 15: A comparison between proximal-attraction (top row) and grasp+motion (bottom row) is shown in terms of (a) the time taken by users to shape a target profile, (b) the quality of users' responses in terms of curvature cross-correlation of profiles, and (c) the distribution of users with respect to the number of trials per target profile.

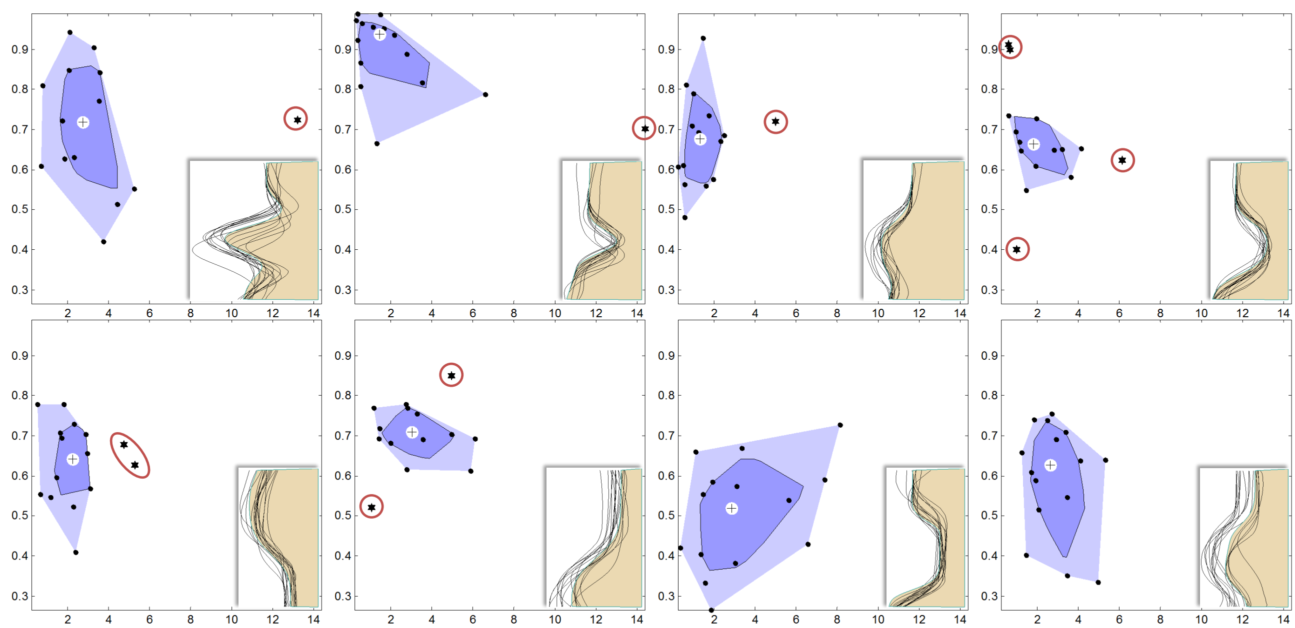

Figure 16: User performance is shown for the each quiz problem as a bag-plot. The x-axis is time in the range [0, 14] minutes and the y-axis is the curvature cross-correlation in the range [0, 1]. The dark and light blue regions show the bag and fence regions, respectively. The white circle is the Tukey depth median and the points marked with red circles are the outliers. The insets show the actual pot profiles (black lines) created by the users in comparison to the target shapes (beige region) of the Quiz. The coordinates of the depth median (C) and the spread (Sp) are provided for each target shape.

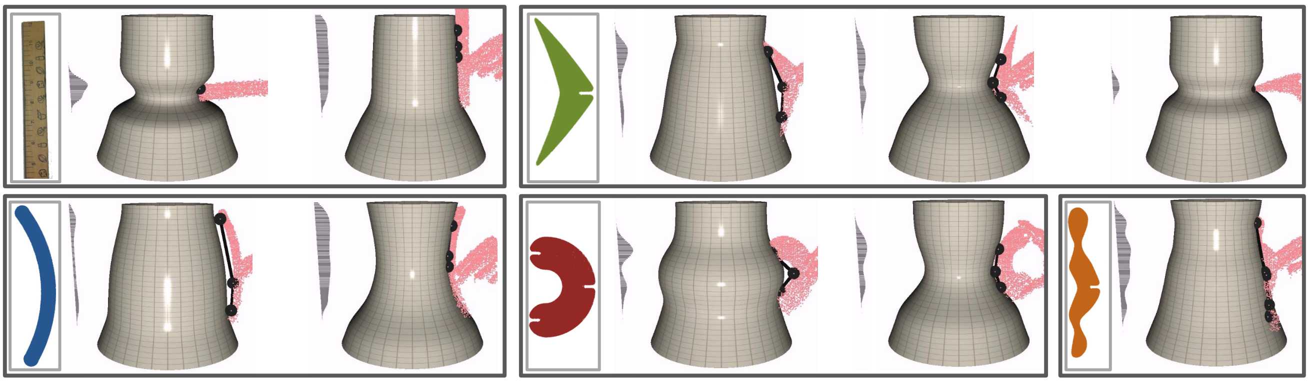

Figure 17: The characterization of tool geometry is visualized for five different physical objects. The objects were chosen to represent concave, convex, flat, and round contacts for deformation.

account for this, we represent the user performance as a bivariate dataset given by the ordered pair of the response quality and completion time. We visualize performance as a bag-plot [32] (Figure 16). Here, the spread of the data (i.e. variations in user responses) is given by the area of the *bag*. Users clearly performed best for thin-concave targets with Tukey median value of (0.94, 1.46). Performance was most consistent for the fat-concave feature (Figure 16(d)). Users also performed consistently for round-and-flat features (Figures 16(e) and (f)). Variations were significant for central flat feature (Figure 16(g)). Further, the pot-profile quality was very low for the central-flat and top-bottom-flat features (Figures 16(g) and (h)). This was mainly because users typically spent considerable time pulling and smoothing the top and bottom regions after performing an initial push. Consequently, the median completion times were also higher for the round-flat and central-flat features (Figure 16(f) and (g) respectively).

Hand Usage (T1). The general user behavior in terms of reaching the pot was similar to the proximal-attraction approach. Both the algorithm and its description was different in this case. The users were explicitly made aware of pushing, pulling and smoothing as three distinct operations. This obviously led to variation in user behavior as compared to proximal-attraction.

Hand Usage (T2). On average, users created 5 pots (max: 12, min: 2) within 5.80 minutes (std: 0.66 min). We made two interesting observations in **T2**. First, we found that users were able to repeat the process of getting from an initial shape to the same final shape across multiple trials. Similarly the users could also deform a current shape back to some previous shape, akin to the *undo* operation, but with the hands. In fact, most participants preferred using their hands to *undo* a pot deformation instead of the keyboard-shortcut. One user stated: "*I thought it was easier to learn the software when I was trying to make my own pot not a model one*". This was expected because of the learning and practice that the users had during the quiz (**T1**). However, during **T1**, users mentioned that their attention was divided due to the need to intermittently look at the target shape during the shaping process. Thus, they generally perceived **T1** to be more demanding than **T2**.

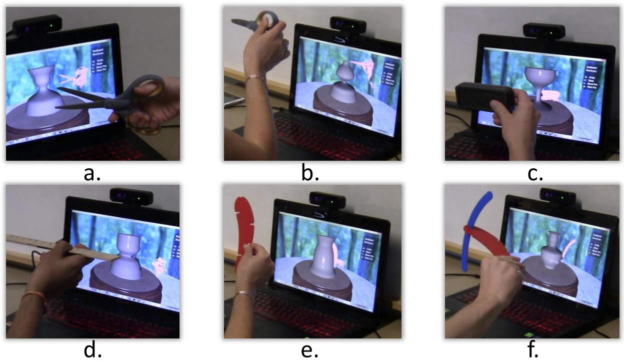

Figure 18: Examples of tool usage are shown.

We made two observations that were not evident in the earlier stages. First, we found that the ability to repeat the process of getting from an initial shape to the same final shape. Similarly, the ability to get to some previous state from the current state was increased substantially. We observed that most of the participants were successfully able to use their hands to *undo* a pot deformation instead of the keyboard-shortcut.

Geometric Characterization of Tools. The choice of everyday objects and ShapeScapes™was mainly helpful in providing a reasonable variety of geometric profiles for pot deformation. However, in order to better understand how users would use these objects, we wanted to pre-determine how the intent of pulling and pushing translates to the use of physical objects. Thus, we conducted a set of experiments (Figure 17) to verify if the users could in fact extend their understanding of the grasp+motion approach and apply it to the use of physical tools. Our experiments showed that the geometry of the tool can indeed be interpreted in terms of the nature of the KDE of the tool's PCL and the grasping angle of the skeleton computed from the KDE. Below, we summarize how this observation came into play during the usage of tools by our participants.

Tool Usage (T3). Users showed immediate enthusiasm during the use of tools. Almost all users first inspected the objects

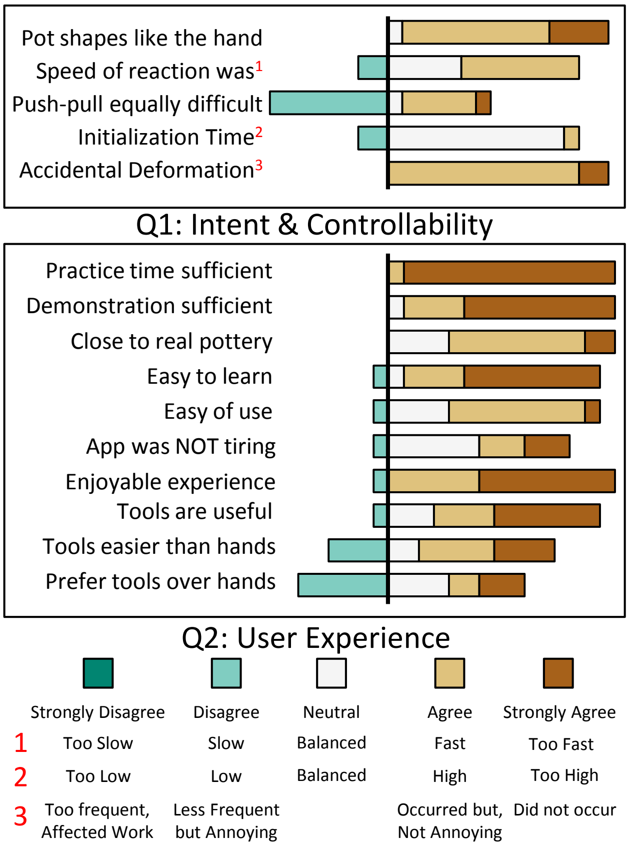

Figure 19: User response to are shown for grasp+motion. While the robustness to accidental deformations was perceived to be negligible (a), many users still perceived pulling to be difficult. Users agreed regarding the usefulness of tools but were not in general agreement about preferring them over hands.

provided to them and planned how to use them for shaping the pots. Users created 4 pots on average (max: 8, min: 2) within 6.0 minutes (std: 0.8 min). In contrast to the use of hands, we observed exploratory behavior in users while using tools. Rather than creating pots, most users were more interested in finding out the effect of each of the objects provided to them. This explained the decrease in the average number of pots in the composition task. One of the difficulties with the use of hands was the inability to create thin concavities. With the use of tools (Figure 18(a),(d)), users could achieve this easily. The most interesting behavior that was observed was the tendency to create convex deformations, which the users achieved by combining two different objects, so as to simulate a grasping hand. This was evident from the users' fascination with scissors (Figure 18(b)). Another important observation was the direct association the users made between the shape of the tool and the purpose it could be used for. The motion of the hand was affected by this association. For instance, while using a whiteboard eraser (Figure 18(c)), the most common motion was that of smoothing the pot. Similarly, for objects with grasp-like geometries, users invariably tried convex deformations by pulling (Figure 18(e)). One user fashioned a new tool by combining different Shapescapes^{TM}parts. This provided the convenience of holding the tool at the "handle" and deforming the pot using fine hand movements (Figure 18(f)).

Intent & Controllability (Q1). We see evident improvements in the perception of intent recognition quality, initialization time, and robustness to accidental deformations (Figure 19). However, despite the decrease in completion time (task **T1**) there was no significant improvement in the user's perception of inconsistency between pulling and pushing. In this case, reason for this perception was primarily related to the visual and tactile feedback rather than the algorithm for pulling itself. This was evident from the user's comments such as: "*I think the reason pushing and pulling were different were because the pulling you had to 2 contacts with the pot and pushing you only needed one. I had a hard time understanding the depth of the pot making it hard to get two contacts on the pot*". One user also suggested: "*I think it would be better if I get some feeling when I touch the pottery. It [would] make me feel more real and easier to control my hand. Then it would be better to have some sounds when I touch the pottery*".

User Experience (Q2). The experience was mostly positive, similar to the proximal-attraction approach (Figure 19(b)). In particular, users liked the use of tools and the smoothing operation the most. One user commented: "*The freeform design with tools was the most fun, as I could spend most of my time focusing on the design aspect as opposed to focusing on minimizing errors.*". According to another user: "*The pottery changing according to my hand shape is so real. While smoothing, I could shape it as well, I like to do it this way a little bit.*".

6.4. Limitations

Our method is currently implemented for pottery, which is essentially a one dimensional deformation. Further, we observed that the use of 2D displays is a factor due to which users tend to use side configurations. We believe that 3D visual feedback will encourage users to access the front and back faces. One user noted: "*This application with haptic feedback could train people for pottery before they actually perform it*". This strongly indicates that the lack of tactile feedback is a critical component that is missing from our current system.

Severe occlusion resulting from camera position and hand orientation is an issue particularly for skeletal based gesture recognition. We partly addressed this challenge using our PCL-based approach which can make use of partial data even when the full hand skeleton is intractable. However, occlusion is an inherent problem in any camera-based method. Investigation of optimal camera position and use of multiple cameras at strategic locations is important. Secondly, we provided a method for temporally adaptive persistence.

In our current implementation, the definition of active regions is in terms of 2D profile topology rather than actual distances in real space. Thus, our implementation is dependent on PCL sampling relative to the mesh resolution of the pot. Independence from the sampling resolution may be addressed with an adaptive approach wherein new sections could be added according to manipulators or old ones removed based on geometric properties of the pot profile such as curvature.

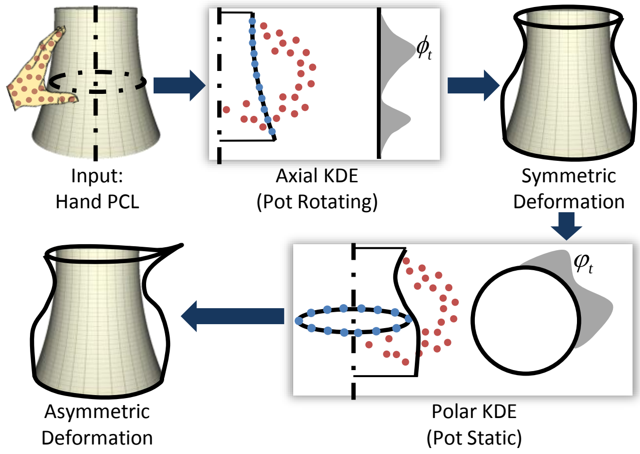

Figure 20: Asymmetric deformation can be applied to a pot in two steps. When the pot is rotating, we apply the *axial* KDE (top row) of the hand PCL for deforming the profile of the pot. Subsequently, users can stop rotating the pot and deform the pot locally using the *polar* KDE (bottom row).

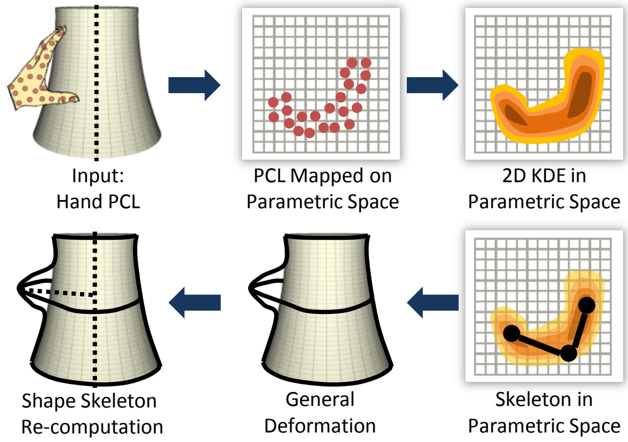

Figure 21: The computation of two-dimensional KDE in the parametric space of a cylindrical surface leads to the computation of grasp and motion for an arbitrary orientation of the hand PCL with respect to the surface. This allows for arbitrary ddeformation of the surface. Recomputing and segmenting the deformed surface using the method of Bærentzen et al. [33] provides a generalized deformation approach using our KDE based approach.

In terms of our evaluation approach, our participants were primarily from science and engineering background. Even though some users had prior experience with creative tasks such as pottery and computer-aided design, studying our approach with art students would provide additional insights on user experience and utility of our approach.

7. Discussion

7.1. *Spectrum of Expressiveness:*

One aspect that is both advantageous and disadvantageous in our approach is that different users can achieve the same target shape using different strategies for grasping, reaching, and deforming a shape. While this provides flexibility and intuitiveness to the user, it also results in increasing the time taken by the user to reach to a desired shape. The evaluation of proximal-attraction evidently indicated that there needs to be a balance between completely free-form interaction and symbolic approaches. This is what we attempted through the grasp+motion approach. The main advantage that our process provided was the discovery of relevant grasp information that is useful to design continuous operations such as shape deformation. Our grasp based approach can serve as a starting point for designing grasp-based interactions using cleaner data such as hand-skeleton [10].

7.2. *Definition of Intent:*

We began with a simple classification of intent through the analogy of structuring operators inspired by Delamé's [28] work. However, users' description of actions and expectation strongly indicates towards a richer and more complex mental model for deformation processes. To this effect, we had to include a third class of operation, namely "smoothing" which evidently improved the performance of the user. Though this aspect is not new in 3D modeling in general, this aspect of refinement is certainly worth investigating from a perceptual point of view.

7.3. *Generalization:*

Although we demonstrated intent classification for rotationally symmetric shapes, the general approach of computing KDE to characterize grasp and motion can be extended to the deformation of arbitrary shapes. Here, we propose such an extension in two steps. First, we will consider asymmetric deformation in the context of pottery itself. For this, we begin by noting that our approach summarizes the hand grasp and motion by computing a one-dimensional *axial* KDE of the hand PCL on the pot's surface. In the same way, we can also compute the one-dimensional *polar* KDE of the PCL (Figure 20). Thus, by combining two one-dimensional KDE computations (axial and polar), we can enable users to create asymmetric features on the pots.

To see how these ideas can be used to conceptualize an arbitrary deformation of a shape, we make two observations. First, the pot is a cylindrical shape with a simple parametric representation and the axis of the cylinder is essentially its skeleton. Thus, given the hand's PCL in an arbitrary orientation with respect to the cylinder's surface, its two-dimensional KDE can be computed in the parametric space as a simple means to determine the grasp and motion of the hand (Figure 21). The consequent deformation of the cylinder would inevitably result in the need for re-computing the skeletal structure of the surface. This is where we invoke our second observation that an arbitrary 3D surface model can be converted to a set of connected cylinders using the recent work by Bærentzen et al. [33] that demonstrates the conversion of arbitrary triangle meshes into polar-annular meshes (PAM). The PAM representation effectively segments 3D shapes into generalized cylinders. Thus, the combination of two-dimensional KDE with the PAM representation can be used for deforming arbitrary meshes.

7.4. *Precise & Selective Reachability:*

One user aptly commented: "*Sometimes it is hard to use the palm because it may deform the surface too much. The context of barely touching does not seem too well implemented. However, if you do this very carefully you can do the barely touching but may make your arm tired a little.*". This is the problem of *precise and selective reachability* wherein one is required to reach and manipulate a local region of an object without affecting neighboring regions. There is extensive volume of work that investigates *distal* selection, manipulation, and navigation [34, 35, 36] of objects. We believe that precision and selectivity are problems worth investigating for close-range, i.e. *proximal* 3D manipulations in mid-air.

8. Future Directions & Conclusions

Our first goal is to extend the grasp+motion approach for arbitrary meshes. This would involve several computational challenges since distance computations and KDE computation would be on 2-manifolds. Secondly, we intend to study how user perception ad performance is affected by adding 3D visual feedback and also tactile feedback. Finally, with our approach, it is not possible to perform deformation using existing hand skeleton tracking approaches. We intend to investigate this in comparison to the PCL based hand representation. One key advantage of using tracked skeletons is that there is a direct correspondence between the fingers and palm which can give useful movement information for better intent detection. This would help segmenting users intentional and unintentional movements [37]. One of the main observations in our preliminary exploration was that users from different backgrounds and age group had different ways of using the pottery tool. In our future works, we want to understand how experience, performance, and creative outcomes will change with respect different user groups such as artists, engineering designers, and young participants.

We presented a spatial interaction technique that uses hand grasp and motion for intent expression in virtual pottery. This approach enables a paradigm shift from existing gesture-based procedural events towards non-procedural and temporally continuous processes in the context of shape deformation. In other words, our work enables users to achieve what they intend in the way they see fit. To the best of our knowledge, no existing hand-based spatial modeling scheme offers such diverse contexts of user input, for instance the use of everyday real objects as tools for virtual shaping, with controllable outcomes. The idea creates new pathways for further research exploring creative design contexts in a "what you do is what you get" framework.

9. Acknowledgments

We thank the reviewers for their valuable feedback and comments. This work was supported by the NSF-AIR Award No. 1312167, the NSF Award No. 1235232 from CMMI-EDI, the NSF Award No. 1329979 from CPS:Synergy, and the Donald W. Feddersen Chair Professorship at the School of Mechanical Engineering. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We thank Dr. Raja Jasti (ZeroUI) for providing valuable suggestions and feedback on our work. We also thank Kevin Lee Jr. (ZeroUI) for the artwork of our pottery interface. Professor Ramani discloses he is a co-founder of ZeroUI that has licensed related technology.

References

[1] Segen J, Kumar S. Gesture vr: vision-based 3d hand interace for spatial interaction. In: Proceedings of the ACM international conference on Multimedia. 1998, p. 455–64.

[2] Wang R, Paris S, Popović J. 6d hands: markerless hand-tracking for computer aided design. In: Proceedings of the ACM symposium on User interface software and technology. 2011, p. 549–58.

[3] Han Y, Han Bj. Virtual pottery: a virtual 3d audiovisual interface using natural hand motions. Multimedia Tools and Applications 2013;:1–17.

[4] Vinayak , Murugappan S, Piya C, Ramani K. Handy-potter: Rapid exploration of rotationally symmetric shapes through natural hand motions. Journal of Computing and Information Science in Engineering 2013;13(2).

[5] Vinayak , Murugappan S, Liu H, Ramani K. Shape-it-up: Hand gesture based creative expression of 3d shapes using intelligent generalized cylinders. Computer-Aided Design 2013;45(2):277–87.

[6] Song J, Cho S, Baek SY, Lee K, Bang H. Gafinc: Gaze and finger control interface for 3d model manipulation in {CAD} application. Computer-Aided Design 2014;46(0):239 –45.

[7] Holz C, Wilson A. Data miming: inferring spatial object descriptions from human gesture. In: Proceedings of the ACM conference on Human factors in computing systems. 2011, p. 811–20.

[8] Leap motion sculpting. https://airspace.leapmotion.com/apps/sculpting/windows; 2013.

[9] Iason Oikonomidis NK, Argyros A. Efficient model-based 3d tracking of hand articulations using kinect. In: Proceedings of the British Machine Vision Conference. 2011, p. 101.1–101.11.

[10] Leap motion skeletal tracking. https://developer.leapmotion.com/features; 2015.

[11] Vinayak , Ramani K. A gesture-free geometric approach for mid-air expression of design intent in 3d virtual pottery. Computer-Aided Design 2015;69:11 – 24.

[12] Vinayak , Ramani K. Hand grasp and motion for intent expression in mid-air virtual pottery. In: Proceedings of the 41st Graphics Interface Conference. GI '15; Canadian Information Processing Society; 2015, p. 49–57.

[13] Ren G, O'Neill E. 3d selection with freehand gesture. Computers & Graphics 2013;37(3):101 –20.

[14] Walter R, Bailly G, Müller J. Strikeapose: Revealing mid-air gestures on public displays. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '13; New York, NY, USA: ACM. ISBN 978-1-4503-1899-0; 2013, p. 841–50.

[15] Karam M, Schraefel MC. A taxonomy of gestures in human computer interactions. Technical Report; University of Southampton; 2005.

[16] Baudel T, Beaudouin-Lafon M. Charade: Remote control of objects using free-hand gestures. Commun ACM 1993;36(7):28–35.

[17] Walter R, Bailly G, Valkanova N, Müller J. Cuenesics: Using mid-air gestures to select items on interactive public displays. In: Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services. MobileHCI '14; New York, NY, USA: ACM; 2014, p. 299–308.

[18] Quek F, McNeill D, Bryll R, Duncan S, Ma XF, Kirbas C, et al. Multimodal human discourse: Gesture and speech. ACM Trans Comput-Hum Interact 2002;9(3):171–93.

[19] Rateau H, Grisoni L, De Araujo B. Sub-space gestures. Elements of design for mid-air interaction with distant displays. Research Report RR-8342; INRIA; 2013.

[20] Jeannerod M. The formation of finger grip during prehension. a cortically mediated visuomotor pattern. Behavioural Brain Research 1986;19(2):99 – 116.

[21] Armbruster C, Spijkers W. Movement planning in prehension: do intended actions influence the initial reach and grasp movement? Motor Control 2006;10(4):311.

[22] Cutkosky M. On grasp choice, grasp models, and the design of hands for manufacturing tasks. Robotics and Automation, IEEE Transactions on 1989;5(3):269–79.

[23] Zheng J, De La Rosa S, Dollar A. An investigation of grasp type and frequency in daily household and machine shop tasks. In: Robotics and Automation (ICRA), 2011 IEEE International Conference on. 2011, p. 4169–75.

[24] Mapes DP, Moshell JM. A two handed interface for object manipulation in virtual environments. Presence 1995;4(4):403–16.

[25] Boulic R, Rezzonico S, Thalmann D. Multi-finger manipulation of virtual objects. In: In Proc. of the ACM Symposium on Virtual Reality Software and Technology (VRST '96. 1996, p. 67–74.

[26] Kry PG, Pihuit A, Bernhardt A, Cani MP. Handnavigator: Hands-on interaction for desktop virtual reality. In: Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology. VRST '08; New York, NY, USA: ACM. ISBN 978-1-59593-951-7; 2008, p. 53–60.

[27] Leyton M. A process-grammar for shape. Artif Intell 1988;34(2):213–47.

[28] Delamé T, Léon JC, Cani MP, Blanch R. Gesture-based design of 2d contours: An alternative to sketching? In: Proceedings of the Eighth Eurographics Symposium on Sketch-Based Interfaces and Modeling. SBIM '11; New York, NY, USA: ACM. ISBN 978-1-4503-0906-6; 2011, p. 63–70.

[29] Holt CC. Forecasting seasonals and trends by exponentially weighted moving averages. International Journal of Forecasting 2004;20(1):5–10.

[30] Stuerzlinger W, Wingrave C. The value of constraints for 3d user interfaces. In: Brunnett G, Coquillart S, Welch G, editors. Virtual Realities. Springer Vienna; 2011, p. 203–23.

[31] Yoo JC, Han T. Fast normalized cross-correlation. Circuits, Systems and Signal Processing 2009;28(6):819–43.

[32] Rousseeuw PJ, Ruts I, Tukey JW. The bagplot: A bivariate boxplot. The American Statistician 1999;53(4):382–7.

[33] Bærentzen JA, Abdrashitov R, Singh K. Interactive shape modeling using a skeleton-mesh co-representation. ACM Trans Graph 2014;33(4):132:1–132:10.

[34] Hand C. A survey of 3d interaction techniques. Computer Graphics Forum 1997;16(5):269–81.

[35] Bowman D, Hodges L. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In: Proceedings of the ACM Symposium on Interactive 3D Graphics. 1997, p. 53–8.

[36] Hinckley K, Pausch R, Proffitt D, Kassell NF. Two-handed virtual manipulation. ACM Transactions on Computer-Human Interaction 1998;5(3):260–302.

[37] Choumane A, Casiez G, Grisoni L. Buttonless clicking: Intuitive select and pick-release through gesture analysis. In: Virtual Reality Conference (VR), 2010 IEEE. 2010, p. 67–70.