Handy-Potter: Rapid Exploration of Rotationally Symmetric Shapes Through Natural Hand Motions

Vinayak *
School of Mechanical Engineering
Purdue University
West Lafayette, Indiana 47907
Email: fvinayak@purdue.edu

Sundar Murugappan
School of Mechanical Engineering
Purdue University
West Lafayette, Indiana 47907
Email: smurugap@purdue.edu

Cecil Piya
School of Mechanical Engineering
Purdue University
West Lafayette, Indiana 47907
Email: cpiya@purdue.edu

Karthik Ramani
School of Mechanical Engineering
Purdue University
West Lafayette, Indiana 47907
Email: ramani@purdue.edu

ABSTRACT

We present a paradigm for natural and exploratory shape modeling by introducing novel 3D interactions for creating, modifying and manipulating 3D shapes using arms and hands. Though current design tools provide complex modeling functionalities, they remain non-intuitive for novice users. Significant training is required to use these tools since they segregate 3D shapes into hierarchical 2D inputs, binding the user to stringent procedural steps and making modifications cumbersome. On the other hand, the use of CAD systems is typically involved during the final phases of design. This leaves a void in the early design phase wherein the creative exploration is critical. We present a shape creation paradigm as an exploration of creative imagination and externalization of shapes, particularly in the early phases of design. We integrate the capability of humans to express 3D shapes via hand-arm motions with traditional sweep surface representation to demonstrate rapid exploration of a rich variety of 3D shapes. We track the skeleton of users using the depth data provided by low-cost depth sensing camera (Kinect™). Our modeling tool is configurable to provide a variety of implicit constraints for shape symmetry and resolution based on the position, orientation and speed of the arms. An intuitive strategy for shape modifications is also proposed. Our main goal is to help the user to communicate the design intent to the computer with minimal effort. To this end, we conclusively demonstrate the creation of a wide variety of product concepts and show an average modeling time of a only few seconds while retaining the intuitiveness of the design process.

∗ Address all correspondence to this author.
1 Introduction

The goal of providing greater and better affordances for natural human-computer interactions (HCI) has largely motivated computational research and technological advancements. The parallel developments of hardware and software technology have continuously provided each other with challenges as well as solutions. The invention of the mouse in 1970 by Doug Engelbart [1] is an example which gives insight about the evolution of our interaction with computers and the corresponding evolution of the ways of creating, manipulating, developing and accessing digital information. The key impact of this invention was that it mapped physical human motion into a corresponding action on a virtual user interface. Recent innovations by Apple® took a significant step forward by further reducing the distance between the computer and the user through capacitive touch based interfaces. The recent success of Microsoft® Kinect™ in the gaming industry is a direct example of the importance of using human motion as an expression, in a non-intrusive way, to create more involving, intuitive and interesting virtual experiences. However, WIMP (windows, icons, menus, and pointers) paradigm is still dominant in HCI despite the development of post-WIMP since the last two decades [2]. Work presented in [3] discusses the basic themes and strategies for reality-based virtual interactions. Wigdor and Wixon [4] discuss that reality-based systems facilitate expert human-computer interaction with little or no prior instructions to the user. To this end, we believe that conceptualization of shapes by human beings is a natural process devoid of any specific tool. Thus, the instructions and training are primarily dedicated towards learning the usage of a modeling tool rather than learning how to think about shapes. This is what drives our research wherein, our intention is to bridge the gap between human expression and digital shape conceptualization during the early exploratory phases of design.

In this paper we propose and emphasize on the metamorphosis of “Computer-aided Design (CAD)” to “Design Aided by Computers (DAC)”. Here we use the term CAD in a sense which is broader than that applied to engineering design. The primary motivation for this paper is to bring the element of design to the forefront of the current engineering process and to eliminate disconnect between the designer and the prototype by naturalizing the design process; particularly when the designer is trying to create rough shapes in large numbers to explore several design concepts. To this end, we introduce a new paradigm for reforming the early-stage design process by developing natural user interfaces (NUIs) to facilitate cognitively simple interactions towards creative and exploratory design without the need for extensive training.

1.1 Human Expression and Shape Conceptualization

Expression and description of shapes via hand and body movements is a frequent occurrence, not only in design, but even in general human communication. Recently, Holz and Wilson [5] gave an interesting and comprehensive description of how natural gesticulation is critical in description of spatial objects. While 2D artifacts like sketches and drawings are better-off being created with 2D interfaces, the creation of 3D shapes using 2D interfaces limits the capability of designers to...
experiment at conceptual and artistic levels [6]. Creation of organic and free-form shapes had been shown in literature using glove-based [7] and augmented reality interfaces [8]. Though these interfaces are promising in providing 3D interaction capabilities for creating a wide variety of 3D shapes, the physical interfaces are often either difficult to setup, involve wearable devices or are expensive to procure.

A device-free alternative is to use computer vision-based gesture recognition to understand the design intent. To this end, classification of gestures is a rich and well-studied problem in computer vision. Work in [9] used compressed 3D shape descriptors using the depth data from ZCam™ for 3D hand pose recovery. Recently, Oikonomidis et al. [10] showed a robust hand pose estimation based on particle swarm optimization on the GPU using depth images from the Kinect™. Despite the vast literature, a holistic combination of computational ease, time efficiency and feasibility of readily usable gesture classifications and their application to shape creation is a challenging task. Further, body skeletal tracking has been applied with greater fidelity for direct use in comparison to gesture recognition at the level of detail of the hands [11]. User satisfaction being the first priority, we take an approach wherein the motion of the human body (particularly the arms) can be robustly used for shape modeling.

Expression of shapes through human actions has been a focus of study in design literature. Horvath investigated hand motion language for shape conceptualization [12]. The naturalization of the process of direct interaction with digitally represented shapes has been addressed in promising field of virtual reality (VR). A recent review of VR-based assembly and prototyping [13] states that “the ultimate goal is to provide an invisible interface that allows the user to interact with the virtual environment as they would with the real world”. However, VR-based technologies are typically more suitable for post-design phases and are less affordable in terms of cost and setup-time. Although we inspire our approaches along the same lines, we do it primarily in the context of the early design phases where iteration of design prototypes necessitates the provision of an affordable and non-intrusive environment which can support exploratory thinking among designers.

Our approach is based on a combination of one-handed and two-handed motions in 3D space towards the definition of the user’s intent in a shape exploration process. Based on the global and relative positions and orientation of each of the hands, our modeling system discerns the action which the user intends to perform. This enables our interaction strategies affordable to a novice user with minimal training since apart from a very few gestures for extremely specific tasks (like switching between modeling modes etc) the user can relate every other motion directly to the primary task at hand - shape exploration.

1.2 CAD and Shape Modeling

The representation of parametric shapes has been studied extensively in CAD literature. A comprehensive description of these representations can be found in [14]. Parametric surface modeling makes use of Bezier, B-Splines and NURBS for modeling complex shapes [15]. Ideas presented in [16] demonstrate the creation and modification of generalized sweep surfaces based on rational motion. Shape deformation, which is an integral part of shape-modeling, has been extensively studied in literature for parametric as well as free-form shapes. Deformation of generalized sweeps have also been investigated extensively in [17, 18], towards applications in human deformation.

In this paper, we approach the shape creation process by using the idea of generalized sweep surfaces. With this approach
we show that a variety of constraints can be applied for generating symmetric as well as free-form models in significantly less modeling time. Further, since sweeps can be represented as a set of cross-sections along a skeletal contour or “spine”, we leverage this representation and propose a novel method for skeleton-based deformation of the shapes created using the generalized sweep concept.

2 Contributions

The main contributions of our work are:

Natural Shape Interaction - With this work, we have initiated the development of a framework to support natural interactions with 3D shapes in 3D space by allowing users to directly create, modify and manipulate 3D shapes without the need for extensive training. Further, we develop this framework using low-cost depth sensing commodity camera (Kinect\textsuperscript{TM}) which requires minimal setup time.

Rapid Shape Exploration - We demonstrate the strength of our natural interaction framework by showing high-speed creation of complex shapes which are other wise difficult to model using existing commercial CAD tools. To this end, we develop the idea of a pottery-based shape modeling metaphor by enabling the on-line creation and modification of generalized cylinders through simple and natural one-handed and two-handed human motion.

Shape Creation Approach - We develop a natural association of human motion with generalized cylinders, which have been extensively studied in literature. We use this association to enable extremely quick creation of a variety of constrained and free-form shapes while retaining the aesthetic characteristics of the shapes. We categorically show that a wide variety of shapes can be created in a matter of a few seconds without the need for training.

Shape Deformation Approach - In contrast to the common inverse kinematics based skeletal deformation methods, we propose a novel spinal-deformation scheme for enabling users to bend generalized cylinders during shape deformation.

3 Shape Exploration Approach

We categorize the shape exploration process into three distinct components, namely, (a) shape creation, (b) shape modification and (c) shape manipulation (Figure 1). We define shape creation as the use of hand and arm motions to create a shape from scratch within an empty working volume. Shape modification is defined as a combination of interactions with shapes with the intention of changing the geometric characteristics of the shape. Shape manipulation refers to the pick-place-orient operations, i.e. rigid body translations and rotations.

The process of 3D interactions is divided into hand and arm motion and hand and arm gesture. Motion is the temporal variation of the spatial configurations of the hand and the arm, i.e. it deals with the dynamic articulations of the hand and arm by obeying some kinematic constraints. In the context of this paper, we define gesture as an interpretation of dynamic hand and arm movements to convey a certain meaningful shape. A typical sweep operation, for instance, can be succinctly defined as a physical process of sweeping a curved piece of wire in 3D space. To this end, we use the pottery metaphor for...
the creation and modeling of sweep surfaces wherein the user moves both hands defining a side profile of the swept surface (Figure 2(a)).

3.1 Shape Creation

In this paper, shape creation is enabled by generalized sweep surfaces i.e. “generalized cylinders” (GC). Given a 2D cross-sectional curve and a 3D skeleton curve, [16] defines a GC as the sweep surface of the cross-sectional curve moving along the skeleton curve. The cross-sectional curve may change its shape dynamically. However, in most conventional methods, the cross-sectional plane is restricted to be orthogonal to the tangent direction of the skeleton curve. GC’s have been well-studied in literature since they have the combined potential of being able to represent fairly complex and free-form shapes and still retaining their parametric nature. In a discrete setting, we define a GC as a surface created by a sequence of sections represented as planar poly-lines along a 3D trajectory or spine represented as 3D poly-lines. A given GC can be described by the position, orientation and size of each of its cross-sections. We use the locations of the hands, \( \vec{p}_i^r \) (right) and \( \vec{p}_i^l \) (left) at a given \( i \)th instance to evaluate these three parameters. Given a cross-section, the distance between the hands, the orientation of the line joining the two hands and the mid-point of the line joining the two hands specify the size, orientation and position of the cross-section respectively (Figure 2(b)). Thus, the temporal variations of the locations of the two hands in 3D space completely define the evolution of the GC for a given cross-sectional shape. Given a planar cross-section \( C \) which tightly fits within a bounding square of unit-length, we represent a GC, \( G \), as a set of scaling factors, \( s_i \), their corresponding center locations, \( \vec{c}_i \) and orientations represented as a \( 3 \times 3 \) rotation matrix \( \theta_i \) (equation 1).

\[
G = \{ C_i | C_i = s_i \theta_i C + \vec{c}_i, 1 \leq i \leq n \} \\
s_i = \| \vec{p}_i^r - \vec{p}_i^l \|_2 \\
\vec{c}_i = \frac{\vec{p}_i^r + \vec{p}_i^l}{2} \\
\theta_i = \text{Rot}(\alpha_i, \beta_i)
\]

In this paper, we confine to closed poly-line sections and open poly-line spines and we consider the creation of constrained generalized cylinder (CGC) and free-form generalized cylinder (FFC) (Figure 3). We envisage a CGC as a surface being created by motion of two hands along a straight spine wherein a cross-section varies in size, retaining a constant orientation and shape. On the other hand, creating an FFC is enabled by additional capabilities of defining the orientation of each section dynamically and curvilinear spines. Thus, from a representational standpoint, the only distinction between CGC and FFC is the level of controllability of \( s_i, \vec{c}_i \) and \( \theta_i \). We implement this controllability by defining snapping of these parameters within prescribed limits with respect to the extent of the motions of the user.

3.2 Shape Modification

In this work we support global shape modifications by providing spinal deformation of a GC. In a minimal sense, a GC can be succinctly represented by two curves, one for the cross-section and the other for the spine. Thus, the general modeling strategy that we follow for the proposed deformations involves the application of univariate functions defined on the curves representing the spine (trajectory)of a given GC. In the following sections, we will demonstrate that the proposed scheme
Fig. 4. Procedure for spine deformation (a) circular bending (b) angular error minimization

allows the designer to conceptualize constrained as well as free-form surfaces. Further, we show that our method allows for free-form deformations while retaining the aesthetic and symmetric characteristics of the shapes modeled.

Spinal deformation entails the bending of the trajectory of the GC. In this paper, we consider bending schemes which preserve the total length of the spine. The spine of the GC is represented as a poly-line i.e. a sequence of points in 3D space. Given such a spine, the user holds one end and drags it to a location of choice. Assuming one end of the spine to be fixed at the base point \( \vec{b} \), the problem is to find a set of rotations of each of the line segments or links of the spine such that the location of the last point moves from the source point \( \vec{r} \) to the location of the user’s hand which we call the target point \( \vec{t} \). This is a typical inverse kinematics (IK) problem for a serial manipulator and has been very well-studied in literature [19, 20]. We propose a novel method of the problem as described in this paper. For simplicity, we assume that the spine is vertical in its initial configuration and the bending happens only on a plane defined by the base, source and target points.

First, we compute the distance, \( d_{bt} = \| \vec{t} - \vec{b} \|_2 \), and the angle \( \gamma = \angle(\vec{t}b, \vec{r}b) \). Then we bend the spine to a circular arc taking \( d_{bt} \) as a chord such that its arc length is equal to the length of the spine (Figure 4(a)). Obtaining the circular bend requires the determination of the radius \( R \) of the circle and the angle corresponding to the chord-length satisfying the constraint preserving the length of the spine (equation 2).

\[
R = \frac{\| \vec{r}b \|}{\theta}
\]

\[
\theta = \arg\min_{\theta \in [0, 2\pi]} \left( \left| \frac{\sin(\theta/2)}{\theta} - \frac{d_{bt}}{2\| \vec{r}b \|} \right| \right)
\]

It can be observed in figure 4(a) that the bending angle \( \gamma^* \) is different than the desired angle \( \gamma \). Thus, in the final step, we minimize the angular error \( |\gamma - \gamma^*| \) by compensating for this error in a piecewise manner according to the hyperbolic tangent function. Equations 3, 4, 5 and 6 give the complete mathematical formulation for the computation of the final coordinates, \( q_i^* \), of the the \( i^{th} \) point \( q_i \) of the bent spine (Figure 4(b)). Figure 5 shows a set of spinal deformations for a bending angle of \( \pi/3 \) and varying target points.

\[
w_i = \frac{\tanh(1/i)}{\sum_{j=1}^{\pi} \tanh(1/i)}
\]

\[
\phi_i = \sum_{j=1}^{i-1} \{ \phi_j + w_j(\gamma - \gamma^*) \}
\]

\[
\phi_0 = 0
\]

\[
T_i = \begin{bmatrix}
\cos(\phi_i) & \sin(\phi_i) \\
-\sin(\phi_i) & \cos(\phi_i)
\end{bmatrix}
\]
Fig. 5. Bending of a unit-length spine for $\theta = \pi/3$ and $0.75 \leq d_{bt} \leq 0.95$ (a) circular bending (b) final bending

$$q_i = T_i q'_i$$

(6)

4 3D Interaction Approach

To provide intuitive interaction with minimal training, our approach uses one-handed and two-handed motions of the user to define what the user intends to do at a given instance during a shape exploration process (Figure 1). For each of the three components of shape exploration, we contextualize the motions of the hand in the following manner. In case of shape creation, we define a thick virtual slab at a specified distance from the global coordinate frame (Figure 6(a)). When both the user’s hands are within this slab, the modeling tool starts creating the generalized cylinders described in the previous section ((Figures 6(b) and 7(a))). While in shape modification and manipulation modes, we use both one-handed and two-handed motions in conjunction with a proximity threshold of the hands with specific parts of the objects. For instance, touching the top face of a generalized cylinder with one hand activates spinal deformation (Figure 7(b)). Holding an object (when either or both the hands are within a prescribed distance threshold of the surface of the object) in manipulation mode with two hands enables the user to translate and rotate the object in 3D space, while only translation is possible when holding an object with with one hand (Figure 7(c)).

Being able to shift between the three components of shape exploration is also an important aspect for a complete experience of shape modeling. Since the prime focus of this paper is the demonstration of the main functionalities of creative shape exploration, defining and implementing the gestural strategies to achieve the inter-mode iterations is out of the scope of the current work.

5 Implementation

To demonstrate the natural shape exploration framework with the pottery metaphor, we developed a modeling tool to support the creation of GC’s. In our current prototypical implementation we assume the the shape of the cross-section to be given a-priori. We particularly experimented with circular, square and hexagonal cross-sections to create shape models. The development of the tool was done using C++ and openGL was used for the real-time visualization of the results. We use an off-the-shelf Kinect™camera in conjunction with the skeletal tracking capability provided by the openNI™library to track the hand locations of the user. The pipeline for the implementation is shown in figure 8.

The interface developed in this tool is very simple in that the user only sees the global frame of reference and a working volume wherein the user can create the GC’s. Taking into account the spatial extent within which a user would typically feel comfortable working, we define the virtual slab at a distance of 1.5 m away from the Kinect™camera and we define the thickness of the slab to be 30 cm which is about double the size of the palm. Figure 9 shows a typical interactive session using the interface developed.

6 Results and Discussion

Figure 10 shows general shapes corresponding to different constraints for the creation of GC’s. We show three varieties of GC’s, (a) fully constrained, wherein the scale of the GC changes with the hand motions while the spine is kept vertical and orientational variations of cross-sections is not allowed, (b) partially constrained wherein the spine can take a free-form on 3D space while the orientation is still constrained and (c) a completely free-form mode wherein the spine, orientation and the section scales can be varied as per the wish of the user. We were able to create these shapes within 2 to
3 seconds which is a significantly low modeling time. However, a better understanding of 3D interfaces for shape modeling warrants a detailed user-study involving the performance of our system and its comparison with traditional CAD tools.

Our modeling metaphor being "pottery", we show a collection of pots which were modeled using the constrained mode (Figure 11). The average time taken per model was about 2.6 seconds. Figure 12 shows a rich variety of shapes and the corresponding time taken to model each shape. The authors would like to stress here that the primary goal of the proposed research is to enable the designer to externalize rough design prototypes for the purposes of exploration in the early phases of design. Thus, it can be appreciated that the description of shapes in the examples shown, involves only the interpretation of what the designer wants to create through his/her own body motion. Moreover, since the actual dimensional details of
a shape can always be specified as a post-process by standard parametric CAD techniques, the 3D shape creation process demonstrated in this paper is more intuitive and natural to use.

7 Conclusions and Future Work

We present “Handy-Potter”, an interactive system for rapid 3D shape exploration through natural hand and arm motions. Our work contributes the paradigm of natural and exploratory shape modeling through novel 3D interactions for creating, modifying and manipulating (the three components of shape exploration) 3D shapes in 3D space. The developed framework
is easy to setup and low-cost with only humans as the actors. We showcased the strength of our modeling framework through high-speed creation of shapes, the stress being on the capability that is provided to the designer to quickly explore a wide variety if shapes with almost negligible training. We present the idea of pottery-based shape modeling metaphor which enables online creation and modification of generalized cylinders through simple human-arm motions. The examples shown in the paper demonstrate that humans can express 3D shapes naturally and intuitively and the description of such shapes involves only the interpretation of what the designer wants to create.

Handy-Potter offers a glimpse at the variety of rich spatial interactions enabled by depth camera for 3D shape exploration. For the interface to be really natural, the designer must be able to seamlessly transition between the three components of shape exploration. Unlike shape modeling with the 2D interaction paradigm, where a significant research effort has been put, 3D interactions for shape modeling are yet to be fully developed and understood from the point of view of usability, ergonomics, efficiency of modeling and visualization. With this in view, the authors see a rich scope for future work in three primary areas. Firstly, designing appropriate gestural interactions and strategies for inter-mode transitions is an important future work for a completely immersive experience in the virtual world. Secondly, in the particular context of the generalized cylinders, naturally inspired methods for a variety of spinal and surface deformations and specification of a rich variety of cross-sectional shapes is the future work which the authors plan to carry out. Finally, we plan to conduct comprehensive user studies to investigate 3D interactions for designing more complex digital artifacts with fluidity. This work shows the potential to move the interactions from a conventional desktop based CAD environment to an NUI enabled spatial environment where the emphasis is on design and the computers become invisible.

Acknowledgements
We would like to acknowledge the support for Professor Ramani by the Donald W. Feddersen Chair professorship for preliminary work in this area. In addition the work in this paper was partly supported by the NSF Partnership for Innovation 3Award No. 0917959 (3DHub) from the Industrial Innovation and Partnership Division, NSF Award No. 1235232 from CMMI-EDI (A-DRIVE) and NSF I-Corps Award No: 1243868. We would like to acknowledge Dr. Raja Jasti for his critical comments and suggestions towards the development of Handy-Potter. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References
Education.


