# *Gesture controlled 3D-modeling and animation interface*

---

User Interface Engineering

**Yavor Stoychev, Stephan Fischer, Daniel Schulz, Plamen Dimitrov**

**30/06/2011**

## Contents

# Requirements Analysis

## Planning

### Background and Motivation

The CG (computer generated) animation is relatively young industry on the entertainment market but develops very rapidly due to the large investments in games and CG movies. Ever more people are trained and employed in the field some of them specialized through education and others starting as amateurs. The importance of the software they are provided with is very important for their productivity or simply satisfaction with their work. A common problem of 3D-modelling is the interface that the user is provided with. Very often manipulation of the objects is not intuitive and there are too many really small buttons structured in a complex way through submenus, vertical and horizontal tabs, multiple view-, property- and toolbars. The keyboard does not have enough buttons for shortcuts, so in 75% of the time button combinations with Ctrl, Shift and Tab have to be used. Navigation is also very often problematic mainly because of the virtual 3D space and the insufficiency of the mouse functionality to implement all of them. With the development of the industry and the sophistication of the final 3D products was significantly raised but with it also the required know-how.

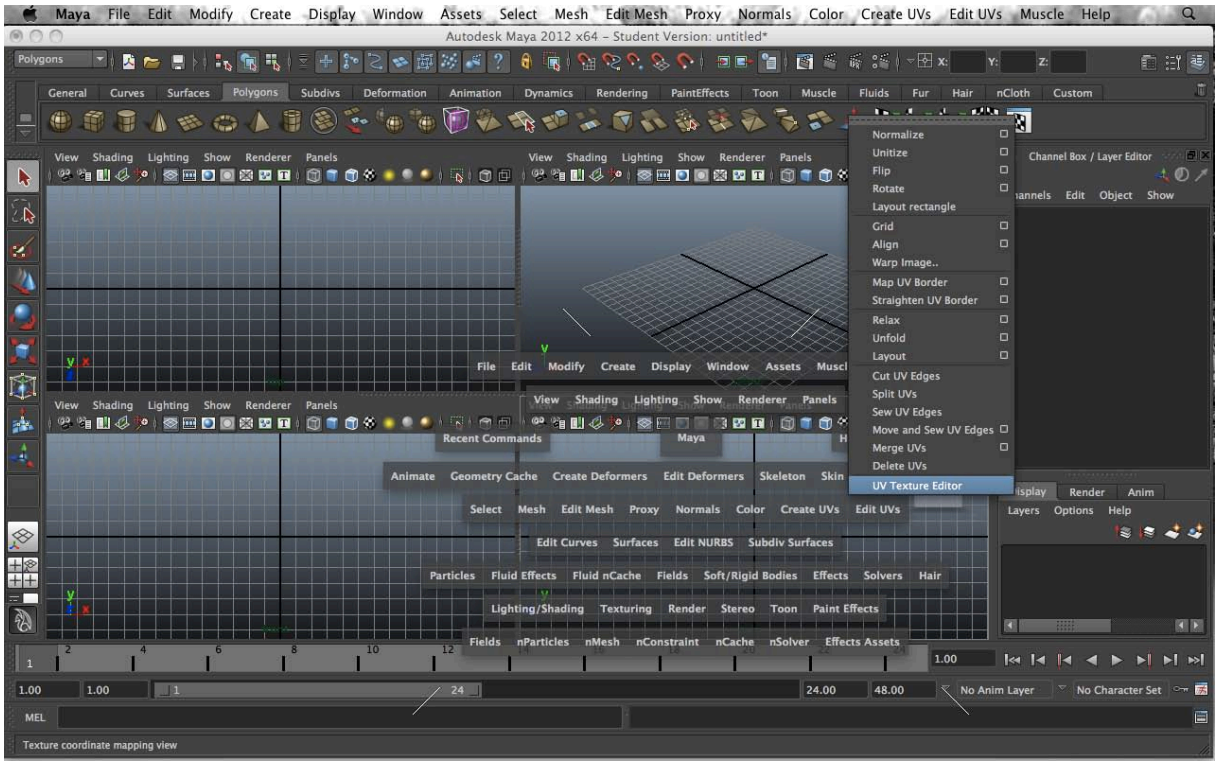


Figure 1.1: The complexity of a contemporary 3D modeling program

The purpose of this project is to simplify and facilitate the process of 3D-modelling by making it more intuitive and interactive. It will help to users that want to create 3D models in a fast and efficient way without obtaining the necessary skills for a 3D modeling program like “Maya” or “3D Max” but also to professionals by increasing their productivity and satisfaction with their work. We want to improve the design process of a 3D model through interactive gesture controlled interface for a more natural and efficient working pipeline.


Figure 1.2: The Kinect camera

The input and output devices to be used in this project are KINECT camera and any type of screen respectively. The Kinect sensor device enables users to control and interact with a computer without the need to touch a input controller, through a natural user interface using gestures and spoken commands. The output can be directed to one or multiple displays including computer screens or projectors.

The underlying data for the project is 3D meshes, composed of points, edges, and faces. There are also textures, materials, colors, norms, light sources and cameras. The project strives to provide free manipulation of the first of these. For the second as well as the final rendering, the model can be saved and exported to Maya, 3DMax or Blender because currently they are not performance critical.

Comparable Work

There is one work particularly close to this idea - "3D Modeling with KINECT, ARDUINO and PUREDATA" by Sebastian Pirch which offers a new innovative approach for modeling in open space with his special gloves on his hands. He uses:

- Kinekt - the 3D camera used in our project
- Arduino – an open-source electronics prototyping platform that can sense the environment by receiving input from a variety of sensors and can affect its surroundings by controlling lights, motors, and other actuators
- Pure Data - a real-time graphical programming environment for audio, video, and graphical processing

Furthermore, Sebastian uses a simplified menu for the four basic operations needed for mesh-modeling: "Create Verticle", "Create QuadFace", "Create TriFace", and "Move". The problem in his work however is the necessity of the gloves as well as the less intuitive interaction relying on sometimes awkward finger combinations. His software can interpret up to a total of 64 (4x4) different inputs most of which are not contextually related to the intended manipulation. Problematic is also the possible long-term health effect of the constant shapes that the hands of the user have to sustain.

More information can be found on the internet address http://www.3rd-eye.at/

## Methods and Materials

Focus Group

As mentioned above the main focus of our project is on professional animators, artists, and 3D-modellers as well as creative users that do 3D-modelling as their hobby. They may have different

experience with 3D modeling and may create CG animations of different sophistication. They may work in an office or at their homes.

User Background / Description of User

Despite the many crossing points in out users' characteristics here we create four different personas that will help us identify possible pitfalls in the design and real customer's needs.

Sheldon

- Profession: Head CG Animator
- Age: 37
- Background: divorced, with two children, likes the small things in life
- Relation to 3D-modeling: well acquainted with older software, solid experience in the history of computer animation

Leonard

- Profession: Game Designer
- Age: 43
- Background: single, unconfident, loves entertainment and MMORPG games
- Relation to 3D-modeling: started from being an artist but later showed interest in 3D-modelling, currently with good experience

Penny

- Profession: Student in Interior Design
- Age: 23
- Background: in serious relationship, loves walking among nature and good parties
- Relation to 3D-modeling: amateur, interested in doing small 3D models

Rajesh

- Profession: Biology Scientist
- Age: 29
- Background: married, careerist, extremely curious and eager to improve
- Relation to 3D-modeling: scientific visualizations of biological processes like cell multiplication, evolution, etc.

Stakeholders

There are different specialized groups when it comes to the creation of a computer animation, some of them less related to the usage of our system, but still necessary to mention.

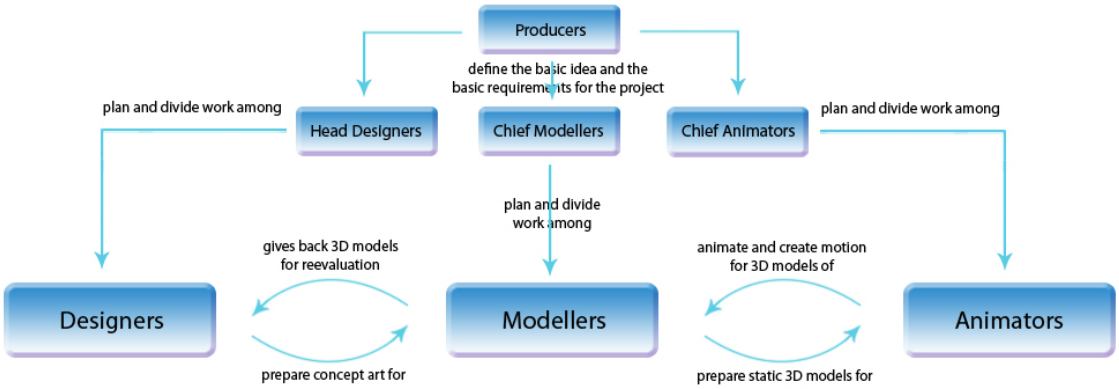


Figure 1.3: Stakeholder diagram

**Animators** create the motions, gestures, and expressions of three-dimensional computer graphics characters and objects. A special feature of our system allows mapping of 3D bones to the actual body of the user which greatly facilitates their working process and therefore proves to be of most benefit.

**Graphic artists** provide graphic art, design and support to feature films, games and related projects. They create still and motion graphics that enhance set and character elements and that support general aesthetic. They usually get in contact with the software for different reasons from drawing some ideas to confirming a CG model as corresponding to their concept art.

**3D-modelers** are the ones responsible for the actual creation of the computer models as they are the connecting gear between the graphic artists and the animators. They are also the main emphasis of our project and our system strives to provide them with the most innovation with regard to the whole interaction process.

There are also head designers, chief artists and animators but we will assume they use the skills of one of the above groups when they interact with the software given the required experience for their position.

## Information Gathering

### Interview and Participant Observation

The goal of our interviews was to get an impression of the work environment and frequently performed tasks of animators, graphic artists and 3D-modelers, but also to get to know their workflow with existing software solutions and their opinions on using a gesture-controlled interface for 3D modeling.

In our interview we included four potential users of our application – two animators, one graphic artist, and one 3D modeler. All of them had a different level of experience and focus. All of them worked with the standard mouse and keyboard input, and used common 3D modeling software - 3ds Max, Blender and MilkShape 3D. All of the interviewees acknowledged that their field has a "medium to steep" learning curve, and most of them attributed this to the un-intuitive and hard-to-use interface found in existing 3D modeling applications. Various toolbars occupy a large portion of the screen, which in its turn reduces the actual working area. Some toolbar buttons have confusing icons, and it is difficult to distinguish between them. Each menu contains at least two layers of sub-menus that in turn lead to a dialog box. This makes even the most trivial of tasks unnecessarily

difficult and tiresome, thus decreasing efficiency and user satisfaction. The alternative, which one of the interviewees suggested was using keyboard shortcuts, but this involves memorizing quite a few key combinations and "turns the whole process into a mortal kombat-like experience".

**Interpretation**

Task Analysis - Outlined Activities

Through observations of the work process in the CG animation field, we have extracted basic set of performed activities that we have to consider. These are also mostly related to our purpose to create an interface that utilizes a Kinect device to provide more intuitive ways of performing basic 3D modeling tasks.

There are three contexts of usage of different subsets of these activities which reduce the large number options and therefore the complexity. Further on, there is transitivity of some of the activities, i.e. their metaphoric meaning can be used in two separate contexts.

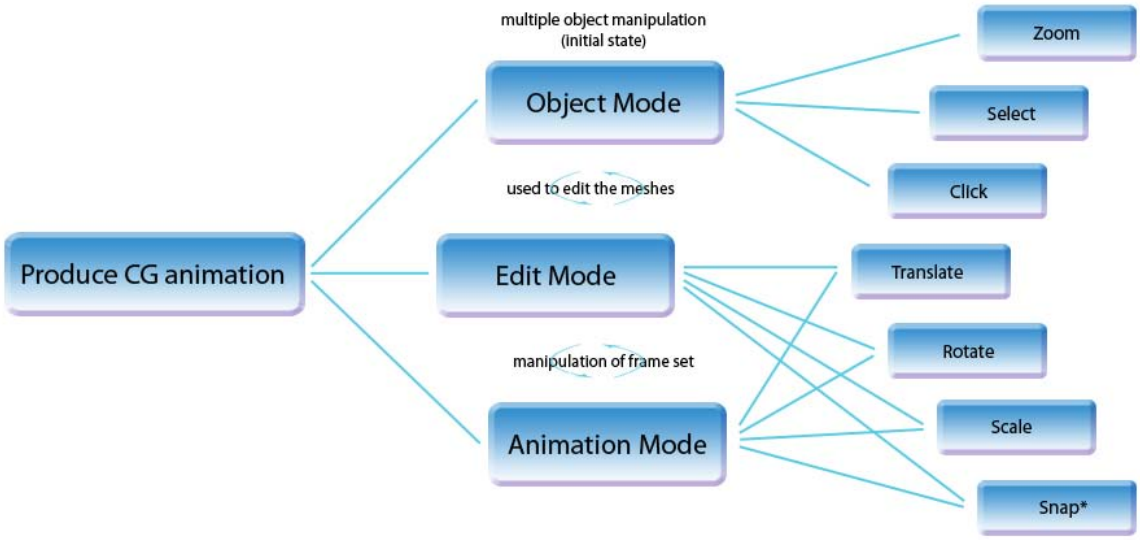


Figure 1.4: Hierarchical task analysis

*snapping will later be used as a special feature in animation mode – snapping bones to modeler's body

**Synthesis**

Problem Scenarios

Scenario 1: Sheldon

Sheldon just started the day early with the usual low-fat cereal. He is happy. He loves sports cars, and now he has to make a CG animation of a brand new Chevy for the new advertising campaign. The Chevrolet representatives have already provided him with the complete painted car model, and all he needs to do is animate it. He creates 15 keyframes, and places the car in a different position at each keyframe. Then, he moves the camera at each keyframe, so that one can get a good look at it

from different directions. Then Maya makes sure that the car switches positions smoothly between frames, and the camera angle slowly changes. He then opens the animation controls, and plays the animation. Sheldon notices, that the viewing angle changes too quickly, and the potential viewers won't have enough time to notice all the features of the design. He then goes back and makes the camera position change less between keyframes. He then runs the animation again. Looks like this time he got it right. Now all he needs to do is fine-tune the motion of the car. He switches to editing mode, and slightly rotates the car at each keyframe, and makes sure that the car position between frames does not make a straight line. This should make it look more realistic. After he is done, he sits back, and places the animation one more time.

Scenario 2: Leonard

Leonard is sitting at home on a warm summer evening. He has no internet and feels disconnected from the world. He cannot play his favorite RPG games on-line. Then he suddenly remembers that somewhere on his external hard drive, there is an unfinished role playing game he started building to impress his friends. The game is mostly finished, but he needs to create more 3D models and include them on the new levels he made earlier. He sighs, plugs in the hard drive to his laptop, and fires up Blender. He adds a series of simple meshes for what would later be a giant blue ogre found at the end of level 4 - very hard to beat. He uses nine cylinders, each of different size for the body, and a UV-sphere for the head. He then switches to editing mode, and focuses on the face. He uses the select tool to select several faces of the mesh. Then he uses the translate and scale tools to add the eyebrows, ears, and the nose. He then uses to the translate tool to make the eyes and the mouth. After he is done with the mouth, he uses the translate tool to make the chin. He then turns his attention to the body. He uses the editing mode tools in the same manner to form the arms, legs, and body. He soon finds out that he needs to add more meshes for the fingers and toes. The work is progressing slowly, and Leonard starts to grow inpatient. So many mouse clicks, and so little work done. He starts dreaming of futuristic user interfaces found in Sci-Fi movies. The LCARS interface from Star Trek would be so much more convenient to use than what we have today.

Scenario 3: Penny

Penny is sitting in the living room and drinking her morning coffee. The party last night was awesome. But today is Monday, and she needs to start working on her class project. It is an interior design project for a new office building in the center of the city. She starts her university-licensed copy of 3ds Max, and creates the empty office space. Then she splits it into rooms. So far, so good. But now comes the difficult part. Penny has to create more complex objects and place them at the appropriate position. She has to arrange a waiting room, and several offices. For this purpose, she has to model a small glass living room table, two sofas, and three chairs and a desk for each office. She starts with the first sofa. For this purpose, she creates two rectangles. Then she uses the select tool to select the second one, and then moves it to combine it with the first. Then she zooms in to make sure that the rectangles are touching. Then she uses the smoothing tool to remove the sharp edges. She uses the same steps to create a table, three chairs, and a desk. Now she would like to replicate the desks and chairs for the remaining two offices, but she is not sure how to do it. She then opens the program manual and starts reading the chapter about manipulating objects. The manual is very technical. It tells Penny where the options are, but she doesn't understand how to use them. This program is not a good choice for inexperienced designers. But Penny knows that Google is her friend, and, after 15 minutes of reading manuals on the Internet she continues her work. And

although it is still hard, and the work is going slow, Penny feels good – she is still a beginner, but she is definitely making progress.

Scenario 4: Rajesh

Rajesh is getting anxious. The Cell Evolution and Mutation conference will take place next Friday, and he will open it by presenting his work on cancer cell mutation and division. He has published his paper two months ago, and his multimedia presentation is almost ready. The thing that is bothering him is that the presentation is very technical, and lacks the multimedia content that would make it vivid and interesting. He decides that he should make a 3D animation of cancer cell evolution and include it in the presentation. He is eager to improve, and now he has a perfect opportunity to do that. He opens MilkShape 3D and creates a sphere. He then uses the scale tools to make it flatter. Then, he creates several duplicates of this "perfect" cell. Later on, he switches to editing mode and plays with each cell to make it look less "perfect". He puts a lot of effort in that and renders the scene numerous times to make sure that it looks the way he wants it to. Then he creates a unique animation for each cell, in which it grows or shrinks, or divides by adding keyframes and changing the respective characteristics of the mesh at each keyframe. Growing and shrinking the mesh is easy, but cell division proves to be a challenge. It is difficult to make the two new cells just touch – Rajesh has to zoom in and out several times to get it right. But as his works slowly progresses, it gets easier and easier. Rajesh is happy – he is getting better at this.

Problem Claims

Problem Claim 1: Sheldon

Sheldon's solid experience helps him to be more efficient and makes it easier for him to model complex motion of objects. However, moving between frames with the mouse is not very convenient, because the frame slider is very thin. Selecting keyframes is also quite tricky, even for an experienced animator, since each frame represents a tiny bar on the playback panel. Furthermore, every time Sheldon needs to switch between frames, he has to move the mouse outside the working area, and drag the slider to the desired position.

Problem Claim 2: Leonard

Even though he's a professional, Leonard faces difficulties when using standard 3D modeling applications. But it is not his fault. These applications are packed with so many features, that even a trivial task like changing the selection mode between vertex, face, and lasso requires traversing several layers of sub-menus. Some of the most common functions have keyboard shortcuts, but it is difficult to remember them all. Thus Leonard often needs to move the mouse outside the working area, and switch between the select, rotate, and scale tools of the editing mode, which distracts him and slows him down. After he has created all the limbs and the head, he has to make sure they are properly attached to the body. Thus, he needs to zoom in, rotate, and zoom out multiple times. Then, if one of the meshes is not touching the other, he needs to move it with great precision, and this is not easy using a mouse.

Problem Claim 3: Penny

Penny is currently in the beginning of a very steep learning curve. Most of the keyboard shortcuts are unfamiliar to her, and she has to use the mouse to switch between modes, tools, and functions. She

has to move the mouse cursor outside the working area to activate the select tool. Then she has to zoom in, rotate, zoom out just to make sure that the building blocks of the sofas, chairs, and desks are actually touching. If they are not, she has to manually move each stray object, very carefully, till it touches. This procedure is usually repeated more than once, and she needs to lock some of the axes before moving each object, which requires entering a sub-menu. This causes many interruptions in her workflow, which slows her down, and makes modeling a tiresome activity.

Problem Claim 4: Rajesh

Rajesh suffers from similar problems. Changing the position of each cell and reshaping it in different keyframes will present a significant bottleneck during his work. Animating cell division will be especially difficult, since the size, the shape, and the position of the cell have to be gradually changed. Rajesh needs to use a lot of keyframes to achieve a smooth animation, and switching between frames is not very convenient. Selecting the right frame usually takes more than a single try. Moreover, every time Rajesh needs to switch between frames, he has to move the mouse cursor to the frame control panel, which is at the bottom, and this interrupts his workflow.

Conclusion

Although current 3D modeling applications provide rich functionality regarding manipulation of objects to bring them to the desired states, they all suffer from a fundamental flaw in their interface – it is too complex. It takes too many mouse clicks to perform even the simplest operations. This contributes to a really steep learning curve, and takes away most of the joy of 3d modeling. It makes it difficult for both beginners and professionals.

# 2 Activity Design

## Exploration

Brainstorming

The results from our explorative brainstorming are the basic intended activities for our project which are also very close to the user's expectation when it comes to 3D animation software. However, in order to optimize the improvements that our project will introduce, here we specify the desired features of these planned functionalities:

| *Activity* | *Needs To Be* |
|---|---|
| **Translating** | Spatially intuitive;<br>Fixed on intervals with precision selected by the user;<br>Controllable (ex. on only in one or two axes);<br>Smooth;<br>Easy and quick to perform; |
| **Scaling** | Noticeable and providing with detailed feedback;<br>Fixed on intervals with precision selected by the user;<br>Controllable (ex. on only in one or two axes);<br>Smooth;<br>Easy and quick to perform; |
| **Rotating** | Rich in functionality;<br>Spatially intuitive;<br>Fixed on intervals with precision selected by the user;<br>Controllable (ex. on only in one or two axes);<br>Smooth;<br>Easy and quick to perform; |
| **Selecting** | Efficient;<br>With adjustable sensitivity;<br>Rich in applicability and selection approaches;<br>Flexible with regard to already selected regions (possibility of interaction with already selected regions like inverting, intersection, addition and difference);<br>Safely undoable; |
| **Zooming** | Spatially intuitive;<br>Fixed on intervals with precision selected by the user;<br>Interactive but not wasting time;<br>Smooth;<br>Easy and quick to perform; |
| **Snapping** | Spatially intuitive;<br>Optimal in precision;<br>Easy and quick to perform; |
| **Clicking** | Comfortable at least as much as clicking with the mouse;<br>Optimal in precision;<br>Easy and quick to perform;<br>Always accessible; |

Table 2.1: Desired features

Conceptual Metaphors

In an attempt to further extend the specifications of the above functionalities we created a collection of metaphors and their implications in the following table:

| *Activity* | *Metaphor* | *Implications Regarding Metaphor* |
|---|---|---|
| **Selecting manipulation mode** | Wearing X-Ray glasses | Transparency of objects, visibility depends on context of manipulation |
| **Selecting menu tools** | Eating chips from a package that refills itself | Always reachable, the constant reaching out is not disturbing and paid any attention to |
| **Translating** | Driving a truck on a road with high traffic; | Caution and obscurity, hard to control and rather controlling the user; |
| | Weather vane; | Disorientation in high dimensional space with few objects, constant direction components required; |
| | Teleporting; | Coordinates-driven, possible to simply specify new coordinates; |
| **Scaling** | Blowing a balloon; | Uniformity but also distortion if needed; |
| | Deepwater fish; | Viewing from different angles creates different impression of size; |
| **Rotating** | Watching pendulum; | No overdoing, essential result somewhere in the cycle; |
| | Decorating only the front of a Christmas tree; | Only certain parts need to be visible, certain angles need to be achieved; |
| **Selecting** | Musical rehearsal; | Frequent repetition, striving for the best final effect; |
| | Coloring a picture; | Constraint to predetermined lines, efficiency of filling the whole figure; |
| **Zooming** | Sewing; | Specific details, focus on operation following it; |
| **Snapping** | Searching for gold; | Many points try to come to the same place and once one of them is finds it all others can find it instantaneously; |
| | Making a tent; | Preservation of supporting points, final shape depends on these |
| **Clicking** | Spicing a meal; | Sensitivity to quantity, even by careful pinching; |

Table 2.2: Conceptual metaphors

System Technology Options

Each activity is realized through the Kinekt camera and a projector or any standard output screen. The focus is on unified input and unified output so that the user is not distracted from the mesh manipulation process and has better grasp on the tools and working artifacts.

**Envisionment**

Activity Scenarios

Scenario 1: Sheldon

Sheldon just started the day early with the usual low-fat cereal. He is happy. He loves sports cars, and now he has to make a CG animation of a brand new Chevy for the new advertising campaign. Chevrolet have already provided him with the complete painted car model, and all he needs to do is animate it. He starts by defines the number of keyframes and importing the car. Creating the keyframes is generally OK, although Sheldon finds it a little bit difficult to work with the tiny slider. He moves the car to the desired position. Then he selects the camera and moves it so that viewers can get a good look at the car from a different direction. He then moves to the next keyframe, and does the same. He repeats the process for all the remaining keyframes. Sheldon expects moving object to be a smooth and lag-less experience. Once the positions are set, Sheldon plays the animation. He notices, that the viewing angle changes too quickly, and the potential viewers won't have enough time to notice all the features of the design. He then makes the camera position change less between keyframes. He then runs the animation again. Looks like this time he got it right. Now all he needs to do is fine-tune the motion of the car. He slightly rotates the car at each keyframe, and makes sure that the car position between frames does not make a straight line. Moving between keyframes with the sliders is a little difficult with the slider. Now that the car doesn't go in an exactly straight line, the animation should look more realistic. Sheldon looks at each of the keyframes and makes sure that everything is the way he wants it to be. After he is done, he sits back, and places the animation one more time.

Scenario 2: Leonard

Leonard is sitting at home on a warm summer evening. He has no internet and feels disconnected from the world. He cannot play his favorite RPG games on-line. Then he suddenly remembers that somewhere on his external hard drive, there is a role playing game he started building to impress his friends. The game is mostly finished, but he needs to create more 3D models and include them on the new levels he made earlier. He sighs, plugs in the hard drive to his laptop, and starts modeling. He adds a series of simple meshes for what would later be a giant blue ogre found at the end of level 4 - very hard to beat. He uses nine cylinders, each of different size for the body, and a UV-sphere for the head. He then focuses on the face. He selects several faces of the mesh. Sometimes he needs to select very specific regions, and he really needs the selection tool to be precise and easy to use. Leonard scales the faces of the mesh. He repeats this procedure several times to create the eyes, ears, nose, and mouth, chin. Scaling should be smooth, and it should support easily switching between scaling with respect to different axes without interrupting his work. Later on, he switches to object mode, and zooms in. Zooming should not necessarily be precise, but it should be easy and intuitive. Leonard selects individual edges and vertices from the face to sculpture it in more detail. He then turns his attention to the body. He shapes the arms, legs, and body. Then he carefully attaches the limbs to the body. He soon finds out that he needs to add more meshes for the fingers and toes. He adds them, models them to his liking, and attaches them to the rest of the body. Putting it all together sounds easy, but Leonard knows better - he zooms in and looks at the Ogre from all sides to make sure that his building blocks are properly connected. This is a rather boring activity, and Leonard wishes there was a way to avoid it.

Scenario 3: Penny

Penny is sitting in the living room and drinking her morning coffee. The party last night was awesome. But today is Monday, and she needs to start working on her class project. It is an interior design project for a new office building in the center of the city. She starts by creating an empty office space. Then she splits it into rooms. So far, so good. But now comes the difficult part. Penny has to create more complex objects and place them at the appropriate position. She has to arrange a waiting room, and several offices. For this purpose, she has to model a small glass living room table, two sofas, and three chairs and a desk for each office. She starts with the first sofa. She creates two rectangles and shapes them to suit her needs by scaling them with respect to individual axes. Penny needs this to be rather easy. Then, she selects the first rectangle, and moves it next to the second one. Then she uses similar steps to create a table, three chairs, and a desk, and put their building blocks together. Then she selects each of them and makes as many duplicates as necessary. The work is going well, and after half an hour, Penny is almost done. Now all that she needs to do is put the objects in their spots. This is just a matter of selecting an object and moving it to the desired position. Dragging 3D objects with the mouse across the screen works, but Penny wishes there was a more fun way of doing this.

Scenario 4: Rajesh

Rajesh is getting anxious. The Cell Evolution and Mutation conference will take place next Friday, and he will open it by presenting his work on cancer cell mutation and division. He has published his paper two months ago, and his multimedia presentation is almost ready. The thing that is bothering him is that the presentation is very technical, and lacks the multimedia content that would make it vivid and interesting. He decides that he should make a 3D animation of cancer cell evolution and include it in the presentation. He is eager to improve, and now he has a perfect opportunity to do that. He starts by creating a sphere. He then resizes it with respect to the Y-axis to make it flatter. Then, he creates several duplicates of this “perfect” cell. Rajesh thinks it will be nice to be able to create a duplicate with just one click. Later on, he plays with each cell to make it look less “perfect” by reshaping it with the scale tool. He puts a lot of effort in that and renders the scene numerous times to make sure that it looks the way he wants it to. After Rajesh is done with creating the cells, he focuses on creating the animations. He sets the length of the animation. Then he creates a unique animation for each cell, in which it grows or shrinks, or divides. He expects to have an unobtrusive way of switching between keyframes at his disposal. All the three changes that appear in the cell can easily be produced using a combination of the scale and sculpture tools. Rajesh is happy – his presentation will look great.

**Rationale**

Activity Claims

Here we present a list of possible and negative aspects of our observation while preparing and reviewing our activity scenarios.

**Object Mode**

| *Feature* | *Pros/Cons* | *Scenarios* |
|---|---|---|
| Zoom | (+) gives an adequate spatial interface and thus makes zooming more intuitive<br><br>(-) lacks fixed-level zooming (ex. 150%, 200%, 250%) | Leonard |
| Select | (+) gives an adequate spatial interface and gives the options to select a vertex, an edge, or a face.<br><br>(-) lacks fixed-size region selection (ex. 320x240x240) | Sheldon, Leonard, Penny, Rajesh |
| Click | (+) easy<br><br>(-) unconventional, might require some getting used to | Sheldon, Leonard, Penny, Rajesh |

**Edit Mode**

| *Feature* | *Pros/Cons* | *Scenarios* |
|---|---|---|
| Translate | (+) very intuitive, just like moving an actual object; supports moving an object along 1, 2 or 3 axes | Sheldon, Leonard, Penny, Rajesh |
| Rotate | (+) intuitive; supports rotating an object around 1, 2 or 3 axes | Sheldon |
| Scale | (+) easy; supports scaling along 1, 2 or all axes | Penny, Leonard, Rajesh |
| Snap | (+) very useful; removes the need to zoom in and manually verify that two objects are connected | Leonard, Penny |

**Animation Mode**

| *Feature* | *Pros/Cons* | *Scenarios* |
|---|---|---|
| Translate | (+)easy to move between frames<br><br>(-) not very precise | Sheldon, Rajesh |
| Rotate | (+) easy to control animation speed | Sheldon |

| | | |
|---|---|---|
| | (-) not very easy to stop the animation (speed 0) | |
| Scale | (+) easy to control animation length | Sheldon, Rajesh |
| | (-) not very precise | |
| Snap | (+) very useful for creating complex animations of humanoid bodies | |

# 3 Information & Interaction Design

## Exploration

### Field Work Artifact

The main work artifacts in the working process of the CG animator are the meshes. The models he or she creates are stored as specific files and when loaded into our software appear as configurations in the workspace. However, the system also supports the final output these meshes are created for – rendered images and animations. It is important to point out that the users are provided with rich export options so that they can later perform post-processing of their models with functionality that is not the focus of our system once they have finished the most intensive work on it which is this focus. Other artifacts are the menu, the 3D cursor which we keep dynamically determined in our application, as well as the teapot (automatically generated mesh) which we decided to keep since it is the charismatic "Hello world" of the 3D modeling.

### Information and Interaction Metaphors

In order to better determine the exact gestures that best fit the mental model of the user we identified additional metaphors regarding the accessing, manipulating and presenting of information.

| *Activity* | *Metaphor* | *Implications Regarding Metaphor* |
|---|---|---|
| **The workspace looks like…** | Different reality; | Focal point of the working process, the user should be engulfed in it; |
| | Movie scene; | Constant review, instantaneous and informative feedback |
| **The menu looks like…** | Rescue squad helicopter; | Appears always and only when really needed |
| | Craftsman's kit; | Wide applicability in almost any case |
| **The buttons look like…** | Natural satellites; | Distinguishable from any background, local in context |
| **Calling the menu is like…** | Surgery | Immediate delivery of required tools, sophisticated situations while getting them |
| **Pressing a button is like…** | Juggling | Balance with things in progress while reaching out |
| **Manipulating the objects in the workspace is like…** | School theater | Expressiveness and creativity with mathematics between rehearsals |

Table 3.1: Information and interaction metaphors

### Requirements for Information Design

We can identify one basic requirement for information design that we strive to fulfill and it is immediate response from the system regarding the manipulation on the meshes. Therefore, it is necessary to ensure quick recognition of the gestures the user is providing as well as quick feedback to be displayed on the workspace. Another requirement regarding the information the system

provides to the user is clean and wide workspace. We elaborate on this requirement further in the "Screen mockup" section where we also show how we fulfilled this requirement.

**Envisionment**

Design Sketches

Based on the extracted metaphors and the claims from the activity design scenarios, we now specify the gestures used in the communication with our system. A compromise had to be made between the Kinekt camera's sensitivity and the improvements on user's experience we introduced in the activity design. We took into account important physical preconditions like physical feasibility and body stress of the gestures when being constantly used as well as the comfort in realizing them.

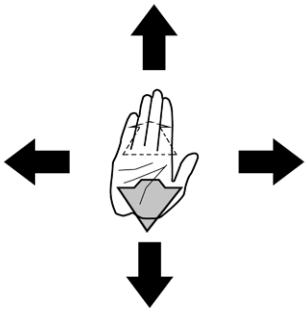

Figure 3.1: Translating

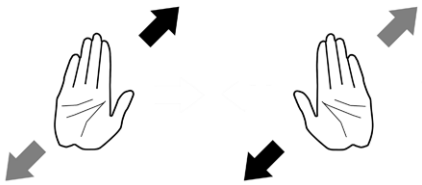

Figure 3.2: Scaling all axes

Translation of object requires only hand so the other can be used for secondary activities like menu selection or simply be ignored by the system and rest, hold a cup of coffee, etc. This hand does not necessarily need to be the right hand – when presence of only one of the hands is detected and the other one is not recognized as present next to it in front of the user operations doable with one hand can be performed.

Figure 3.3: Scaling x-axis

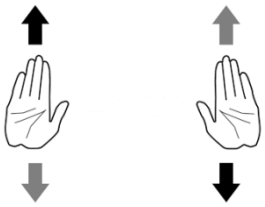

Figure 3.4: Scaling y-axis

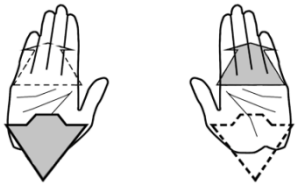

Figure 3.5: Scaling z-axis

While the translation gesture is directionally intuitive for the scale activity we implement general scale gesture but also axis-specific scaling which serves to provide the user with the same flexibility.

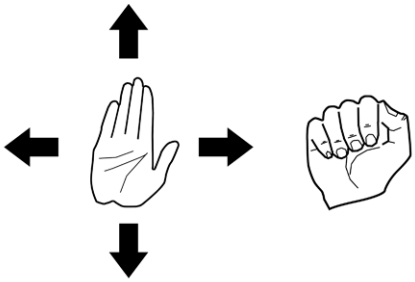

Figure 3.6: Rotation

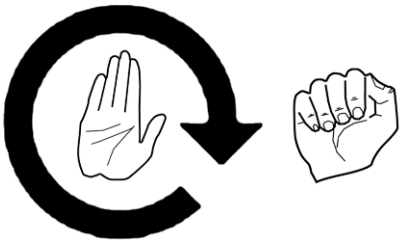

Figure 3.7: Rotation along z-axis

Compared to the previous two, rotation is spatially intuitive in two dimensions but needs additional specification for the third – the z-axis. Therefore we use two gestures for reaching its full functional capacity.

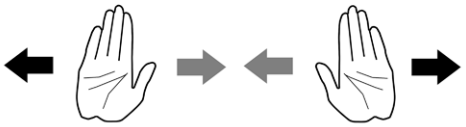

Figure 3.8: Zooming

Figure 3.9: Clicking

Although the zooming gesture is identical to scaling along the x-axis, here we use different context (selection mode) compared to the context above (edit more). This simplification is part of our initial goal to reduce the complexity originating from the great number of options through emphasizing on the context of usage.

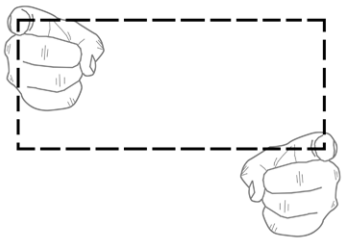

Figure 3.10: Selection box

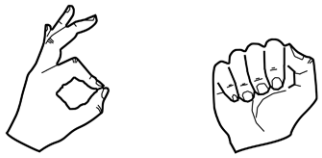

Figure 3.11: Snap

Besides by clicking selection can be done also with the help of selection box. Snapping occurs at currently selected point in space instead of fixed 3D cursor, but can be kept for longer if the selected snapping point coincides with existing point.

Figure 3.12: Change view left-to-right

Figure 3.13: Change view top-down

Two additional gestures represent quick way to change between views which increase the desired accuracy and provides with better 3D orientation.

Screen Mockup

As shown in our mockup below we try to keep clean and visible as much of the main workspace as possible. The main reason for this is that it is the focus of the working process and the user needs to engulf into it in order to feel and be more efficient. Another feature supporting this design goal is the small simplified menu that only appears when it is called by the user.

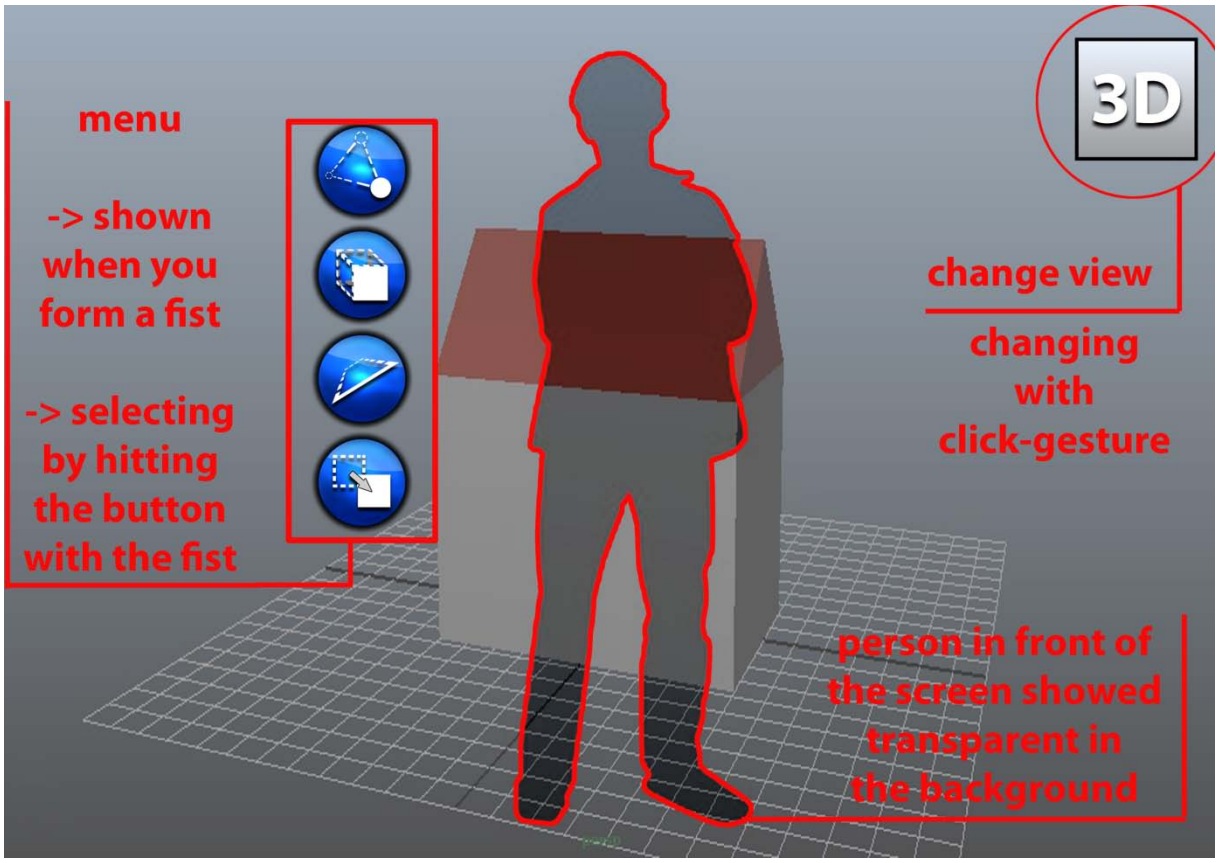


Figure 3.14: Mockup screen

Selection of an option from the menu is realized through hitting button with a fist which however does not imply any requirement for speed or strength, rather simply forming a fist and reaching out for the menu.

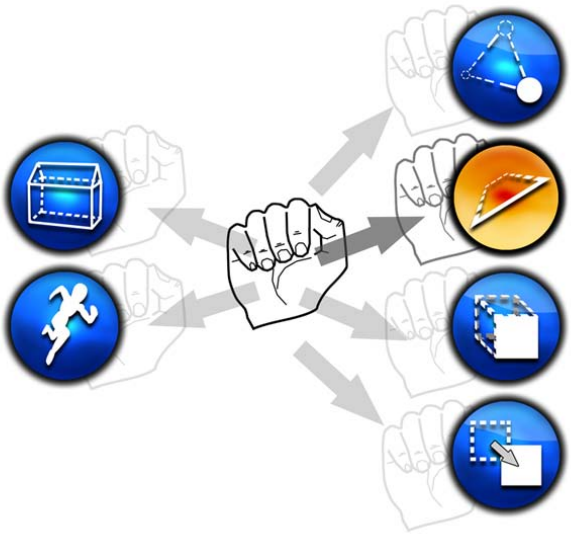

Figure 3.15: Working with the menu

Switching between modes is realized through the menu as well as the most frequent operations that remain too specific for gesture representation as duplicate, select face, edge, and vertex.

Information Scenarios

Scenario 1: Sheldon

Sheldon just started the day early with the usual low-fat cereal. He is happy. He loves sports cars, and now he has to make a CG animation of a brand new Chevy for the new advertising campaign. Chevrolet have already provided him with the complete painted car model, and all he needs to do is animate it. He starts by importing the car model. Then he switches to animation mode, and defines the length of the animation. Then he moves the car to the desired position. Later on, he selects the camera, and then moves it, so that viewers can get a good look at the car from a different direction. He then moves to the next keyframe, and does the same. He repeats the process for all the remaining keyframes. The response time of the system is almost instant. It almost seems to Sheldon that he is moving an invisible object, and can only see its shadow on the screen. Once the positions are set, Sheldon plays the animation. He notices that the viewing angle changes too quickly, and the potential viewers won't have enough time to notice all the features of the design. He then makes the camera position change less between keyframes and runs the animation again. Looks like this time he got it right. Now all he needs to do is fine-tune the motion of the car. He slightly rotates the car at each keyframe, and makes sure that the car position between frames does not make a straight line. He can see the car moving and rotating just as if he is really grabbing and moving it. Now that the car doesn't go in an exactly straight line, the animation should look more realistic. Sheldon looks at each of the keyframes and makes sure that everything is the way he wants it to be. After he is done, he sits back, and places the animation one more time.

Scenario 2: Leonard

Leonard is sitting at home on a warm summer evening. He has no internet and feels disconnected from the world. He cannot play his favorite RPG games on-line. Then he suddenly remembers that somewhere on his external hard drive, there is a role playing game he started building to impress his friends. The game is mostly finished, but he needs to create more 3D models and include them on

the new levels he made earlier. He sighs, plugs in the hard drive to his laptop, and starts modeling. He adds a series of simple meshes for what would later be a giant blue ogre found at the end of level 4 - very hard to beat. He uses nine cylinders, each of different size for the body, and a UV-sphere for the head. He then focuses on the face. He selects several faces of the mesh. The system responds instantaneously, and he can see the selected parts colored in green. Leonard scales the faces of the mesh. It is almost as if he is actually sculpting them with his own fingers. He repeats this procedure several times to create the eyes, ears, nose, and mouth, chin. Later on, he zooms in. He can feel the model closing on him as he performs the relevant gesture. Then he snaps all the pieces together. He soon finds out that he needs to add more meshes for the fingers and toes. He adds them, models them to his liking, and attaches them to the rest of the body using the snap gesture. Using this gesture, he can move an object close to another one, and it automatically snaps to the other object. Leonard finds this feature really cool, as it saves him a lot of time zooming around and exploring the Ogre for improperly connected limbs.

Scenario 3: Penny

Penny is sitting in the living room and drinking her morning coffee. The party last night was awesome. But today is Monday, and she needs to start working on her class project. It is an interior design project for a new office building in the center of the city. She starts by creating an empty office space. Then she splits it into rooms. So far, so good. But now comes the difficult part. Penny has to create more complex objects and place them at the appropriate position. She has to arrange a waiting room, and several offices. For this purpose, she has to model a small glass living room table, two sofas, and three chairs and a desk for each office. She starts with the first sofa. She creates two rectangles and shapes them to suit her needs by scaling them with respect to individual axes. Then, she selects the first rectangle, and, using the relevant gesture, snaps it to the second one. Then she uses similar steps to create a table, three chairs, and a desk, and put their building blocks together. The system responds immediately to her gestures, and the meshes are moving and resizing just as Penny is making the gestures. It almost feels like the computer is predicting her every move. Then she selects each newly-created object and makes as many duplicates as necessary. The work is going well, and after half an hour, Penny is almost done. Now all that she needs to do is put the objects in their spots. This is just a matter of selecting an object and moving it to the desired position. Penny switches to a top view of the scene, and this makes selecting objects a lot easier. It turns out to be a lot of fun. Penny almost feels as if she is inside a Tetris game, moving the pieces with her own hands.

Scenario 4: Rajesh

Rajesh is getting anxious. The Cell Evolution and Mutation conference will take place next Friday, and he will open it by presenting his work on cancer cell mutation and division. He has published his paper two months ago, and his multimedia presentation is almost ready. The thing that is bothering him is that the presentation is very technical, and lacks the multimedia content that would make it vivid and interesting. He decides that he should make a 3D animation of cancer cell evolution and include it in the presentation. He is eager to improve, and now he has a perfect opportunity to do that. He starts by creating a sphere. He then resizes it with respect to the Y-axis to make it flatter. The system responds to his command immediately, and it almost feels like he squeezes it with his hands. Then, he creates several duplicates of this “perfect” cell. Once he activates the duplicate button, another identical cell immediately appears next to the first one. Later on, Rajesh plays with each cell to make it look less “perfect” by reshaping it with the scale tool. He puts a lot of effort in

that and renders the scene numerous times to make sure that it looks the way he wants it to. After Rajesh is done with creating the cells, he focuses on creating the animations. He sets the length of the animation. And he can see the number of frames changing as he moves his hands. Rajesh creates a unique animation for each cell (where it grows or shrinks, or divides). To achieve this, he changes the cell size, shape, and position at each keyframe. Rajesh is really excited about doing this, as it almost feels like he is holding a single magnified cell in his hand, and he moves it and stretches it.

Interaction Scenarios

Scenario 1: Sheldon

Sheldon just started the day early with the usual low-fat cereal. He is happy. He loves sports cars, and now he has to make a CG animation of a brand new Chevy for the new advertising campaign. Chevrolet have already provided him with the complete painted car model, and all he needs to do is animate it. He starts by importing the car model. Once the car is imported, Sheldon switches to animation mode by first making a fist to bring up the menu buttons, and then punching the newly-appeared animation button, and uses the scale gesture to set the animation length. When he is done, he brings up the button menu and punches the edit button. Then, using the translate gesture, he moves the car to the desired position. Now it is time to play with the camera. He uses the pointing fingers of his hands to select the camera, and then moves it, so that viewers can get a good look at the car from a different direction. He then moves to the next keyframe (by switching to animation mode, and using the translate gesture), and does the same. He repeats the process for all the remaining keyframes. He then switches to animation mode and plays the animation by rotating his right hand clockwise. Sheldon notices, that the viewing angle changes too quickly, and the potential viewers won't have enough time to notice all the features of the design. He goes through the keyframes again and makes the camera position change less. He then runs the animation again. Looks like this time he got it right. Now all he needs to do is fine-tune the motion of the car. Using his hands, Sheldon slightly rotates the car at each keyframe, and makes sure that the car position between frames does not make a straight line. This should make it look more realistic. Now he switches back to animation mode, and uses the translate gesture to browse through individual frames and make sure that everything is the way he wants it to be. He uses the rotate gesture to control the speed and direction of the animation playback, and he slows it down at times to verify that everything looks OK. After he is done, he sits back, and places the animation one more time.

Scenario 2: Leonard

Leonard is sitting at home on a warm summer evening. He has no internet and feels disconnected from the world. He cannot play his favorite RPG games on-line. Then he suddenly remembers that somewhere on his external hard drive, there is a role playing game he started building to impress his friends. The game is mostly finished, but he needs to create more 3D models and include them on the new levels he made earlier. He sighs, plugs in the hard drive to his laptop, and starts modeling. He adds a series of simple meshes for what would later be a giant blue ogre found at the end of level 4 - very hard to beat. He uses nine cylinders, each of different size for the body, and a UV-sphere for the head. He then focuses on the face. He uses his two fingers to select several faces of the mesh. Then he scales the faces using a two-hand scaling gesture. He repeats this procedure several times to create the eyes, ears, nose, and mouth, chin. Later on, he switches to object mode, and uses the zoom gesture to zoom in. Leonard selects individual edges and vertices from the face to sculpture it in more detail. He then turns his attention to the body. He uses the edit mode gestures in the same

manner to form the arms, legs, and body. Then he uses the snap gesture to snap the limbs to the body. He soon finds out that he needs to add more meshes for the fingers and toes. He adds them, models them to his liking, and snaps them to the rest of the body with a simple gesture.

Scenario 3: Penny

Penny is sitting in the living room and drinking her morning coffee. The party last night was awesome. But today is Monday, and she needs to start working on her class project. It is an interior design project for a new office building in the center of the city. She starts by creating an empty office space. Then she splits it into rooms. So far, so good. But now comes the difficult part. Penny has to create more complex objects and place them at the appropriate position. She has to arrange a waiting room, and several offices. For this purpose, she has to model a small glass living room table, two sofas, and three chairs and a desk for each office. She starts with the first sofa. She creates two rectangles and shapes them to suit her needs using the scaling gesture. Then, she draws a rectangle using the pointing fingers of her hands to select the first rectangle, and, using a snapping gesture, snaps it to the second. Then she uses similar steps to create a table, three chairs, and a desk, and snap their building block together. After all the unique objects are created, Penny selects each of them, makes a fist to bring up the buttons, and punches the duplicate button to create as many duplicates as necessary. The work is going well, and after half an hour, Penny is almost done. Now all that she needs to do is put the objects in their spots. This is just a matter of selecting an object and moving it to the desired position using the translate gesture. To make things even easier, Penny uses a click gesture (just touching her thumb with her pointing finger), and switches to a top view of the scene. Now selecting and moving the objects is a child's play.

Scenario 4: Rajesh

Rajesh is getting anxious. The Cell Evolution and Mutation conference will take place next Friday, and he will open it by presenting his work on cancer cell mutation and division. He has published his paper two months ago, and his multimedia presentation is almost ready. The thing that is bothering him is that the presentation is very technical, and lacks the multimedia content that would make it vivid and interesting. He decides that he should make a 3D animation of cancer cell evolution and include it in the presentation. He is eager to improve, and now he has a perfect opportunity to do that. He starts by creating a sphere. He then uses the Z-axis scale gesture to make it flatter. Then, he creates several duplicates of this “perfect” cell by making a fist and then punching the duplicate button. Later on, he plays with each cell to make it look less “perfect” by reshaping it using the scale gesture. He puts a lot of effort in that and renders the scene numerous times to make sure that it looks the way he wants it to. After Rajesh is done with creating the cells, he focuses on creating the animations. He switches to animation mode by first making a fist, and then punching the animation button, and using the Scale gesture, sets the length of the animation. Then he creates a unique animation for each cell, in which it grows or shrinks, or divides. To animate a growing or shrinking cell, he simply uses the scaling gesture to scale it a little bit more at each keyframe. To animate a dividing cell, he uses the duplicate button and creates two cells that almost overlap, and slowly moves them apart in each keyframe using the translation gesture. Rajesh is happy – his presentation will look great.

**Rationale**

Information & Interaction Design Claims

From out information and interaction scenarios we extracted the following claims:

**Information Claims**

| *Feature* | *Pros/Cons* | *Scenarios* |
|---|---|---|
| Animation Mode | (+) allow the user to focus entirely on his work without any interruptions; provides an intuitive way of previewing animations; | Sheldon, Rajesh |
| Object Mode | (+) allow the user to focus entirely on his work without any interruptions; | Sheldon, Leonard, Penny, Rajesh |
| Button Panel | (+) makes the buttons visible only when the user activates it; preserves space and avoids distraction;<br><br>(-) two gestures are required to activate a button; | Sheldon, Leonard, Penny, Rajesh |
| User "shadow" in front of screen | (+) allows the user to "calibrate" himself to the display, and makes interactions with on-screen objects more natural;<br><br>(-) might be distracting at first; | Sheldon, Leonard, Penny, Rajesh |

**Interaction Claims**

| *Feature* | *Pros/Cons* | *Scenarios* |
|---|---|---|
| Translate | (+) smooth and intuitive; allows easy manipulation with respect to one, two, or all three axes; | Sheldon, Leonard, Penny, Rajesh |
| Rotate | (+) allows easy manipulation with respect to one, two, or all three axes;<br><br>(-) not very precise | Sheldon, Leonard, Penny, Rajesh |
| Scale | (+) intuitive; allows easy manipulation with respect to one, two, or all three axes; | Sheldon, Leonard, Penny, Rajesh |

| | | |
|---|---|---|
| Snap | (+) very useful; saves the trouble of zooming and manually attaching two objects; | Leonard, Penny |
| Toggle Button Menu | (+) easy; does not interrupt the workflow; after toggling the menu by making a fist, the user is ready to hit a button; | Sheldon, Leonard, Penny, Rajesh |
| Activate Button | (+) easy; does not interrupt workflow; naturally flowing gesture after toggling the button menu;<br><br>(-) somewhat unconventional; | Sheldon, Leonard, Penny, Rajesh |
| Change View | (+) the gesture version of a mouse click – easy; saves the hassle of manually rotating an object to a left/right/top/bottom/front/rear view; | Penny |

Usability Inspection

The main goal of our project is to provide means to manipulate 3D objects more intuitively for the purposes of 3D modeling and animation. We tried to design an interface that is unconventional, but at the same time easy to use by both professionals and amateurs. We designed gestures, that are easy to perform, and they remain so even after many repetitions. We expected that most of the potential users will not have much experience with gesture-controlled interfaces. We made our best to present our users with an interface that is easy to get used to. We reuse existing gestures to perform various actions in animation mode while making sure that these gestures are still intuitive in the animation context. We also included snap functionality in both the object and animation context, to make the traditionally difficult actions of animating human motion, and snapping objects together easier. This is especially useful for people like Penny and Leonard, who have to create complex objects by combining smaller parts. We also made sure that the interface provides adequate tools for professionals like Sheldon and Rajesh. All of them can perform their tasks using our prototype, and, although some drawbacks are unavoidable, our interface provides certain advantages over traditional interfaces for all of them. It offers an easy and intuitive way of controlling animation playback and editing keyframes with minimal interruptions of the workflow. It allows easy shaping and snapping of complex objects. It allows performing typically complex operations without traversing layers of submenus, and it turns shaping objects into a game.

# 4 Prototype

To demonstrate the features of our prototype we created a video mockup. This video shows the basic features of our system: translating, rotating, scaling, duplicating, changing view, animation, vertex, edge and surface manipulation.

The video is made out of a split screen where you can see a person standing in front of a monitor interacting with our interface and the screen itself to watch what's happening on the screen. The prototype and materials related to it can be found on the CD accompanying this documentation.

# 5 Conclusion

The focus of our course work is placed upon the research and development of software that through new technologies attempts to fill an old gap in the field of CG animation. The problem with the usability of computer animation modeling software is one of the classical problems of insufficiently friendly user interface and has already been approached by some researchers. However, it remains unsolved due to many factors like the strong specialization required by the field, the rich functionality that is essential and yet cannot be handled by friendly enough interfaces, etc. The growing industry of computer animation in movies, games and other media speaks for itself. Ever more people get involved in the field and the necessity to put the scope of their work back to the more human aspect is becoming more critical with time.

In this course work we propose the idea of using Kinekt camera and replace most of the encoded functionalities with gestures and keep the rest in a compact helpful menu. Using Scenario-Based Design and going through Analysis and Activities, Information and Interaction Design we reached certain conclusions. By developing the sketches, mockups, a prototype, metaphors and scenarios, and considering wider range of possible users, we made the observation that if simplification is realized in an efficient enough way it may not lead to loss of functionality and rather identify and reduce completely redundant complexity. We decided to concentrate our efforts on dealing with the most important issue – the one of object's manipulation, and targeted further functionality as later improvements. Making the life of the computer animation modeler easier was of highest priority for us together with widening the accessibility to other types of users. These are also our suggestions for the future development of this kind of software. They actually gave use one more lesson – that if we deal with a product where people's imagination is applied to produce another product, we have to do our best to make it easy and accessible, but most of all continuation of their own creativity.

## References

http://siliconangle.com/blog/2010/11/08/kinect-camera-feeds-already-cracked-hacker-not-interested-in-the-bounty/

http://usa.autodesk.com/maya/

http://www.3rd-eye.at/