A METHODOLOGY USING ASSISTIVE SKETCH RECOGNITION

FOR IMPROVING A PERSON'S ABILITY TO DRAW

A Thesis

by

DANIEL MEYER DIXON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2009

Major Subject: Computer Science

A METHODOLOGY USING ASSISTIVE SKETCH RECOGNITION FOR IMPROVING A PERSON'S ABILITY TO DRAW

A Thesis

by

DANIEL MEYER DIXON

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Tracy Hammond
Committee Members,	Richard Furuta
	Ann McNamara
Head of Department,	Valerie Taylor

December 2009

Major Subject: Computer Science

ABSTRACT

A Methodology Using Assistive Sketch Recognition

for Improving a Person's Ability to Draw. (December 2009)

Daniel Meyer Dixon, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Tracy Hammond

When asked to draw, most people are hesitant because they believe themselves unable to draw well. A human instructor can teach students how to draw by encouraging them to practice established drawing techniques and by providing personal and directed feedback to foster their artistic intuition and perception. This thesis describes the first methodology for a computer application to mimic a human instructor by providing direction and feedback to assist a student in drawing a human face from a photograph. Nine design principles were discovered and developed for providing such instruction, presenting reference media, giving corrective feedback, and receiving actions from the student. Face recognition is used to model the human face in a photograph so that sketch recognition can map a drawing to the model and evaluate it. New sketch recognition techniques and algorithms were created in order to perform sketch understanding on such subjective content. After two iterations of development and user studies for this methodology, the result is a computer application that can guide a person toward producing his/her own sketch of a human model in a reference photograph with step-by-step instruction and computer generated feedback.

DEDICATION

To God the Father in Heaven, whose works and guidance have sustained me despite my lack of faith at times and my generally undeserving character. He promises to make good come from any hardship to ultimately show how awesome He is, but you have to prevent yourself from digging your own hole deeper.

To Jesus Christ, the Son of God, whose steadfast willingness to do what God the Father had for Him to do made the way for everlasting life for anyone who truly desires it. "Jesus said to her, 'I am the resurrection and the life; he who believes in Me will live even if he dies, and everyone who lives and believes in Me will never die. Do you believe this?'" (John 11:25-26 NASB)

ACKNOWLEDGEMENTS

These are the friends, peers, and advisors who have had more patience with me that I deserved: Benjamin Spratling, Dr. Tracy Hammond, Manoj Prasad, Dustin Talk, the members of the Sketch Recognition Lab, my co-workers at the Information Technology Issues Management group, Dr. Andruid Kerne, Dr. James Caverlee, Mike and Karen Davis, Beth Marinari, Dr. Hank Walker, Tina Broughton. Thank you for being part of God's continuing work to refine me into a man of integrity. I pray for you all.

I must also thank my committee in its entirety: Dr. Hammond, Dr. McNamara, and Dr. Furuta.

I cannot justify through words the blessing that my family is except that they got me to this day because they're awesome and God is at our core.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

Many people have adopted the idea that they are simply unable to draw well, and can be quick to confess this belief if asked to put pen to paper [1]. To such a person, drawing seems to have a barrier since the user knows that there is something not right with the look of the sketch, but s/he does not know what should be done to fix it. Much of the ability to draw comes from how one sees and processes the object being drawn. An experienced drawer has the ability to see contours and spaces formed by an object and the relationships between the two, while the novice drawer produces more simplified representations of the object's features. A classic example of this is how a novice sketcher will incorrectly draw the eyes too high on the head – rather than drawing them in the vertical center of the head where they actually reside – because all of the perceptually important features on a face occur on the lower half of this face. In other words, people draw what they perceive rather than what they see. This makes it difficult for people to draw pictures with realism. Helping the timid drawer overcome this mental obstacle in his/her drawing skills leads assistance in drawing to be a perfect domain for sketch recognition.

As a sample scenario, a person learning to draw may want to create a portrait working from a reference image of a famous individual, a family member, or even him- or herself. A computer application could provide feedback to such a student to help him/her improve the sketch while drawing. This thesis proposes a methodology where a

This thesis follows the style of *IEEE Transactions on Visualization and Computer Graphics*.

computer application would first process the image using face recognition software to obtain information about the human form (e.g. the location of the nose, eyes, mouth). As the student draws, the application would use sketch recognition to understand and quantify the error of his/her freeform strokes to aide in the artistic instruction. The application would provide real-time interaction with reference media that would enhance how the student perceives it. If the student becomes confused or the application determines that s/he drew something wrong, constructive feedback in multiple forms can guide the student to learn how to fix the mistake. By interpreting the user's drawing and knowing the desired goal, the application is the guide that assists in making the drawing meet the goal yet the student is the one actively creating all of the strokes on the page and thus still owns the artwork.

This thesis presents the first artistic methodology and application to use sketch recognition to assist the user in creating his/her rendition of a reference image with the intent of improving that person's ability to draw. Given that the artistic domain poses many challenges for use with sketch recognition, unique approaches and principles had to be developed for determining what to recognize within a drawing, how to provide helpful assistance, how to correct drawing issues, and, most importantly, how to interact with the user.

2. THE ART OF DRAWING

Drawing is considered a skill that is learnable at any age and not an innate ability [2, 3], of similar difficulty to learning how to read and write or play baseball. Everyone is able to establish a basic competency. Much of drawing is the act of reproducing what is actually seen, maintaining proportions and alignment of elements, and not falling victim to what one's perception may wrongly produce. The key requirements of drawing could be that of any skill: learn the principles, find a challenge, and practice often. Drawing is also not a skill of muscle memory or extraordinary coordination. As Edwards (1999) states, "…drawing is a skill that can be learned by every person with average eyesight and average eye-hand coordination—with sufficient ability, for example, to thread a needle or catch a baseball. Contrary to popular opinion, manual skill is not a primary factor in drawing." Without the defunct misconceptions that drawing is reserved for only the gifted, the idea of providing assistance in drawing within a computer application should not be out of the question.

2.1 Human Perception and Drawing

How does one describe how an artist processes a scene? One popular theory involves a division of functions between the left and right hemispheres of the brain. During the 1950's and 60's, studies using animal and human subjects led researchers to believe that each hemisphere of the brain may develop asymmetrically and be specialized in complementary, but different, modes of thinking [4]. This idea led to the right-brain versus left-brain, or "R-mode" versus "L-mode," theory [3]. This theory states that these two halves of the human mind respond to stimulus differently and that a person is generally predisposed to respond using one half over the other. For a majority of the time, these two modes do work in tandem and contribute their abilities for processing the task at hand. At other times, however, this theory states that one mode may appropriately become the leader and dominate the activity or the two may even be in conflict, where one is attempting to do a task that the other is better at. Those with a perceived inability to draw are responding to the task using their L-mode, when the R-mode is best equipped to process visual information. Aiding the mind in switching to the R-mode is the crucial first step in preparing oneself to begin a drawing task. As Edwards describes, the R-mode is where “we understand metaphors, we dream, [and] we create new combinations of ideas,” while the L-mode “analyzes, abstracts, counts, marks time, plans step-by-step procedures, verbalizes, and makes rational statements based on logic.” A further side-by-side comparison of the two modes is provided in Table 1.

Table 1. Summarization of the two modes of perception. There modes are L-Mode and R-Mode and were proposed by Edwards (1999).

L-mode	**R-mode**
Verbal – Uses words to label and describe. *Analytic* – Solves problems in a part-by-part manner. *Symbolic* – Uses symbols to stand for objects. *Abstract* – Selects a piece of the information available to represent the entire thing. *Linear* – Inclined to arrange ideas via links in an order and toward a convergent conclusion.	*Nonverbal* – Processes data without the use of words. *Synthetic* – Perceives the whole by putting items together. *Actual, Real* – Processes things as they are really seen at the current moment. *Analogic* – Determines likenesses between items and establishes relationships. *Holistic* – An object is seen all at once, including its overall patterns and structures and toward a divergent conclusion.

Looking at Figure 1 of sketches by two male participants in their twenties working from the same reference image, the sketch on the left is a byproduct of the L-mode and marked by symbolic representations for facial features. For example, an eye is drawn in a basic representation, using two curves and a black dot; this neither reflects a realistic eye nor the one found in the reference photograph, but is merely symbolic of one. In contrast, the perception of the R-mode processes the facial features as they are seen at the time and without the need of symbols or words. The drawing on the right of Figure 1 was formed using the lines, shapes, and spaces that the model's face consists of, using the object as a whole and resulting in a more realistic representation. The thought process of the person desiring to draw must become less about recognizing known features of a familiar object (e.g. eye, nose, mouth) lest one give into the tendency to recreate these as his/her L-mode wants them to be. This has been stated as "draw what you see, not what you think."

(a) (b)

Figure 1. Contrasting example of two participants (a,b) drawing the same reference image (middle). Each perceived and produced two very different products.

One part of assisting the person in drawing, therefore, becomes creating the proper conditions for a dominant L-mode to reject the task, shifting to the R-mode, and allowing the person to truly "see" an object. A simple example is to flip the view of an object to be upside down (Figure 2). When normally encountered in its usual orientation, the L-mode can recognize and categorize the features seen based on matches with existing memories and facts. However, when turned upside down, the recognition problem may not be so trivial and can cause the needed pause to realize and see the details that are not normally given attention (e.g. shadows, shapes, curves). It is this shift of consciousness to detail that should be fostered.

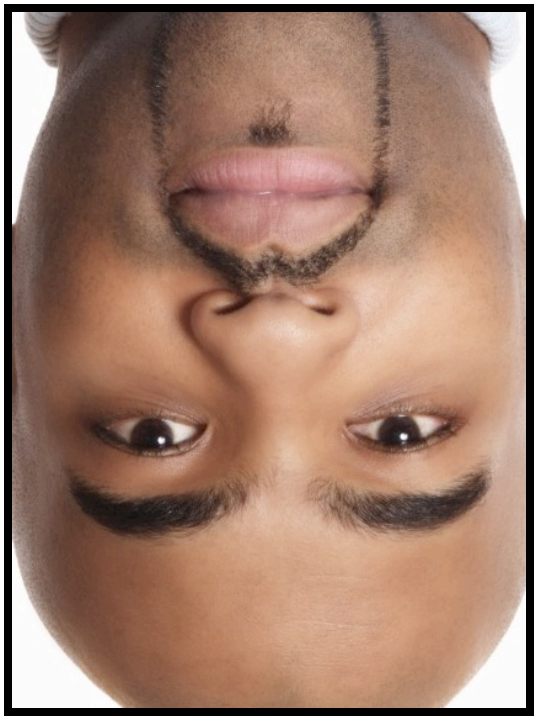


Figure 2. Example of presenting material to help a shift to R-mode.

Some [5] have pointed out that theories such as this one that come from educational psychology obtain their backing from the scientific understanding of the mind, not the scientific understanding of the brain. Bruer specifically states that the notion of brain lateralization (left brain versus right brain) is a speculation adopted more by educators and that actual brain research does not defend such a disjoint handling of task types between the two hemispheres. Edwards acknowledges the critique and placates any additional ones by leaving others to show what actual neurological processes are occurring behind her theory. She states years of empirical evidence amongst herself and her students that support the use of the L-mode and R-mode analogy in art teaching and instruction. The author of this thesis also adopts the analogy as a premise for constructing his application and methodology, using the terms of the analogy (i.e. "L-mode" and "R-mode") instead of those that claim localization of functionality in the brain (i.e. "left-brain" and "right-brain").

Another study [6] analyzed the drawing ability of a group of children ages 5 to 17. Willats characterized the drawings of each child based on the projection (orthographic, perspective, or none), the perspective (classified by the angle of convergence and the angle of obliquity), and the relationship between objects chosen by the child. From this study, Willats created classifications based on these features and then concluded that learning to draw for a child means the progression through a series of stages where the child gradually moves up each classification. As the child grows older, s/he can become frustrated with the limitations of his/her current drawing ability and subconsciously seeks an alternative classification (e.g. one that includes perspective projection) that lets him/her achieve the steps toward the realism desired. Willats also comments that much of this progression comes from explicit teaching, and finds a correlation from drawing to language and communication. The author of this thesis finds a relationship to this work and that of Edwards [3], who could be said to believe that Willats' progression of stages could extend past childhood. When children experience frustration or discouragement with their drawing ability, Edwards believes that the many simply stop seeking to improve. Then, as adults, they are capable of being stuck in an early stage and still producing drawings of a more simplistic childlike classification. The conclusion is that one's drawing quality reflects his/her level of instruction and not the lack of an innate ability.

2.2 Survey of Art Instruction and Education

Among the approaches for teaching a drawing task, not one is universal [1, 3, 7-10]. In the classroom setting, the teacher may not show an example of what to draw, but instead openly define the task for the student to expand upon [7]. The argument is that this is to the student's advantage to learn how to see his/her own path to the task while avoiding design fixation, which the notion that the student's final creation will be very similar to the example given. On the opposite end, a book or online tutorial can guide the student quite well in a step-by-step fashion to recreate a well-defined end product. While in art class, having students imitate the original work of others may be discouraged [11], though some feel that pupils can initially learn much by reproducing the masters [3, 10]. The young child also learns from and finds simple enjoyment in tracing a favorite action hero and other iconic figure. The author of this thesis has concluded that a distinction exists between encouraging creativity in an art student's work and helping the beginner to achieve the desired level of realism in his/her sketches. This work is built upon the latter notion.

If the teacher feels something is incorrect with a student's drawing, some say that s/he should never mark or motion a "correct" stroke but instead use an open question or two to discover the student's thought process [11]. Two art students of college-level art programs who were interviewed as part of a literature review both said it would be appalling for a professor to make a mark on a student's work. However, this may not always be applicable such as when the student is learning through an online art school or is taking a class for beginners, or when the one teaching is a peer helping the student.

When it is just the student and a book or an online tutorial or any other situation where there is a concurrent disconnect from an art teacher, there may be no one to provide constructive feedback, leaving the student possibly stumped as to what improvement is needed to make his/her drawing look right. An application such as the one posed in this thesis should be seen as a complementary tool for a book on drawing instruction that is able to support an art teacher's requirements for a student's homework assignment.

Proving the validity of this work from another angle is the lack of importance put on art instruction in grade school [3, 12, 13]. Art education in the early 20th century was more crafts-oriented and focused on practical needs for life but a shift occurred with the introduction of the concept that art was an innate ability of the gifted [13]. "As a result," Orde states, "the role of the art teacher changed from instructor to nurturer. Instead of the rigid training of specific skills, the teacher became a provider of materials who encouraged self-expression." Even with research dating back to the 1920's, standardized testing in art education has been absent, especially on a national level. Some state this inability to quantify artistic skill affects the perception of administrators in education, as well as the general public, when comparing with more measurable disciplines [12]. The use of and motives behind standardized testing for visual arts are also topics of debate by those in the art education community. Many of the visual arts disciplines also lack a standard curriculum, which inhibits standardized testing. Drawing and sketching in particular are not part of any widely accepted standardized testing though some fields of the visual arts are, such as dance, music, and other performance-based disciplines. However, the claim is that artistic ability is normally distributed among the public [3, 12, 13]. Orde opens a discussion that ties the visual perception and spatial ability learned through drawing with other areas that require these same traits (e.g. mathematics, sciences, engineering, design, even computers) and continues by asking if learning to draw might improve general intelligence. In her words, "Drawing training, because of its relationship with visual literacy and other cognitive areas, is essential to the total educational development of a child and has particular significance with respect to the ever-advancing communication technologies." If the fundamentals of sketching can be captured and assessed by a computer, whose presence and accessibility is becoming ever prevalent, the work of this thesis might well be a step toward opening the currently unknown artistic facets of many people, whether inside or outside of an academic environment. While the addition of drawing instruction will not guarantee better designs or engineering schematics, or even a bolstered interest to draw [2], this learned ability will still always be an asset in communication, expression, and problem solving [1].

3. SKETCH RECOGNITION

The discipline of sketch recognition employs the modalities and techniques that allow a computer application to identify hand-drawn shapes that a person creates. This is in addition to the normal scenario of a graphics application simply rendering drawn strokes to the screen without any interpretation. Input can include a handheld stylus whose movement is digitized across a surface or a sketch on paper whose strokes are obtained from a scanned-in image. The common digitizing surfaces are specially built monitors, tablet PCs, and USB-attached tablet pads. Applications are possible in any domain where the affordance of drawing lowers the barrier of a human to interact with the computer as a tool. A few such domains include design, modeling, military planning, mathematics, and education.

3.1 Understanding Digitized Strokes

As a user moves a stylus or other input device against a digitizing surface, the movement is sampled as points with temporal and spatial data (Figure 3). These points then represent the strokes drawn by the user and are analyzed to determine what form they embody. Sketch recognition employs one or more of three techniques to understand these strokes: gesture recognition, vision-based recognition, and geometric recognition.

Figure 3. Example of a digitized stroke. When a curve is drawn on a digitizing surface (left), the movement is sampled as data points with coordinates and a time (right).

Gesture recognition relies on how the strokes are drawn, making it dependent on attributes like the order in which strokes are drawn and their direction [14]. An example of where this form of recognition is employed is in personal handheld devices that use a stylus to input letters of an alphabet, such as Palm Graffiti (Palm 2009). Gesture recognition may require a training period for the user as s/he must learn the gestures. Gesture recognition can also be dependent on the drawing style of the user, requiring a similar training period for the system to learn how the user draws the gestures.

Vision-based sketch recognition relies on the look of the shape and is generally inflexible to the variance in how a shape is drawn because [15, 16]. This form of sketch recognition can be used when the drawing is done offline and then brought into a computer, such as when a paper sketch is scanned in. Like other computer vision techniques, the stroke can be analyzed on a pixel-by-pixel level instead of using the sample points.

Geometric recognition defines shapes by their geometric features and constraints [17-19]. For example, an arrow might be described as three lines, their lengths, and the angles between them. This approach allows for more flexibility as the user draws since it is not dependent on what order strokes are drawn in or their direction.

Sketch recognition also has two levels of functionality. One is the lower level recognition of primitive shapes [14, 18-20], such as lines, curves, circles, and rectangles. The second is the higher level recognition of the symbols and forms of a particular domain that are defined using domain-knowledge and are made up of primitive shapes [17, 21], such as UML diagrams and military symbols. As a primitive or shape is recognized, it may also undergo "beautification," if the domain allows, where the hand drawn strokes are manipulated or replaced by a cleaner or ideal representation (Figure 4).

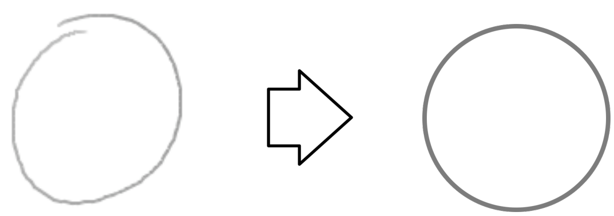


Figure 4. Example of sketch recognition and beautification. A person draws a circle and the hand drawn stroke is replaced by a representation of a perfect circle.

The methodology outlined in this thesis makes use of all three recognition types and employs a high-level recognition approach. Gestures let the user explicitly mark strokes as reference lines, erase strokes, and undo strokes s/he has drawn. Vision-based and geometric recognition are used to compare the user's drawing to that of an example template generated from the reference imagery. Strokes made by the user are collectively assessed as a group of strokes/primitives on each step while s/he constructs his/her rendition of the face. Further discussion of the sketch recognition techniques employed in this thesis is outlined in section 6.

3.2 Related Works

Research work by Alvarado [22] sought to define the proper techniques and conceptual model for a sketch recognition user interface, or "SkRUI," as the author coined. Key findings were as such: (1) Do not give feedback until the user explicitly says s/he is done sketching; (2) Do provide an obvious indication of recognition mode and free-sketch mode; (3) Only recognize one explicit domain of shapes at a time as automatic domain detection is not available; (4) Do have pen-based editing for copy, paste, etc.; (5) Do support distinct editing gestures and sketching gestures without a modal switch; (6) Do use large buttons; (7) Always have the pen respond in real-time. All seven principles were effectively used, though some of the actions implied by (4) were excluded as the author adhered to the draw-and-erase conceptual model found in a normal pen and paper situation.

Arvo and Novins [23] sought to introduce a new paradigm for free sketching that beautified a shape while in the act of drawing. Sketch recognition occurred as soon as possible after a pen-down action and then a user's stroke was gradually fitted to its ideal representation. This provided a different type of corrective feedback as, once identified and fitting has begun, the user should understand how to complete the shape. Such feedback has a different intent than the work in this thesis and is mathematically limited to certain simple shapes.

Bae [24] created an interactive application called "ILoveSketch" that borrowed many sketching affordances, such as ink that could "dry," for drawing curves in a 3D space to produce a model. Interesting features included automatic stroke merging, pen-based perspective camera movement, and automatic paper rotation. Created with the professional designer in mind, ILoveSketch was not meant for one with weaker drawing skills and did not understand the sketch, making it unable to provide corrective feedback on how to improve the user's sketch.

Another older work [25] that is still quick relevant was done by some of those in the same group as the latter [24]. By preprocessing example imagery, Tsang created a 3D sketching application that provided a suggestive interface for fitting, closing, excluding, and general manipulation of user strokes. This application moved strokes toward the representation found in an example image unless parts were "glued" or "pinned" by the user. Strokes were also suggested to the user based on if his/her stroke could be closed, extruded, or matched a previously drawn stroke from a database. Also intended as a design tool for the professional, a product of this application was meant as a derivative work and an not a rendition of the example imagery given, nor was the tool meant to improve the user's ability but play off ability that already existed.

Fan [26] introduced a framework that assists teachers by automating the assessment of student sketches of diagrams for biology, chemistry, computer science, and other related fields. The project eliminated the need for handwriting or sketch recognition by reducing the drawing vocabulary of the user to a limited set of primitives that are accessible from a toolbar. A teacher was responsible for using the application to generate the correct and incorrect versions of the diagram s/he wished the students to create, and these were used to assess student sketches. The methodology presented in this thesis gives the student much greater freedom in drawing and reduces the burden placed on the teacher.

Another work [27, 28] created a heuristics engine that verified a user's sketch for a floor plan, providing text feedback as needed. The sketch recognition used was limited as rooms were boxes or blobs and doors and windows were represented as lines. Users could also label rooms and features with text. The application then ran through its heuristics to verify that certain architectural features met proper criteria (e.g. "the Emergency Room should be close to the Entrance Lobby"). These heuristics were manually entered by an expert and were defined in a simple tuple form: (<requirement> <room> <room>). For example: (SHOULD-BE-ADJACENT EMERGENCY-ROOM INTENSE-CARE-UNIT). This allowed for text feedback when a heuristic was violated. The sketch could also be converted into a simple walkthrough 3D model using VRML. While this work does provided feedback to improve a user's sketch, this thesis presents techniques for such an application to also give visual feedback and on more complex shapes.

Plimmer (2008) created a collaborative teacher-student system for visually impaired children to learn handwriting and their signatures through haptic and auditory feedback. Their system, called "McSig," allowed a teacher to sit at a tablet PC and draw a shape or create a stencil that the student was to draw. A corresponding device then guided the visually impaired child along the stroke using the multimodal feedback of Dutch paper (a surface that raises when drawn on) and an auditory pitch based on position. The most striking parallel between this work and that of Plimmer's is that they both help the user "see" the shape s/he is to draw by manipulating the source imagery so it can be envisioned. Both works also share the notion that the stroke must belong to the user and not the system as Plimmer reported that interaction with a device that physically moved the hand in order to teach the movement was not received well by non-sight participants.

Taele [29] developed a computer-assisted language learning application called Hashigo that taught how to properly write Kanji, or Chinese characters that are used in the Japanese writing system. This application improved a student's proper Kanji comprehension by verifying for technique (e.g. stroke order, stroke direction) in additional to visual correctness, highlighting sketches that were correct as feedback. All characters exist as a series of geometric constraints, which is one way this work differs from that of this thesis as what the user should draw is very rigidly defined. This work also pertains to line drawings and the user's ability to remember their sequence, versus perspective drawings and the user's ability to learn universal techniques.

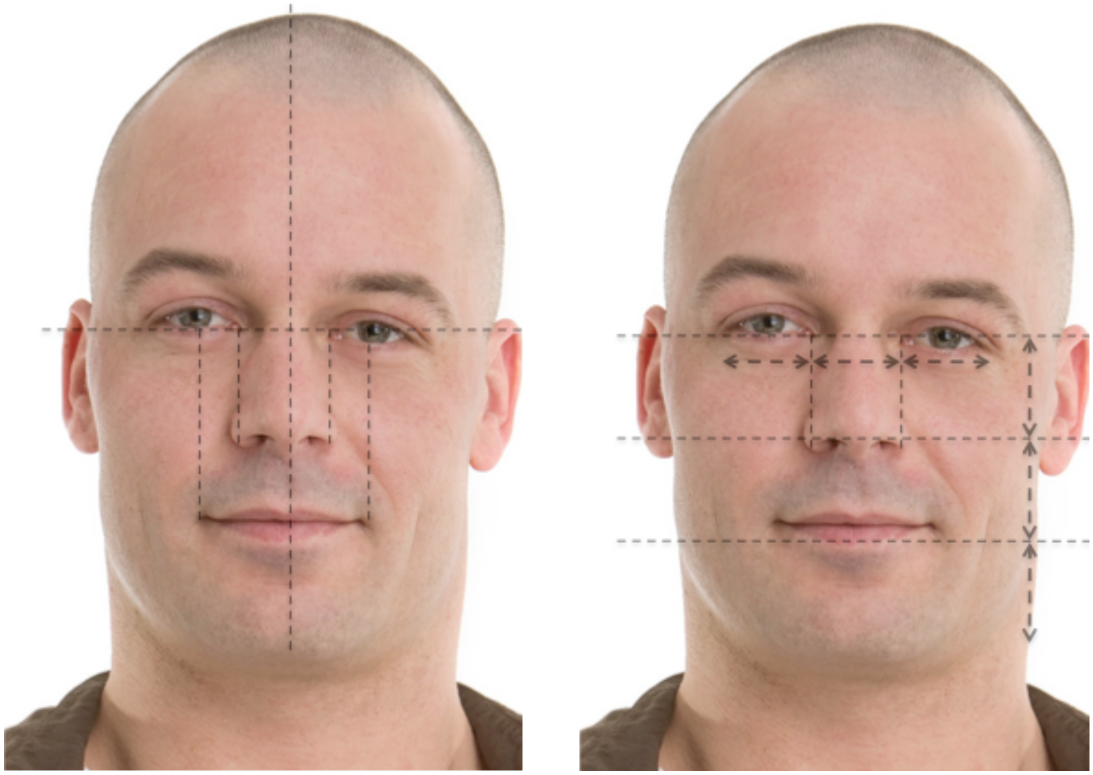
4. A METHODOLOGY FOR ASSISTIVE SKETCH RECOGNITION

With the prior literature review and presentation of the sketch recognition discipline, this section now opens the research problems that exist when merging art instruction from a computer application and art creation from a human user. Given the challenges of the artistic domain, many questions arise when designing an application that is meant to improve a person's ability to draw. What does the person learn to draw? How is the person instructed on what to draw? How does the application obtain understanding of the drawing subject? When is a person done with a drawing task? How does the application evaluate what is correct? When does the application provide help? How does the application know if the result could be considered a good drawing? These are the research problems explored by the author and answered through the methodology presented in this thesis.

Overviews of the two iterations of development for the methodology are described in the first two subsections, withholding in-depth discussion of the techniques for interaction and recognition until latter sections. The subsequent subsection presents the nine design principles for assistive sketch recognition learned from user studies that make this methodology effective. The last subsection recounts an exploration of ideas and concepts that took place between the two iterations of the development.

4.1 Learning by Drawing the Human Head

Empirical evidence suggests that the human form is a common centerpiece to many works of art, whether a stick figure done in crayon or a painted portrait of a human model. The author of this thesis found that imagery of the human head could serve as a good basis to build a methodology on, to assist the act of drawing with, and to evaluate sketches against. First, the student learning to draw usually desires to achieve realism or the ability to make his/her sketch look right before venturing elsewhere in his/her ability [3, 6]. The author proposes that realistically sketching an actual person from a photo is a task many would like to achieve. Second, the human head is seen everyday yet most will make common mistakes when sketching it. For example, many draw the eyes much larger than they actually should be, the ears smaller than they should be, etc. This means that any improvement in drawing ability should be noticeable. Third, step-by-step instructions, corrective feedback, and evaluative metrics can be created from the universal proportions shared by the features on the human face (e.g. sides of the mouth should fall in line with the pupils; Figure 5) [3, 8]. Lastly, positional data for the facial features of the human model in an image or a drawing can be determined using face recognition techniques, thereby generating a template for the sketch recognition to work with.



(a) **(b)**

Figure 5. Example of universal proportions found in the human face. (a) The face has left/right symmetry. The eyes are generally in the middle of the head. The width of the mouth and the pupils line up. The inside of the nostrils and the inside of the eyes line up. (b) The width between the eyes is the width of an eye itself. The eye line, tip of the nose, and bottom of the mouth will divide the lower half of the face into thirds.

The result of these four properties is the ability to create an end-to-end methodology for producing an application with step-by-step instructions that leads a user to his/her rendition of a reference headshot of a human model. The rest of this thesis outlines the author's proof of concept for such an application (Figure 6) along with the principles and techniques learned in the process. In this application, each step gives a written explanation describing the task and a series of questions for the user to self-check his/her work. Each step also provides some manipulation of the example image to assist the user in seeing what to draw. The user can freely draw but also has tools and gestures for drawing, erasing, and marking strokes. When the user is satisfied with an attempt at drawing the task of a step, s/he explicitly tells the application to check the work. If the attempt fails to meet constraints against an underlying example template generated prior through face recognition on the image, text remarks from the application inform the user and s/he can elect for visual feedback (if available) to see adjustments that should be made or reattempt without it. If the user's attempt succeeds, s/he is given congratulatory feedback and can advance to the next step. The user can also advance a step if s/he chooses to ignore the corrective suggestions. The final drawing is then evaluated by using face recognition to generate its representative template and comparing this to the one gained from the reference imagery.

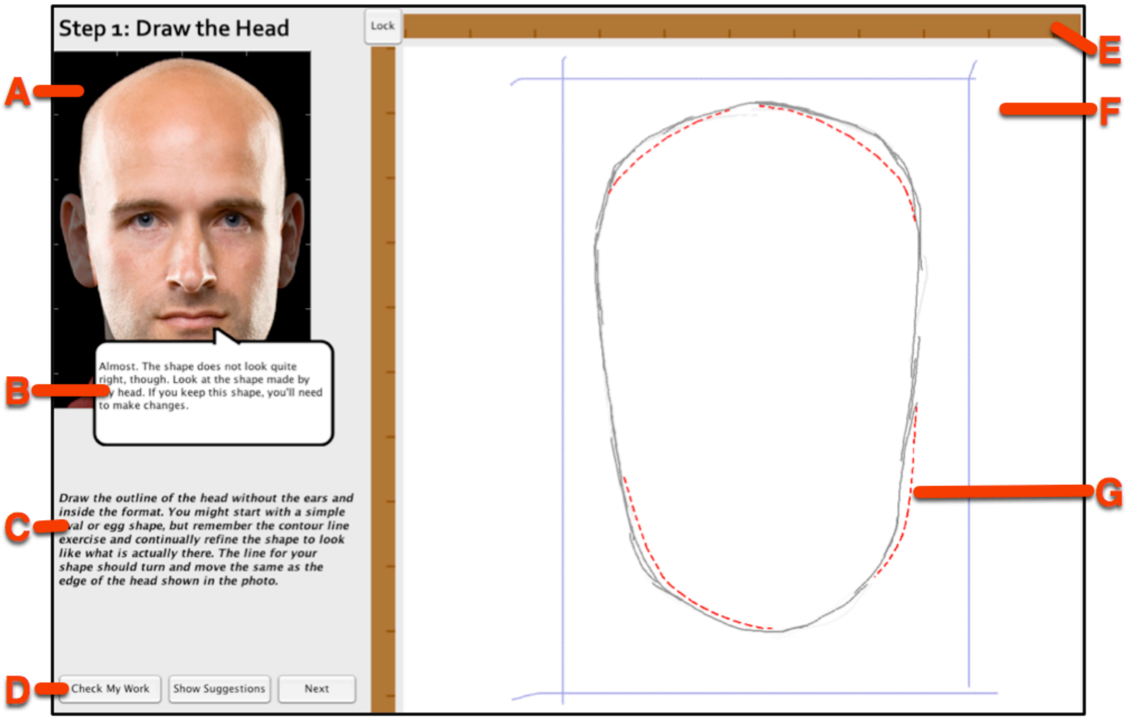


Figure 6. Screenshot demonstrating most of the interaction techniques. (A) The reference image manipulated to show the head. (B) Text feedback. (C) Instructions for the step. (D) Options available to the user. (E) Straightedge tool. (F) Drawing area. (G) Visual feedback on the drawing.

4.2 A Prior Attempt Using Cartoon Characters

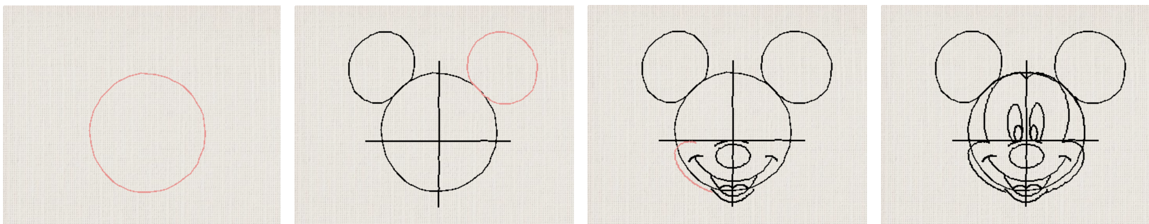


Figure 7. Selections of the example template in the first version. These were shown to the user in steps as a guide for drawing.

The first iteration of this work was built around the idea of teaching people to draw iconic cartoon characters (Figure 7). This version also had step-by-step instructions and introduced using corrective feedback to lead the user to his/her rendition of the cartoon character. Each step of the exercise included a concurrently built example image based a hand drawn template (Figure 8) and a written explanation of what, why, and/or how to draw the new piece. When a piece from the example template should have been drawn differently by the user, corrective feedback explained and showed the adjustments to be made. When a piece was drawn well, congratulatory feedback greeted the user (e.g. "Looks good!" "Well done!") and s/he was automatically advanced to the next step of the exercise. The underlying notion was that such example imagery would be enjoyable to the user as anyone from a child's age and up enjoys the sight of iconic figures, and step-by-step guides for the construction of such cartoon characters already existed and could be expanded upon. However, this approach failed to address many important methodological issues, such as how an example template could be properly obtained without error for sketch recognition to work with. The subject matter itself, while enjoyable to create, was also not a good example for learning the ability to draw realistic. As proof, the size and placement on the head of the character's ears could vary greatly from that seen in the example and he still be recognizable and aesthetically pleasing. However, any large amounts of variation in a drawing from the ears of a human model in a reference photograph would obviously be noticed as incorrect as proportions are much more vital. Thus, the author went back to the drawing board.

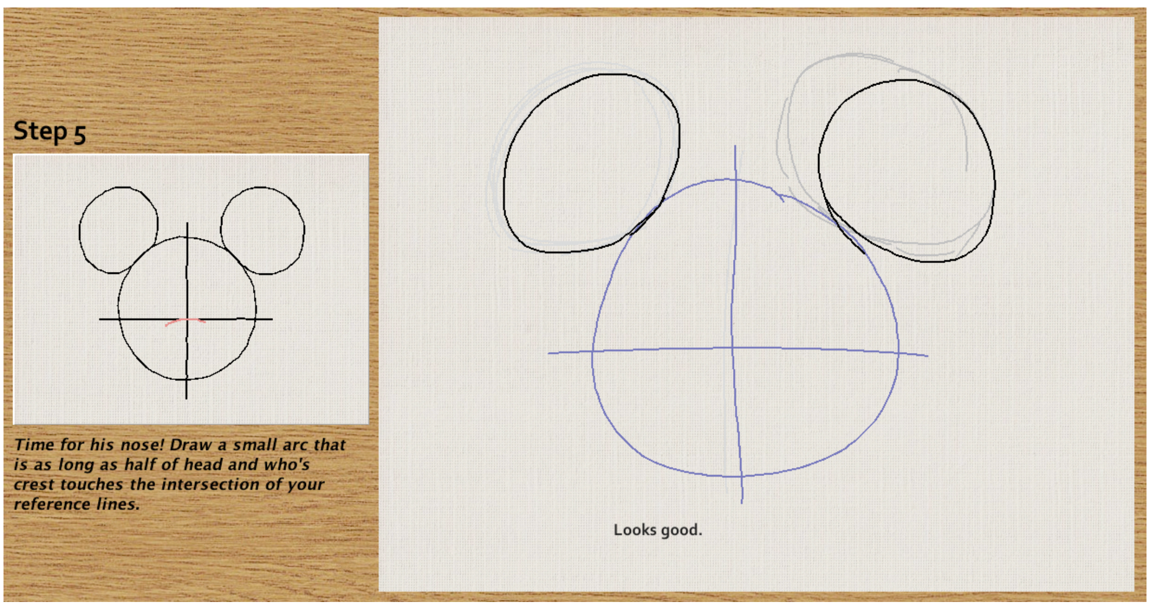
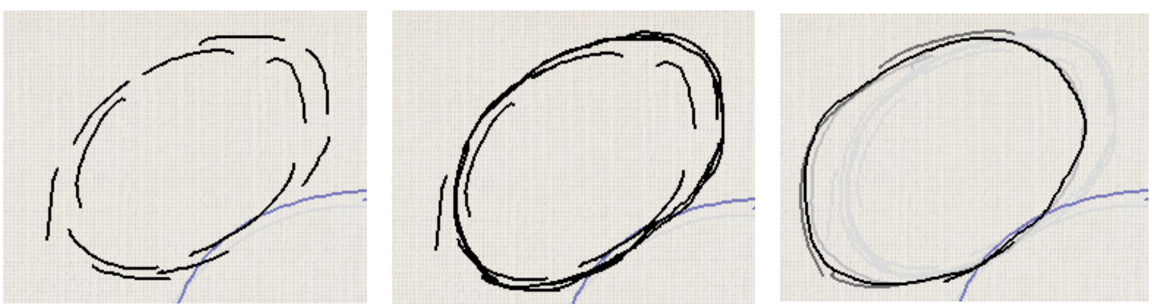


Figure 8. User interface of the application from the first iteration. An example image with instructions is on the left and the drawing area is on the right.

This early version adopted interaction techniques that were also not effective. On the predecessor of this application (i.e. a sheet of paper), strokes can be erased, are not beautified as they could be in a sketch recognition application, and can be traced over to emboss or emphasize them. Adopting this model, the first two properties were combined to allow for "prep-sketching" (Figure 9.a) and for "beautifying" (Figure 9.c) the drawing by removing strokes that would normally be erased. The third property of over-tracing a shape was used as, what the author thought would be, a novel way to cue the functionality of the application (Figure 9.b). While prep-sketching allowed the user to mock up the look of the shape with small strokes before committing to it by over-tracing them (i.e. drawing the correct shape in full), user studies yielded that it was not a universally intuitive approach. This over-tracing technique was also not an effective way for the user to make requests of the application. These human-computer interaction techniques are discussed further in Section 5.



(a) (b) (c)

Figure 9. Example of interaction techniques for first iteration. (a) An example of "prep-sketching" for a circular shape. (b) An example of over-tracing or embossing a shape when satisfied with it. (c) Automatically cleaning up unneeded strokes for the user when an attempt passes all constraints.

The look of the corrective visual feedback also adversely affected the user experience in this version. A sample of the visual guidelines is shown in Figure 10, using animated lines to suggest a wider shape and a dashed line to suggest where to make adjustment. Further discussion on corrective feedback is covered in the Section 5.

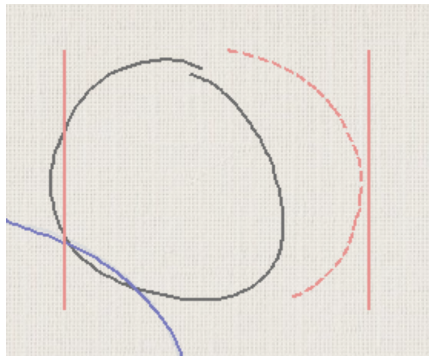


Figure 10. Example of corrective feedback in the first iteration. Visual guidelines of the corrective feedback tell the user to make his shape (the gray circle) wider. The accompanying textual remark is, "Almost, this shape will need to be wider. Redraw the shape with the adjustments."

4.3 Design Principles of This Methodology

After an iteration of development, the resulting application was evaluated through a user study. Six participants performed the first user study, each in one sitting, for the first version of the application (Appendix A). Five participants tested the second version of the application in another two-part, two-day user study (Appendix B; a control group is shown in Appendix C). The purpose of these studies was to determine what effect use of the application would have after short-term exposure and how participants would respond to the application's assistive drawing exercises. As one result of the studies, nine important design principles for assisting the act of drawing using sketch recognition were developed (Table 2) and detailed here. Further discussion on the user experience for each study is detailed in Section 7.

Table 2. The nine design principles gained from the user studies.

1. A formal process must exist to determine error in the example template used by sketch recognition.
2. To limit the user's subjectivity when perceiving an object, the application should help him/her process and "see" it.
3. When creating subjective content, the user should explicitly trigger feedback so the application does not interrupt or thwart his/her though process.
4. Corrective feedback must clearly show what is being asked for and why it is being asked.
5. Previous or "erased" strokes must be temporarily visible as a form of corrective feedback.
6. Sketch recognition must be adaptive to some deviation from the example template.
7. A larger and properly primed drawing area is advantageous to the user creating subjective content.
8. Using additional image analysis on the reference photo to gain information for sketch recognition is advantageous.
9. Sketch recognition should incorporate other drawing principles such as symmetry and light/dark values.

1. A formal process for template creation. The first guideline stems from the ability of a teacher to draw an example template with error, thereby leading the student astray or to frustration as s/he is incorrectly corrected by the sketch recognition. The error could be any number of items: a wrong proportion, a lack of symmetry, missing strokes, etc. To avoid the need to tolerate such error, a formal process should exist to ensure that the exercise is correct. This process might be as simple as having another teacher verify the exercise. It might also be done systematically, by having the teacher draw repeated versions of the example template for use by recognition, or through creation of the example template through image-based analysis (e.g. face recognition). The teacher might be further assisted with the addition of CAD tools for drawing or selective beautification on desired strokes.

2. Help the user to "see" the reference object. To limit the range of the user's subjectivity when perceiving an object to draw, the application should help him/her process and "see" the reference content correctly. Looking at any of the freehand exercises in Appendices A, B or C, one can see the need for this when eyes are drawn too big, proportions of a head are not correct, or a character is not symmetric. The user needs to perceive the edges, spaces, and relationships that make up the object, instead of seeing the eyes, nose, or mouth, for example [3]. The first version used some techniques for the R-mode, such as flipping the reference object upside down, and the second version manipulates the image to highlight the facial feature asked for. The second user study also began with a pure contour drawing (introduced in Section 4.4) as a warm-up exercise for the R-mode. Addressing this guideline is even more important, given that the act of drawing while in R-mode has been described as a complete mindset shift where one will most likely lose track of time and usual thought processes [3]. Everything about the application should be complimentary to this shift.

3. Sketch recognition should be triggered by the user. With subjective content, what the application asks for or expects and what the user creates can be different. This can cause moments where the application is looking for one type of shape while the user is waiting, believing s/he has already drawn what the application has asked for. Conversely, the application could interrupt the user when not desired and therefore adversely contribute to the user experience. Lastly, a complete shape does not necessarily mean the user considers the task done as drawing can be a very iterative process of draw, erase, look, and repeat. An appropriate solution is to have the user explicitly inform the application when s/he is done with an attempt at a task.

4. Corrective feedback should be consistent and clear. Conveying to the user what needs to be corrected must be done properly. One simple rule is that the corrective feedback should be relative to the task, and not the current attempt of the task. For example, if the user drew the outline of a head in response to the instructions and it was wrong, the corrective feedback shown should be relative to the example template's piece for the head and not the user's strokes. Otherwise, when s/he redraws the shape, the corrective visual feedback will suggest different proportions and placement than originally asked for.

More importantly, some users expressed frustration when they did not understand the corrective visual feedback or know what the application was asking for, such as animated lines motioning to make the shape wider or thinner. This principle must work in tandem with Principle 2 by helping the user "see" what is wrong, as well as Principle 6 by providing clear feedback while allowing for some natural deviation from the template.

5. Previous strokes should persist as feedback. When informing the user that an attempt for a drawing task is incorrect, his/her previous strokes are the best source for preventing a repeat of the same mistake. This is very similar to how strokes erased by a pencil on a piece of paper will still persist faintly and serve as a reference for a previous attempt.

6. Sketch recognition should allow for natural deviation. With subjective content, it is the user's right to reject the application's corrective feedback and continue on. Since it is also very difficult for someone to produce a direct copy of a reference image, natural deviation from the example template should be allowed and supported by sketch recognition. However, there is a fine line between allowing a user's shape to pass constraints even through it should still be drawn slightly different and asking the user to continuously repeat his drawing of the shape until it is correct. If constraints are too tight and yet, to the user, his/her shape looks right visually on each attempt, this causes frustration and inhibits learning. On the other hand, when the user is constructing a sketch from an example template, each piece in the sequence of steps will have a spatial relationship to another that must be honored. This is why the author found it important to allow the user to work on any facial feature while on any step of the exercise, as a person will sometimes recognize a discrepancy and attempt a fix it after new features have been drawn.

7. A larger and properly primed drawing area helps the user. Given that proportions are one of the first thing the beginner struggles with, many users will fail to gauge proper sizes while laying down earlier pieces, thereby possibly causing their latter pieces to run off the page. The approach adopted here is to instruct the user to draw a "format" [3, 8] or a viewfinder on screen, thereby creating a buffer between where s/he'll draw and the edge of the paper. Another approach could be to inform the user when s/he is drawing base pieces that his/her proportions will be too large for the remaining parts of the exercise. Simpler approaches to compensate for this are using the application where the most screen real estate is available, or allowing for scaling or moving of the drawing to provide more space as needed.

8. Additional image-based analysis is advantageous. For the sketch recognition to best assist the drawer, it must have as much information to work with as possible. Employing face recognition technologies is one approach. For the purposes of sketch recognition, it provides a way to understand what is in the image and a simple enough template that a student can be helped to establish proportions in a step-by-step fashion. However, this approach falls short when assisting the student with the look of his/her sketch. To determine if the nose in the drawing really is the nose found in the photograph will require additional image-based analysis techniques to interpret the user's strokes and expand the paradigm for corrective feedback.

9. Sketch recognition should incorporate other drawing principles. While establishing proportions and placement are first steps to successful drawing, the student also needs to learn and have guidance on other key aspects of drawing. The step-by-step guide presented by the application in this work did teach how to lay down reference lines and setup a drawing, with some commentary on symmetry. However, the drawing task is made easier by having a fairly straight on view of the face. Further instruction could be added on maintaining proper proportions when the face is slightly tilt, looking off to the side, or each in a profile view. Discussion and assessment of light and dark values made by different pencil types, pen pressure, and/or crosshatching would literally add another dimension to the user's drawing.

4.4 An Exploration in Using Reference Photography

The second iteration of development began as an exploration of mock-up screens for how assistive sketch recognition could function using any reference photography as example imagery instead of line drawings and cartoon characters. The underlying notion was that drawing from reference photography would be closer to the act of drawing a real world object, involving the same aspects and interplay with human perception. As Edwards states, “The key to drawing is to set the proper conditions to look” [3]. The exercise shown as a mock-up in Figure 11 was meant to make the user’s perception aware of the details. Studied by Pariser [10] and used heavily in classes taught by Edwards, a contour drawing exercise is often used was a warm-up to drawing and has many variations. To begin such an exercise, the drawer will choose an object and attempt to reproduce the contour, or edges, and lines that the object consists of. Edwards called this a “pure contour drawing exercise” where the student attentively and intently will draw the lines of his/her non-dominant hand for a set amount of time without looking at the paper. In the mock-up, the user would spend a full five minutes to slowly move his/her hand to draw at the same rate that his/her eyes moved along the lines of the hands in the picture. The strokes would not appear until the five minutes were up. Such an exercise is meant to shift the participant into a state of mind primed for drawing, or R-mode. This mock-up became a warm-up exercise for the second user study.

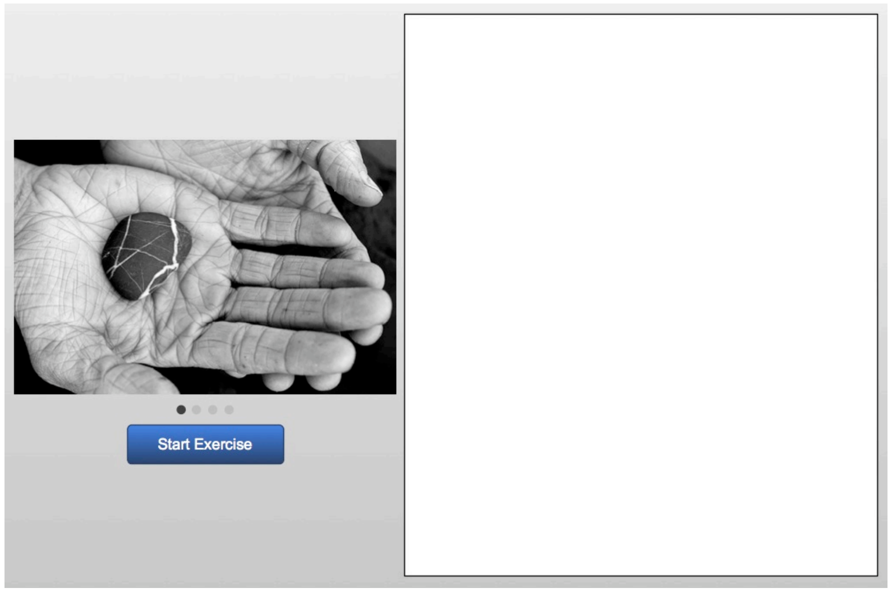


Figure 11. Mock-up screen for a blind contour drawing exercise. This would have been a warm-up exercise. A version of this was used in the second user study.

The next mock-up screen (Figure 12) demonstrates a proposed drawing mode, which has many concepts that were used in the second version of the application. One concept in particular is the horizontal and vertical straightedge tools that the user can drag onto the stage to create straight lines. Another important concept is the buttons for the user to explicitly prompt for suggestions and advancement of the current step. These features are discussed more in Section 5.

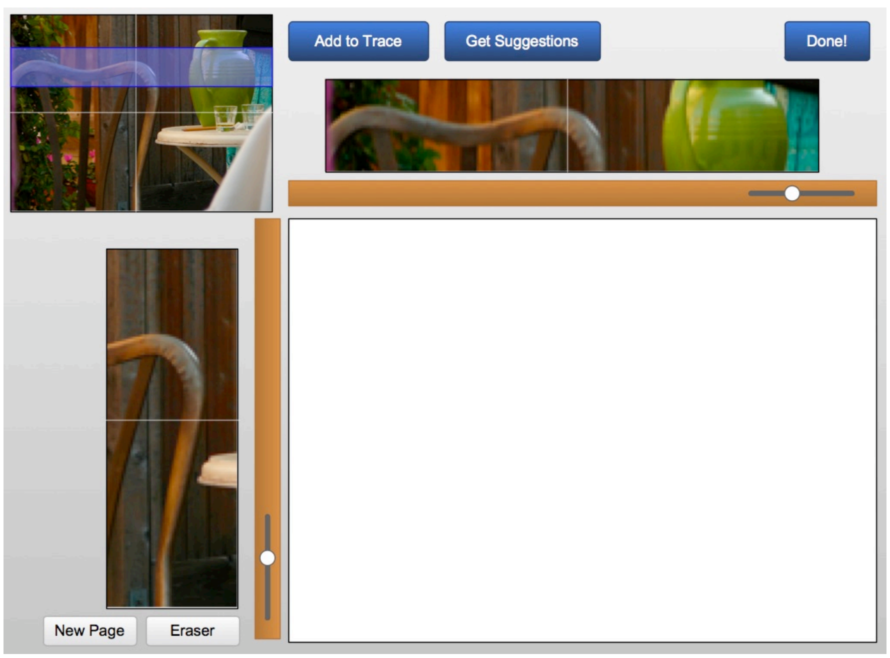


Figure 12. Mock-up of a proposed drawing mode. This screen has many concepts that made it to the final version.

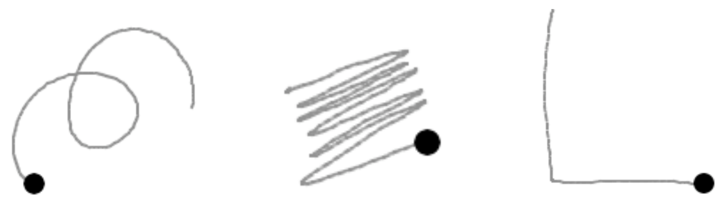
This mock-up (Figure 12) also introduces a first attempt at manipulating the example imagery to change how the user will see it. The full image is shown in the top left corner but a horizontal and vertical segment of it are also seen. Instead of seeing a whole object in the image (e.g. the chair), these segments help the user see only part of an object (e.g. the top portion of the chair's back rest) so s/he can focus on the primitives, spaces, and relationships that make up that part. A click-and-drag motion across either segment with the pen stylus would slide its view accordingly. The underlying notion was that such a view helps the user move this drawing task from a L-mode activity (i.e. "I see a chair") to a R-mode one (i.e. "I see two parallel curves"). Some art students use a similar technique when drawing or painting by using another piece of paper to cover over portions of their reference material [3]. The author also speculated that the two elongated segments would help the user establish relationships with an object's parts along the corresponding horizontal or vertical axis. This exploration yielded new ideas but many of the research questions posed in the beginning of this section could still not be answered until the methodology worked around human faces.

5. HUMAN-COMPUTER INTERACTION FOR DRAWING

Having defined a methodology for assistive sketch recognition in the prior section, the discussion now moves to cover the development of the user interface components and interaction techniques that supplement this methodology. The user interface consists of a drawing area, a reference image to draw, and an area to provide instructions and options to the user. Each of these plays a part when providing corrective feedback to the user.

5.1 Providing an Area to Draw

The predecessor to the display for this kind of application is a piece of paper, whose visual and aesthetic affordances can be replicated. The author did not concern this work with the tactile qualities of paper under the assumption that development and testing would be done on an existing digitizing surface (i.e. a Wacom monitor or tablet PC). The drawing area supports these interaction techniques: draw a freehand stroke; draw a stroke with a straight-edge tool (Figure 6); mark a stroke as a reference line with a right angle gesture (Figure 13.a); erase a stroke with a scribble gesture across it (Figure 13.b); undo any of these actions with a rollback gesture (Figure 13.c).



(a) (b) (c)

Figure 13. Supported gestures on the drawing area. (a) Rollback for undo. (b) Scribble for erase. (c) Right angle to mark reference line.

The straightedge tool is required because some of the instructional steps ask the user to draw reference lines for use in constructing proportions on the drawing. The first user study found that users could have difficulty drawing a straight line over an extended length, which could affect recognition and the placement of strokes on subsequent steps. To rectify this, the straightedge tool was added for a user to drag into the drawing area and use to create straight horizontal and vertical lines, mimicking the affordance of a pencil against a ruler. This tool also helps to eliminate unneeded deviation from the example template that's referred to in Principle 6 (Table 2). Lines drawn with the straight-edge tool are automatically marked as reference lines by a blue color, distinguishing them from the user's strokes intended for his/her final sketch.

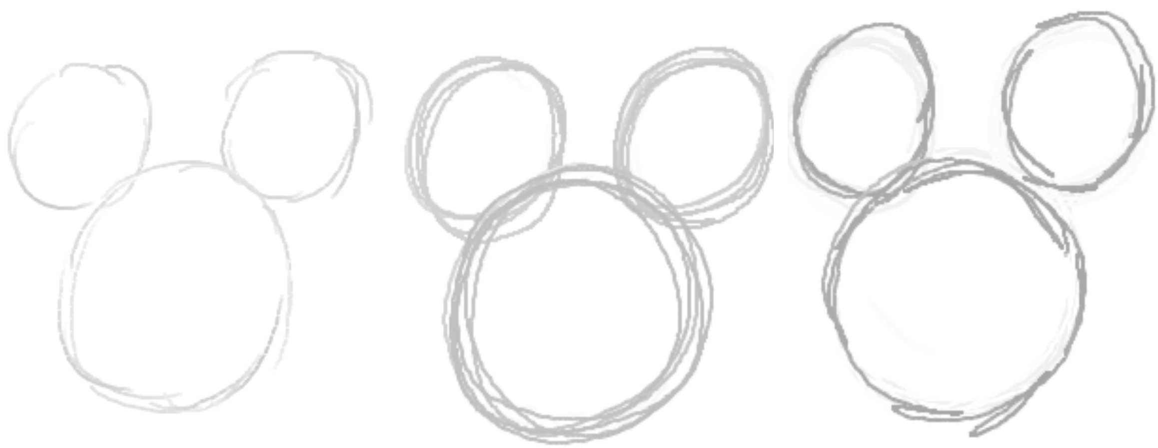
The user can explicitly mark the last non-reference stroke drawn as a reference line or mark with a right angle gesture (Figure 13.c), also changing the stroke to blue. Originally, this gesture was defined as a flick of the pen stylus on the surface, but the author found that some user study participants were unsuccessful in triggering the gesture. The right-angle gesture is the replacement because this shape is not generally found in the human face and it conceptually relates the rigid grid work of reference lines created by the user. This gesture might also be used to beautify references lines to a horizontal or vertical axis, thereby replacing the straightedge tool, but this interaction technique was not studied for this thesis. Reference lines can also be locked to prevent an accidental erase or hidden so the user can cleanly see the outcome of his/her drawing effort.

The remaining two gestures are a rollback gesture (Figure 13.a) for an undo action [24] and a scribble-out gesture (Figure 13.b) for an erase strokes. Strokes that are erased are not immediately removed as they still serve as a reference for the user after s/he has received corrective feedback (Principle 5). For the first version, erased strokes were gradually faded out and set to a color that varied slightly from the paper background, making them appear similar to lines erased on a piece of paper that leave a faint remnant. For the second version, they were gradually faded out but completely removed after a period of a few minutes to make sure that the drawing area remained clean. All gestures are recognized using low-level geometric recognition [18] on all strokes drawn. If a stroke is not interpreted as a gesture, it will persist on the drawing area as a stroke meant for the stretch.

As stated earlier, a goal is to mimic a normal pen and paper situation. Since paper does not have buttons, the author purposefully sought to avoid the use of any at first and the user interface even had the aesthetics of a desk and a piece of paper (Figure 8). The latter version did incorporate large buttons to gather explicit actions from the user and the aforementioned desk aesthetics can be considered merely a preference in light of the application's overall intent. It is, however, to the advantage of the user to make this drawing area as large as is possible, as formulated by Principle 7 of Table 2 and section 4.2.

Discussed briefly in Section 4.2, three other interaction techniques were tried. These were to use “prep-sketching” to mock-up shapes, to mark embossed strokes as significant, and automatically clean up unneeded strokes. To explain the thought process behind these techniques stemmed from the plausible workflow of a sketch artist.

When drawing a human body, a sketch artist will make light strokes using general shapes to first establish the human form and its proportions. These light strokes could be short reference points to mark out locations, or might be long and continuous, flowing with the hand and stream of thought. “Prep-sketching” was meant to allow the user to mock up the placement and proportions of a shape using short strokes and adjusting his/her mock-up as needed. If the user were to decide his/her mock-up of a circular shape (Figure 9.a) needed to be wider, s/he could simply continue prep-sketching to elongate the shape without erasing strokes or triggering sketch recognition functionality. Strokes that were part of the mock-up would then be removed automatically, as described later. Results from the first user study showed that prep-sketching was not defined to one drawing style nor universally intuitive. While some might use short and light strokes to mock-up a shape (Figure 14.a), others may wish to use over-tracing strokes to mock-up a shape (Figure 14.b), sometimes drawing multiple shapes at a time before committing to any. Another group might slowly draw a shape in short over-traced sections while never lifting the pen (Figure 14.c). Still others may not prep-sketch at all even when explicitly told the action is allowed, either because this is not a drawing technique they are familiar with or because prep-sketching just does not exist in their conceptual model when drawing on a digitizing surface. If prep-sketching is to be supported and become an asset to sketch recognition, the author's conclusion is that an application would need to provide a tutorial to the user on a specific and recognizable style or obtain other contextual information in order to determine if a stroke is meant for mock-up or final display.



(a) (b) (c)

Figure 14. Examples of varying prep-sketching styles. These were the kinds of styles observed during the first user study.

Returning to discussion of the sketch artist, it is his/her natural inclination to emboss, or over-trace, the strokes s/he intends to keep once the desired shape is formed with prep strokes. These can bring out the outer form of the human body from his prep-sketch strokes or accent features like a nose or ear. Similarly, when the user likes a shape, s/he could commit to it by over-tracing the portion of the prep-strokes that s/he desires (Figure 9.b). The author's intent was for embossing strokes to then serve as a trigger for functionality, such as to provide corrective feedback or continue with the drawing lesson. It would also be the point in time to remove any unneeded strokes, as discussed next. While the idea was novel, it did not create a good user experience. As formulated in Principle 3 of Table 2, the user should be the one to explicitly trigger recognition.

During a clean-up pass, a sketch artist will erase the lighter, unneeded strokes used for mock-up. Normally, beautification in a sketch recognition means to automatically adjust a user's input strokes to some ideal representation, such as replacing a drawn shape with a perfect circle (Figure 4) or a domain symbol. While it is desirable to clean up unneeded strokes, all strokes must still be those of the user when s/he is learning to draw. Originally, strokes that would be erased, like prep-sketch strokes and incorrect strokes, were gradually removed as the user continued his work (Figure 9.c). With the lack of necessity for the prior two properties, there became no need to automatically clean up strokes. However, this interaction technique became Principle 5 in Table 2 for strokes explicitly removed by the user and they were found to be a valuable form of feedback.

5.2 Instructing How to Draw and What to Draw

The use case scenario for this methodology is a person learning to draw who wants to work from any reference image of a human model. In order to lead the user in a step-by-step fashion, the instructions were written by the author to be generic and centered around the genre of portrait photography. The author made the instructions based on a literature review [3, 8, 10, 11, 30] and the prior interaction techniques of Section 5.1 were incorporated. The full instructions with example imagery are provided in Appendix D.

When the methodology assisted in drawing cartoon characters, the application was based on the use case scenario of an art teacher kick-starting the exercise. For example, an art teacher could form a lesson on depth perception, symmetry, perspective, etc. by creating one or more example templates for his students to recreate, marking the steps to take. A similar approach is mentioned by Fan [26]. Instructions would have been put into the application by the teacher so s/he could give guidance and tips, while the application itself provided corrective feedback and suggested adjustments to student's drawings based on the example template. For user study purposes, the author of this thesis generated the instructions. A sampling of the instructions from the first version is shown in Table 3.

Table 3. Sampling of instructions from first version.

"Next, look at your pen. How are you holding it? With your fingers near the tip? If you are, try sliding the pen forward to hold it almost near the middle. There are many times when its best to draw more with your arm and less with wrist."

"Now draw a line that splits your circle right down the middle. Most faces are symmetric so this will help with the placement of features, like the nose."

"Draw a large smile that follows the bottom part of the head. It should start and end around the center of the nose."

"Now create the chin. Draw it similarly to the lower part of the mouth. Start and end just above the lower circle for the head."

5.3 Showing What to Draw

Complementary to the written instructions, a visual demonstration will additional show the drawing task to be performed by the user. With each step, the reference image is manipulated to help the user see what to draw (Appendix D). When the step calls for a reference line on the drawing, a corresponding red reference line is shown on the image (Figure 15.a). By default, these reference lines collectively build on the image and persist with the remaining steps. The user can tap the image to see only the last reference line(s) asked for, tap again to remove all lines and view the image, or tap once more to return all lines (Figure 15.c). When the step calls for drawing a facial feature, that feature is highlighted on the image by dimming all others (Figure 15.b). This adjusted view assists the user in focusing on the shape and contours of the facial feature, presenting the feature as it is not normally encountered and making the user less likely to produce a mere symbolic representation for the feature [3]. This is part of setting the proper conditions for a dominant L-mode to reject the task so an R-mode can take control. These manipulations are generated from the template gained during the face recognition stage. Figure 16 shows a mock-up of another concept for highlighting a facial feature for the user to draw. While this concept does provide a zoomed-in view of the feature, it may not be appropriate for wider features such as eyes and ears.

When using cartoon characters, the example template was concurrently built with the steps (e.g. the nose was not shown on the example template until it was asked for in the instructions) and all new additions were highlighted in a light red (Figure 7). This visual demonstration was simple as was the example template itself.

(a) (b) (c)

Figure 15. Example manipulations of the reference photo for assisting the R-mode. (a) Asking the user to drawing a horizontal and vertical reference line. (b) Accentuating the eyes with reference lines still visible. (c) Accentuating the mouth, but the user has toggled off the reference lines built up from this point.

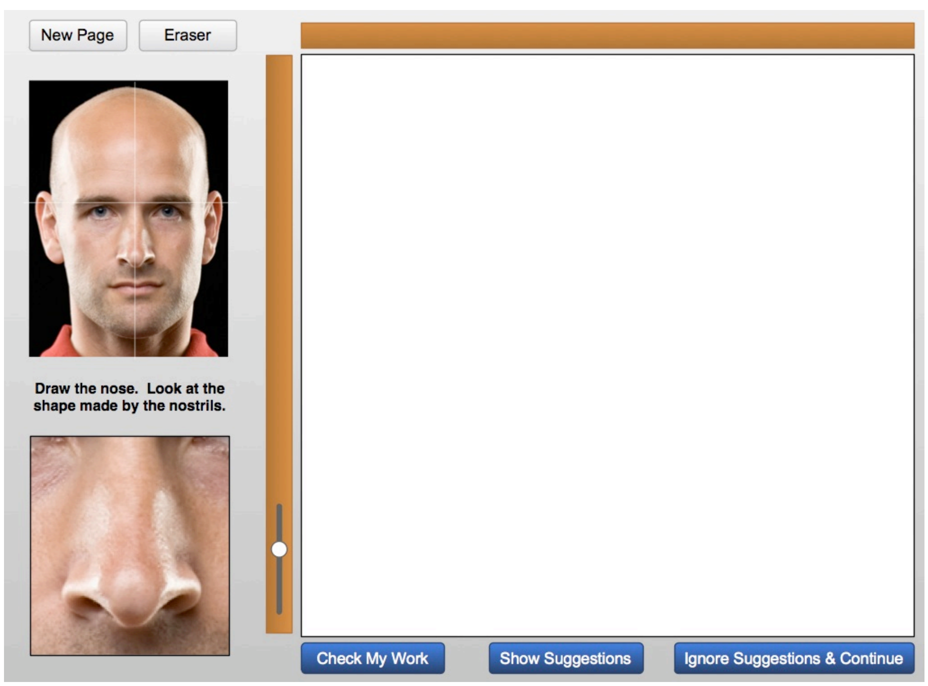


Figure 16. Mock-up screen showing another means of highlighting a facial feature (the nose in the bottom left). This is to assist the R-mode with visual processing.

5.4 Providing Congratulatory and Corrective Feedback

The application provides two forms for feedback: text remarks and visual guidelines. Congratulatory feedback is provided as text. Corrective feedback is provided as text and visual guidelines. There are three explicit options (Figure 17) that the user has for engaging and responding to the feedback: “Check My Work,” “Show Suggestions,” and “Next.”

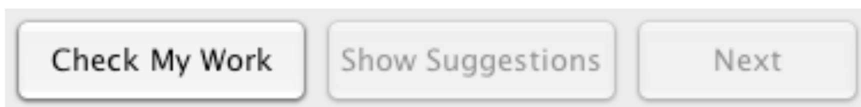


Figure 17. Three options for requesting and responding to feedback. Shown is the initial state of the options at the start of each step, where the latter two are not available yet.

The first option, "Check My Work", performs sketch recognition on the user's drawing and verifies whether what s/he has drawn meets what was asked for in the instructions for that step (all steps are listed in Appendix D). When initiated, corrective feedback is provided within 1-2 seconds as text informing the user of what is wrong or the user is told that his/her attempt is correct and s/he can continue. The first version sought to infer when the user was done with a task by an implied over-tracing of his/her work, but this contributed to a poor user experience. After the initial request to "Check My Work", the second option, "Show Suggestions", becomes available when the user's attempt is incorrect and the application has visual suggestions to make. Otherwise, this option is not available. Visual suggestions are not automatically shown on a "Check My Work" request to first allow the user to make the corrections himself/herself based on the text. The third option, "Next", allows the user to advance to the subsequent step anytime after an initial use of "Check My Work". The prior version prevented step advancement, but would sequentially loosen thresholds, until the user drew as expected. The author found that allowing deviation from the template is beneficial to the user, thus allowing him/her to disagree and to not be thwarted by the application (Principle 6 in Table 2).

Text remarks are only provided whenever the user asks the application to check his/her work and come as a comment bubble from the reference model (Figure 18). If the user's attempt passes all constraints against the example template, the application replies with a complimentary remark and informs the user to advance (Table 4). However, if the constraints fail the user's attempt, the application will first provide text with what is wrong (e.g. “My nose is too wide.”). If the text contains ambiguity in terms of which strokes are being evaluated, the strokes referred to will be temporarily changed to yellow (Figure 18). For example, the text may state, “My right eye is two small,” and the strokes for the right eye of the sketched head, which are on the left of the user, will be temporarily changed to yellow.

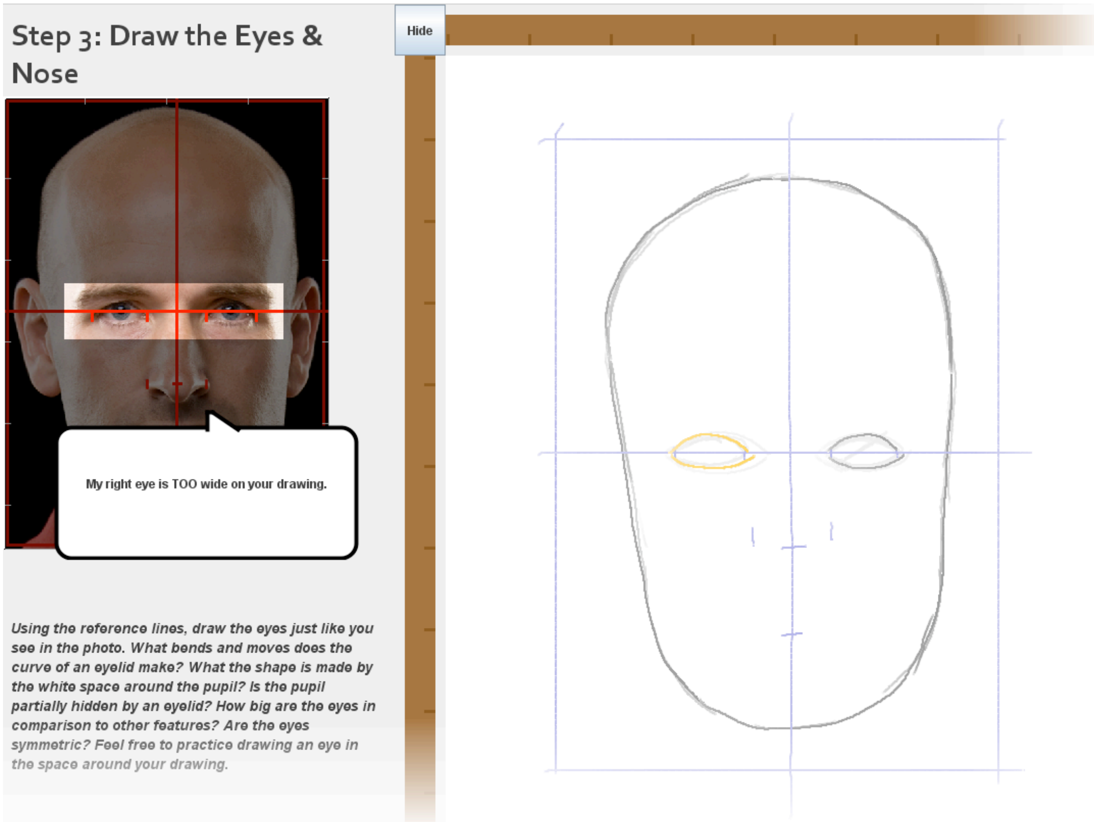


Figure 18. Example of highlighting a shape for ambiguous suggestions. The text remark reads, "My right eye is TOO wide on your drawing," while the eye in question is on the left of the user

Table 4. Sampling of congratulatory feedback when the user has completed a step.

"Looks good! Why don't you 'head' to the next part of this step."
"The eyes look good! Why don't you 'see' what's next?"
"I like my nose! Follow your 'nose' to the next step."
"The mouth makes me happy. Continue on to the next step."

The user may elect to gain additional information from the computer by pressing "Show Suggestions." For a correction to a reference line, this visual guideline would be a red line in the proper location for the reference line on the user's drawing. For a facial feature, the application will generate the proper shape to show in red using the example template (Figure 19), use a copy of the user's strokes that satisfies the constraints (Figure 20), or not provide visual feedback. It is desirable to leave only user-created strokes on the drawing area, so visual guidelines are removed after the user has drawn a single stroke and do not leave a mark on the drawing area like an erased stroke.

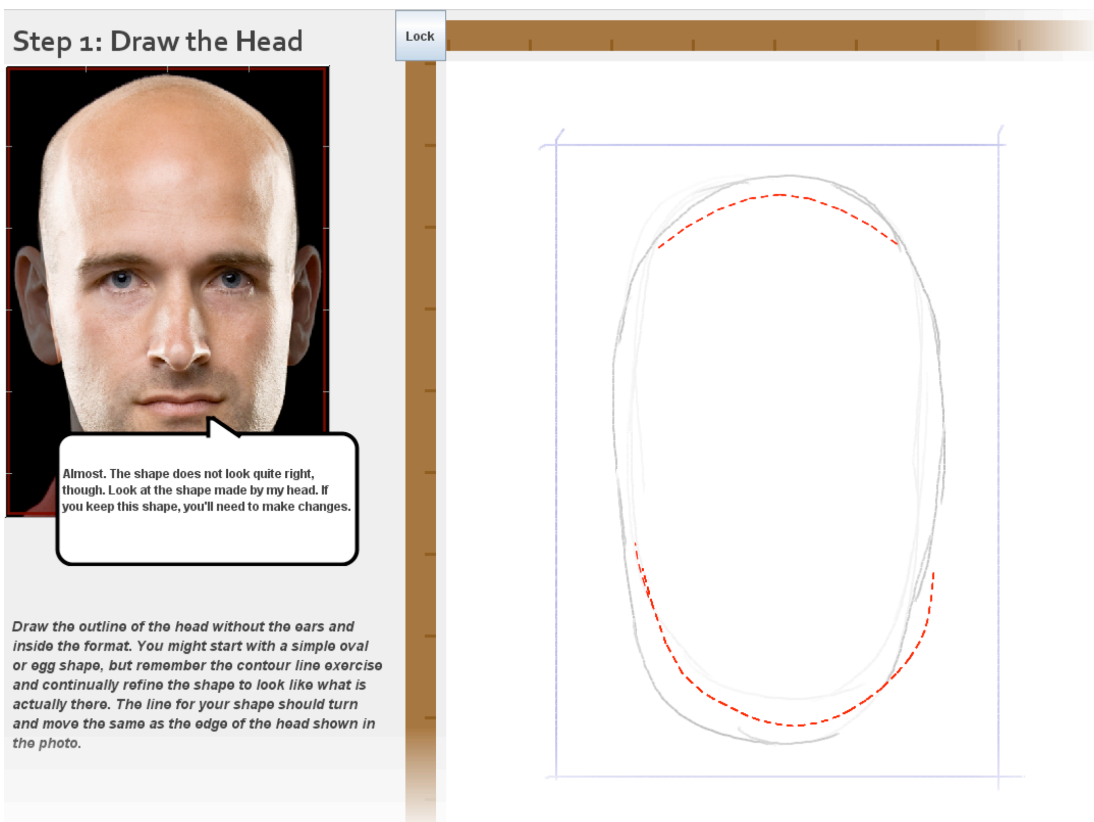


Figure 19. Example of visual guidelines from the example template. The application is displaying the areas to changes based on the template and the text remark reads, "Almost. The shape does not look quite right though. Look at the shape made by my head. If you keep this shape, you'll need to make changes."

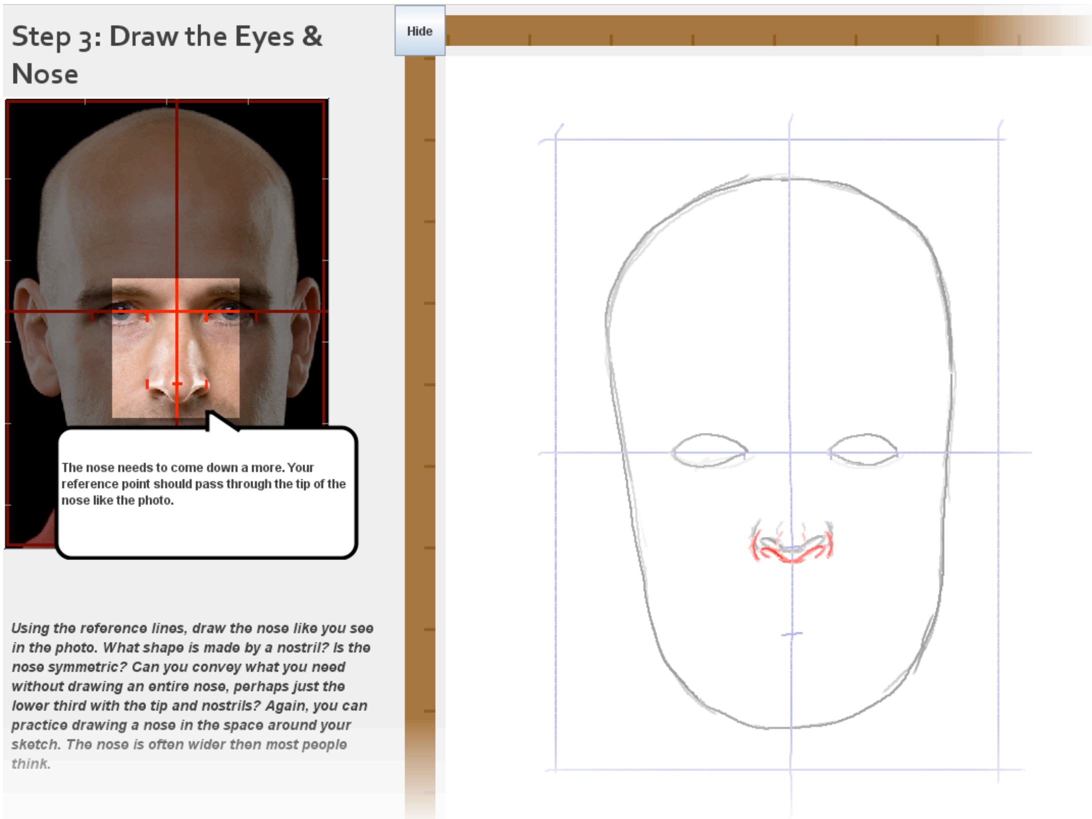


Figure 20. Example of visual guidelines from the user's strokes. The application is displaying a red copy of his/her strokes and the text remark reads, "The nose needs to come down more. Your reference point should pass through the tip of the nose like on the photo."

Determining how the visual guidelines should look was an iterative process, as their presence affects how the user will redraw a shape. A first attempt showed the entire template piece in a light red color, but the author determined that this would prevent the user from learning. Another attempt simply showed tick marks on the middle left, middle right, center top, and center bottom of the bounding box, but the author found that there was an urge to draw the shape passing through these marks, even if it should not. Yet another attempt only marked regions of difference between the user's shape and the template piece for all three constraints.

The decision for the first version of the application was to provide the visual feedback in two forms: one for the proportions and alignment constraints and another for corrections to the drawn shape itself. For a correction to a shape's proportions, solid red lines that were as tall or as wide as the bounding box were animated to signify how much to adjust the height or width. Using the example of when a shape should be wider, two vertical lines would animate from the width of the user's shape to a suggested width from the corresponding template piece. Alignment corrections that should be made to any of the edges are marked in the same way. Corrections to the look of the shape were shown in the visual feedback in two ways. First, the template piece was flashed in place for one second on the screen for the user. Flashing the shape instead of making it persist was meant to help establish a mental image for the user, but not let him/her simply copy the shape desired. Second, the regions of difference between the user's shape and the corresponding piece from the example template were shown with a dotted line (Figure 21). However, this separation of visual guidelines for the different constraints was found to be confusing by users as they could not always grasp what the application was asking for, even with text remarks.

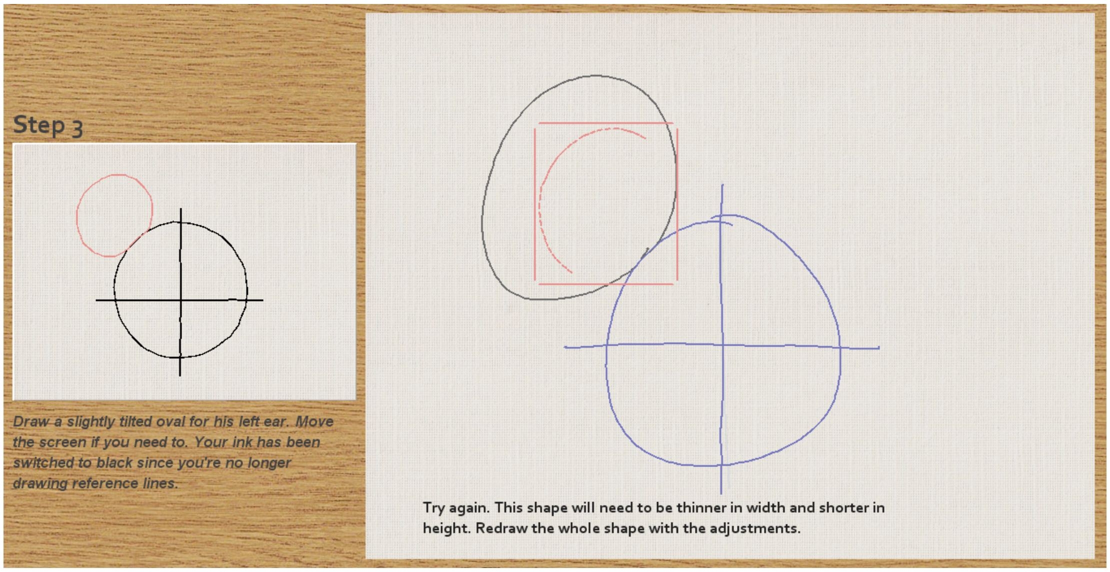


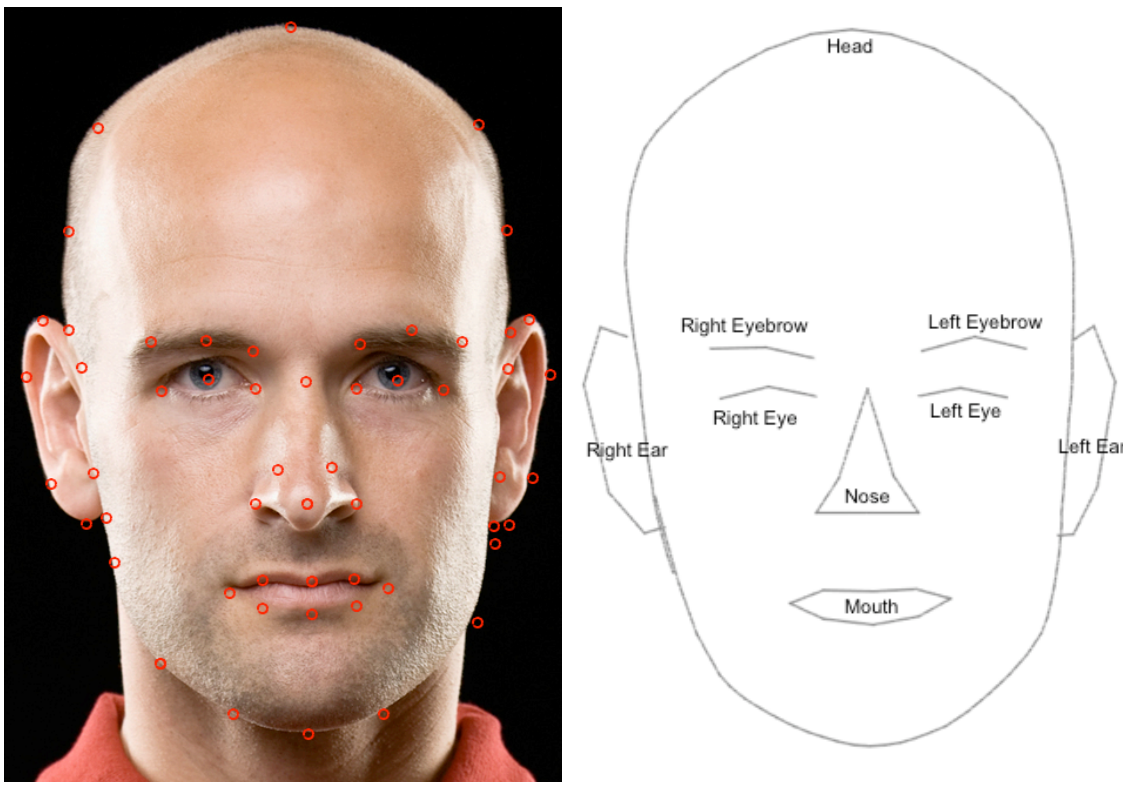
Figure 21. Example of older, two-form visual guidelines. The red lines would animate inward to indicate the application should be thinner. The red dashed curve would indicate where the shape should be.

6. TECHNIQUES FOR ASSISTIVE SKETCH RECOGNITION

Prior sections have discussed the methodology and the interface components for an application meant to assist a user in drawing. The content of this section consists of the underlying methods and algorithms that allow this application to then process, interpret, and verify the drawn strokes of the user. An overview of the fundamentals of sketch recognition is provided in Section 3.

6.1 Obtaining a Template for Comparison

To obtain any interpretation of and, in turn, provide corrective feedback on the user's drawing, the application must have a reference to which it can compare. As explained in Section 5.2, the original methodology called for an art teacher to hand draw the template that was to serve as a basis for the exercises. However, the possibility of error in the template because of only one originator was significant enough that Principle 1 came about (Table 2). A more systematic way was introduced as the methodology evolved. Before use in an exercise, a reference image undergoes offline pre-processing through face recognition software [31] to obtain 40 data points for the facial features. Those facial features captured include the following: corners and pupils of each eye; contour of each eyebrow; tip, bridge, and wings of the nose; contour of the mouth; and contour of the head. An example template for use with sketch recognition is generated from these data points. Additional points are generated for the edge of the nostrils using the positions of data points from the eyes and nose, resulting in a total of 42 points (Figure 22.a). A Bezier spline is fit to the head contour to produce a smooth representation for each and to provide more data for shape matching. Criteria for selecting a reference image was a subject with little to no hair on the scalp, a fairly straight-on look at the camera, and a normal expression. Subjects with limited hair were chosen because the author desired the user study participants to learn using a full view of the human head.



(a) (b)

Figure 22. Outcome of the face recognition step. (a) The final data points after all pre-pressing is completed. (b) The underlying example template generated.

However, it was discovered that the returned positions of these data points could vary by an undesirable amount from the actual positions of the facial features that they are meant to represent. While this result could be considered as degrading of the methodology presented in this thesis, the author assumes that the generic face recognition library chosen will achieve an ideal outcome. The offending points were manually moved to their proper location during development and the probable case was deemed as not a significant issue. Possible solutions would be to use additional face recognition libraries to produce a weighted set of data points or create an interaction mode that would allow a teacher to visually inspect and manipulate the generated template with the assistance of CAD tools. This latter option could also allow for the addition of other details to the template by the teacher. The author did not test with another face recognition package. To continue with the ideal case, data points were manually inserted for the contours of the ears, two facial features not included by the particular commercial package used [31], resulting in 53 points total.

This example template allows the application to know the relative position and size for the facial features of the model in the reference photo. For a few features (i.e. the head), the contour of their shape is also known and can be used verification of the look of a user's shape. Those without sufficient data points served as the basis only for the alignment and proportion constraints that must be satisfied by the user's drawing.

Before use in comparison of facial features, the example template must be scaled to the same size that the user will draw his/her sketch at. To infer this data, the first step of the assisted exercise requires s/he to draw a boundary (Figure 23), also called a "format" [3], that is of the same aspect ratio as the reference photograph (i.e. this box has a width and height of similar proportions). Once the application verifies that this drawing boundary is indeed proportional to the photograph and not too small, displaying corrective feedback as needed, the underlying example template can be scaled to the same size and centered on the format. Asking the users to first identify his/her drawing boundary is not just for sketch recognition, but is in fact equally important for the user. It was noticed with the first version that users had difficulty in appropriately framing their drawing when given the whole page (a.k.a. drawing area), resulting with many having a drawing cut off by the edge of the page because correct proportions where not set initially. Drawing such a bounding box helps the user establish proper proportions from the onset.

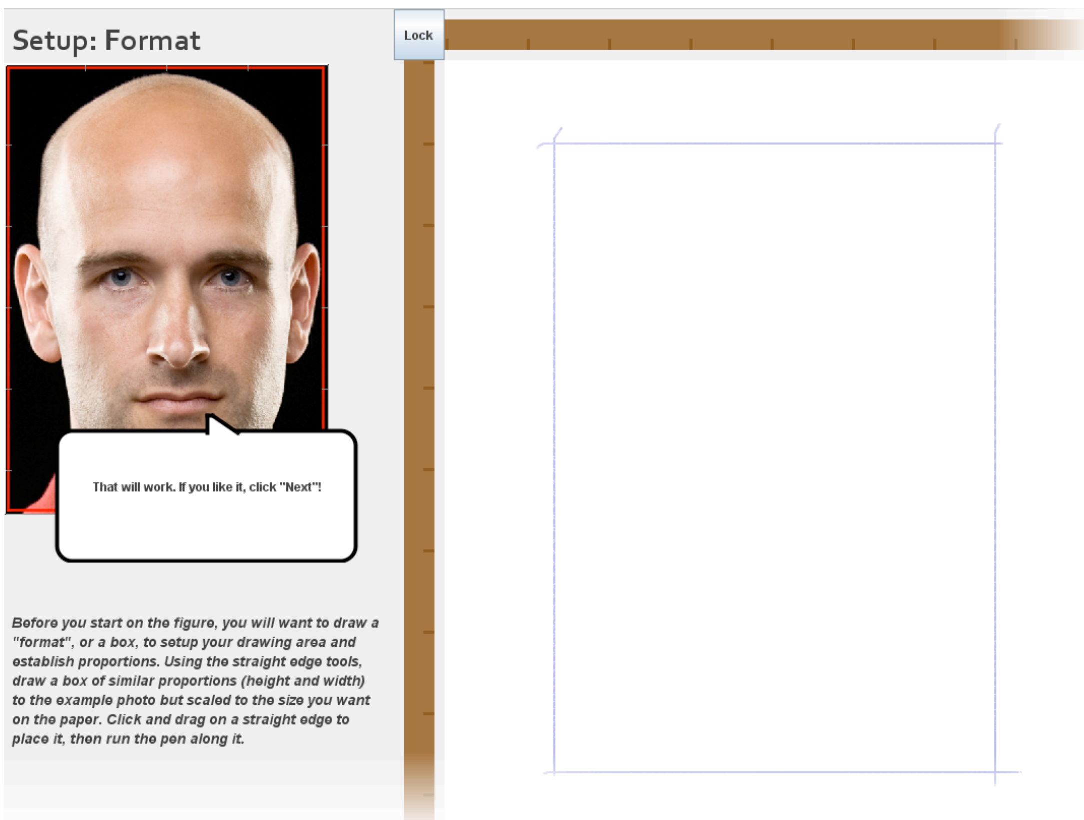


Figure 23. Example of user creating a drawing boundary, or "format." This helps establish proper proportions for the remainder of the exercise.

The underlying example template is not permanently set, however, until the user has effectively drawn the contour of the model's head. This drawing task is the second step of the assisted exercise. After the user has drawn strokes for the head and selected "Check My Work," the template undergoes a best fit to these strokes before being used for comparing constraints and generating feedback (Figure 19), which are described in Section 6.3. A best fit means finding the proper position and scale for the template in order to begin comparing. The template is already scaled to and centered on the aforementioned bounding box. If the center of the bounding box of the shape made by the user's strokes coincides with that of the template (with at least a 70% confidence), the template is repositioned to the shape's center point. Otherwise, the template remains in place as the user will need to move his/her head shape accordingly, and will be informed of such by the visual feedback. If the height and width of the user's shape is close to the height and width of the head piece from the template (with at least a 70% confidence), the template is also rescaled. Otherwise, the user will need to adjust the proportions of his/her shape. The action of fitting the template lets the user decide the placement of the head with the drawing boundary and falls in line with Principle 6 (Table 2) that requires sketch recognition to permit allowable deviation when drawing.

Probabilities for confidences are calculated using a univariate Gaussian distribution. The assumption is that head in the reference photograph is centered and cropped appropriately, something that could be done automatically given the face recognition data. Only when the template has been fit the user's shape for the head and that shape has passed all constraints (Section 6.3) is the template permanently set for the remaining steps based on the head contour's position. The user never sees this template and none of these transformations are ever witnessed.

When exploring how any reference photograph might be used as example, it was realized that the application would not have a-priori knowledge of the reference material but instead rely on preliminary user input and generalizations to provide instruction. Figure 24 shows a mock-up of a tracing mode that the user would have performed with a pen stylus on a digitizing surface at the start of a main drawing exercise. The instructions for this task would be, "Trace all edges of the foreground objects and not trace patterns (e.g. the wood grain)." The purpose of this task was two-fold: familiarize the user with the image and provide a template for sketch recognition to work. The reader of this thesis can probably see obvious problems with this approach. A user could draw something wrong or nothing at all, though it might be possible to use the templates of multiple users on the same image to "learn" the image from combined template. The application would also not know the objects in the picture and have to compare against the template as a whole. Obviously, such an approach would not form a foundation that sketch recognition could stand on. Building on face recognition and photographs of human models overcomes these limitations.

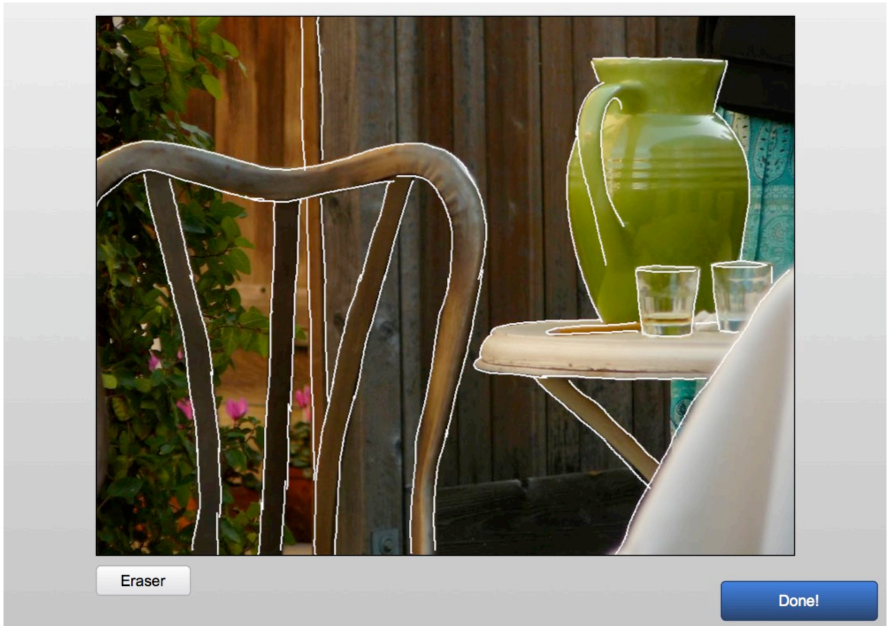


Figure 24. Mock-up screen for an initially proposed tracing mode. This trace would be used as a template for sketch recognition on the drawing later. This approach has many limitations and was not developed further.

6.2 Classifying the Strokes of a Drawing

As mentioned in Section 5.1, all strokes are first checked to determine if they are gestures via low-level geometric recognition [18] before any other classification occurs. An undo gesture is a helix primitive. A scribble-out gesture is a function of time and bounding box density. A right-angle gesture is a polyline with two straight segments that form a ninety-degree angle.

Though the application guides the user step-by-step through creating a sketch, s/he may still wish to modify work done on a prior step and therefore it cannot be assumed that strokes drawn for a step only pertain to the facial feature asked for in the instructions. For example, the user may decide to erase and redraw part of an eye from step three after noticing something about its relationship to an eyebrow drawn in step five. It would be unnatural to limit the user to working on only one facial feature at a time and burdensome to require him/her to explicitly tell the application when backtracking a few steps. To allow the user to draw anywhere but still have the ability to only select the strokes desired when checking constraints of a facial feature, the application reclassifies all strokes of the sketch using a k-Nearest-Neighbor (kNN) classifier (k = 3) on each "Check My Work" request from the user to determine the intent of each.

A first consideration is how strokes that do not relate to the template at all should be handled. For example, the user may want to practice drawing a shape in the extra space on the page, which the instructions encourage at times, or add detail or another facial feature not found in the template, such as wrinkles on the forehead or a beauty mark on the cheek. The approach for drawings on the extra space is to ignore all strokes drawn outside of a buffer area around the drawing boundary. The approach for strokes drawn on the template is to enhance the template to handle any anomalies drawn by the user. Just after the example template is permanently set, an "ignore space" grid is superimposed on it to define the areas that should be ignored by sketch recognition, such as the cheeks and forehead (Figure 25). This "ignore space" is computed by creating a ten column by twenty row grid of placeholders for the template. Each placeholder is assigned to the template and defines a space to ignore if it is outside a threshold distance from any existing data points on the template obtained during the face recognition phase. The threshold chosen is half of the distance between the eyes. This threshold is relatively constant across all models despite any facial expression and allows for an ample buffer between existing data points and any spaces to ignore. Thus the example template consists of a set of data points, each labeled with the facial feature or "ignore space" it belongs to, and can be used as a feature space for classification purposes. One might consider the simply tactic of requesting the user to not draw any features or detail outside of those that have been asked until the very end. However, this limits the drawing freedom of the user and the participants in the user study worked in the iterative manner that's supported by classifying the strokes each time.

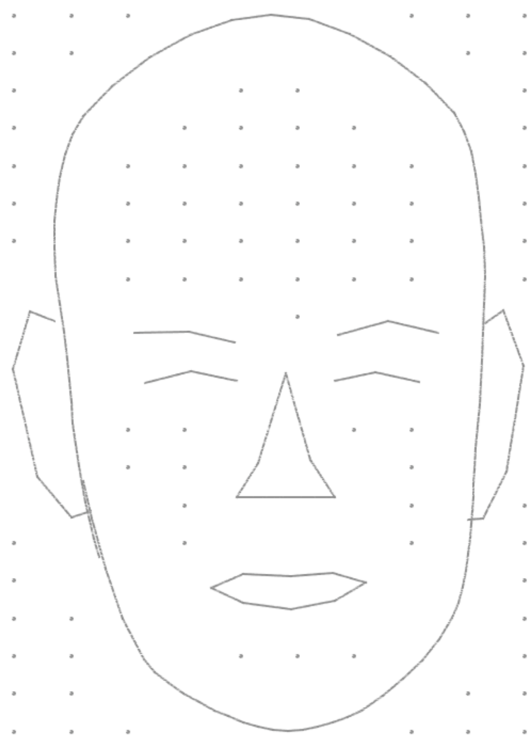


Figure 25. Example template after the "ignore space" grid is superimposed. This grid of dots is part of classifying all strokes on each "Check My Work" request so that the user can freely work on any part of the sketch despite the current step of the exercise.

The second consideration of how to determine what strokes belong to which facial features is answered by use of the kNN classifier [32] with k = 3 using the data points of the example template. A kNN classifier is one of the simpler machine learning algorithms that determines what an object should be classified as by a majority vote from its already classified neighbors. To classify a stroke, its endpoints and center sample point are each matched to the closest data point on the example template via Euclidean distance and then the stroke is placed into its appropriate "pool" of similar strokes (Figure 26). For example, if each selected point of a stroke is matched with a data point for the head contour from the example template, it is put into the pool of strokes for the head. An odd number (k = 3) is used for the points to classify in order to break ties, as opposed to an even number where the classifications could be split down the middle between two classes. If all three receive different classifications, the classifier defaults to that of the first end point of the stroke. Again, all strokes drawn outside the example template are immediately classified as ignore strokes. The classifier also ignores reference lines, which are explicitly marked by the user or through use of the straightedge tool or a right-angle gesture. Of importance is that this technique is only used on the steps that it can support. For latter steps where no recognition data points are known or cannot be inferred from the facial recognition data (i.e. the eye brows, the neck), the classifier is not used, existing strokes assume their prior classifications from previous steps, and new strokes are analyzed based on their positions to the facial feature of that step.

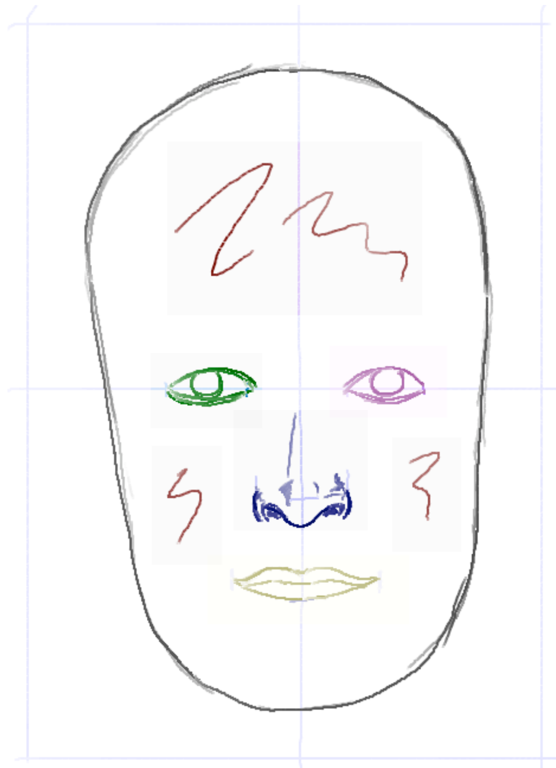


Figure 26. Dramatization of stroke classifications. Each classification is shown by a color. Reference lines are ignored by the classifier.

In comparison to the strokes drawn when rendering a model's portrait, those created for a cartoon character are generally much simpler (e.g. the nose is a oval instead of a group of refined strokes). In the first version of this work, the purpose, or classification, of a stroke was determined by the step the user was on and how it was drawn, as discussed with the interaction techniques presented in Section 5.1. If the shape called for by that step's instructions was not found from the user's stroke via a geometric-based recognizer [18], it was deemed a prep-sketch stroke that would be gradually removed. This technique put obvious limitations on the user's freedom to draw as purposed strokes might be deemed as strokes to remove.

6.3 Determining the Correctness of Strokes

Given classifications and pools of all strokes, the pool for the facial feature of the current step must then pass through constraints (i.e. alignment, proportion, and shape) against its corresponding piece in the example template. For alignment of reference lines, the location and end points of the user's line are checked for an 80% coincidence or greater with the corresponding data points on the example template. For example, the eye reference line (a horizontal line used to place the eyes) asked for as part of the second step is checked against the data points for the pupils of the eyes. For a facial feature, the proportions and alignment of the user's shape (i.e. the pool of strokes from the classifier) are checked against the template piece by comparing the bounding boxes to determine if the two are the same using an 80% confidence level. This constraint helps the user get his/her shape for the current step in line with the example template while still allowing variance (Principle 6).

Next, for the features with ample data points (i.e. the head), the user's shape and template piece pass through elastic matching [33] to determine any differences between the two and verify the user's shape. Elastic matching is a vision-based technique of comparing an object to a template but is flexible by allowing some variance. The algorithm uses a window consisting of neighboring points on each side (here w = 2, so five points per window; [34] to scan along the re-sampled points of the user's shape and find regions that vary greatly between it and the template piece (Figure 27). A value is given to a window instance by matching each of its points to its closest point on the template piece using Euclidean distance and then summing up the distances. When the algorithm has finishing scanning, the average and standard deviation of all the windows is calculated. If a distance for a window is greater than the average plus one standard deviation, it is a region that varies enough from the template to need correction and is becomes part of the visual guidelines shown to the user. This shape constraint is only satisfied when the standard deviation is less than 60% of the average. The goal of this simple algorithm was that it be invariant to the direction, start and end points, and over-tracing so that the user is not required to draw in a particular manner.

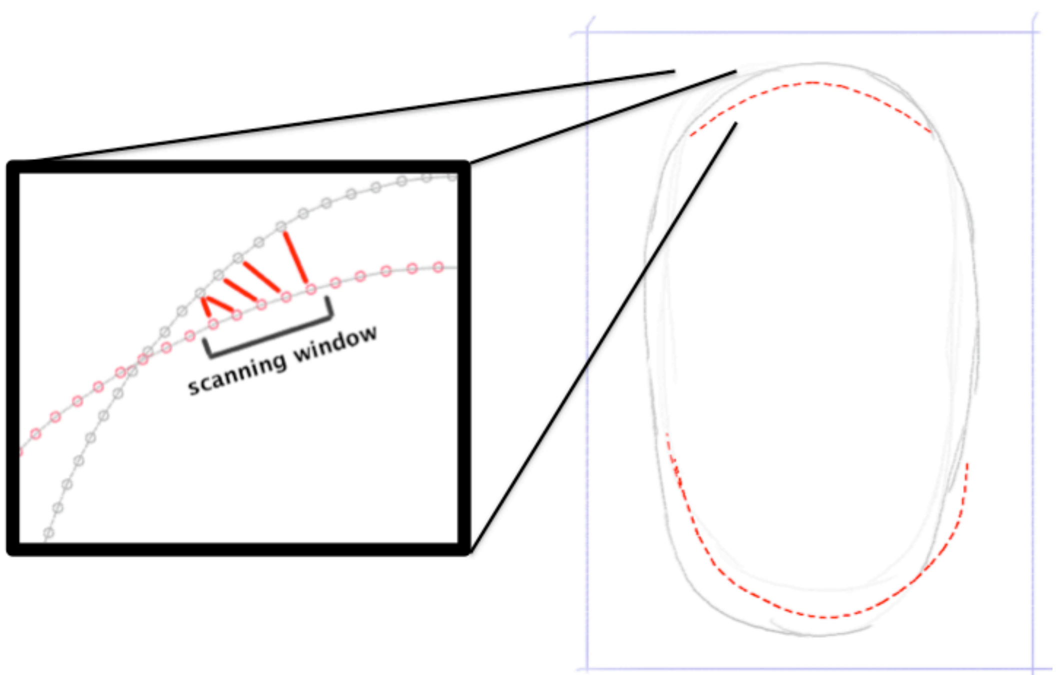


Figure 27. Visualization of the scanning window of the shape constraint algorithm. This simple algorithm finds regions of difference between two shapes.

When a shape is in need of correction, the visual guidelines that are show are generated in one of two ways. For facial features with sufficient data points for their look, the windowing algorithm is used to generate the areas of difference from the template as visual guidelines (Figure 19). For facial features without sufficient data points, the user's shape is reused and fitted to the constraint (Figure 20). If neither case is an option, no visual feedback is provided for the look of the shape though text remarks are always given. When the methodology worked around cartoon characters, what the user was asked to draw for each step could pass through all three constraints (alignment, proportion, and shape) because of the relative simplicity found in the reference object and drawing task, thus visual guidelines could always be generated. With the introduction of more complex shapes, this ability was lost.

7. EVALUATION

Concluding the methodology and principles presented prior, this latter section brings the discussion to cover the user studies and evaluation performed on them. The first subsection describes the quantitative metrics developed for applying a score to a user drawn sketch. Subsection 7.2 and 7.3 then recount the happenings of and the insight learned from user studies. The last subsection presents exploratory work preformed in order to quantify the aesthetics of a user drawn sketch.

7.1 Quantitative Evaluation of User Drawing Ability

While one may empirically label a drawing as better than one done prior, it is desirable to quantify that improvement. For the first user study, no such quantitative evaluation was created. To numerically evaluate the sketches created by the second user study participants (Appendix B), the author initially created a non-exhaustive list of eleven metrics for manually comparing the ratios of a user's sketch to those of the reference image (Table 5). If a ratio from the user's drawing deviated by a difference greater than 0.05 from that of the reference image, it was marked as an error. For example, using the first metric, the width of the head divided by its height for image A in Appendix B is 0.735 and the same ratio for the drawing of the first participant (Male #1) is 0.805. With a difference of 0.07, it was counted as an error. However, it is not desirable for the final error metric to be hand calculated as the user would rather have immediate feedback and, even though these ratios could be incorporated into the application as it has an understanding of what is in the sketch, it would be better for comparison purposes to also handle drawings done outside the assisted exercise.

Table 5. Hand-calculated evaluation metrics originally used to quantify a drawing. These were later replaced by another method.

• Head Width to Head Height Ratio • Chin-to-Eye and Eye-to-Top-of-Head Ratio • Bottom-of-Mouth to Head Height Ratio • Mouth Width to Head Width Ratio • Bottom-of-Nose to Head Height Ratio	• Nose Width to Head Width Ratio • Eye-Spacing to Head Width Ratio • Top-of-Left/Right-Ear to Head Height Ratio • Bottom-of-Left/Right-Ear to Head Height Ratio

To determine the amount of variance that a drawing contains from its reference source, face recognition software is again employed under the ideal case to generate a template from the drawing itself. The new template T_d is then compared with the template originally generated from the photograph for use with sketch recognition T_p. A summation of the distance found between the corresponding pairs of points of each template leads to an overall error for the user's drawing. Similar to the approach used by [35], two corresponding pairs of points (p_i^d, p_j^p) between the templates are then compared using these equations:

$$e\left(p_i^d, p_j^p\right) = \sqrt{e_\theta^{\,2}(p_i^d, p_j^p) + e_{xy}^{\,2}(p_i^d, p_j^p) + e_\ell^{\,2}(p_i^d, p_j^p)}$$

$$e_\theta(p_i^d, p_j^p) = \left|\theta_i - \theta_j\right|$$

$$e_{xy}(p_i^d, p_j^p) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

$$e_\ell^{\,2}(p_i^d, p_j^p) = \left|\ell_i - \ell_j\right|$$

The orientations of each pair of points are represented by θ_i and θ_j, (x_1, y_1) and (x_2, y_2) are the locations of each pair's calculated midpoint, and lastly ℓ_i and ℓ_j are the lengths each pair. Points are not compared between facial features (i.e. each point in a pair is part of the same facial feature). With *P* standing for the number of corresponding pairs for a facial feature *f*, the overall error is then calculated as:

$$e(T_d, T_p) = \sum_{f \in T} \frac{1}{P} \sum_{(i,j) \in P} e\left(p_i^d, p_j^p\right)$$

Figure 28 shows a visualization of comparing a face recognition template from a user's drawing to one from a reference photo. To align the templates, the one for the user is scaled to have a matching height and moved to have the same bounding box center point as the template for the photograph. The lower the error, the more proportionally correct the sketch was drawn, thus the objective is to achieve as low a score as possible. The errors calculated from template matching for each drawing of the second user study are shown in Appendix B, however this approach was not implemented as the time of the study. The errors from the hand-calculated metrics are also shown on the side in parentheses. In order to create a control group, a supplemental user study was also performed a time after the second one (Appendix C). For this latter study, the assistive sketch recognition was turned off so that no corrective feedback was given but the participants still followed the same instructions and saw the same manipulations to the reference imagery.

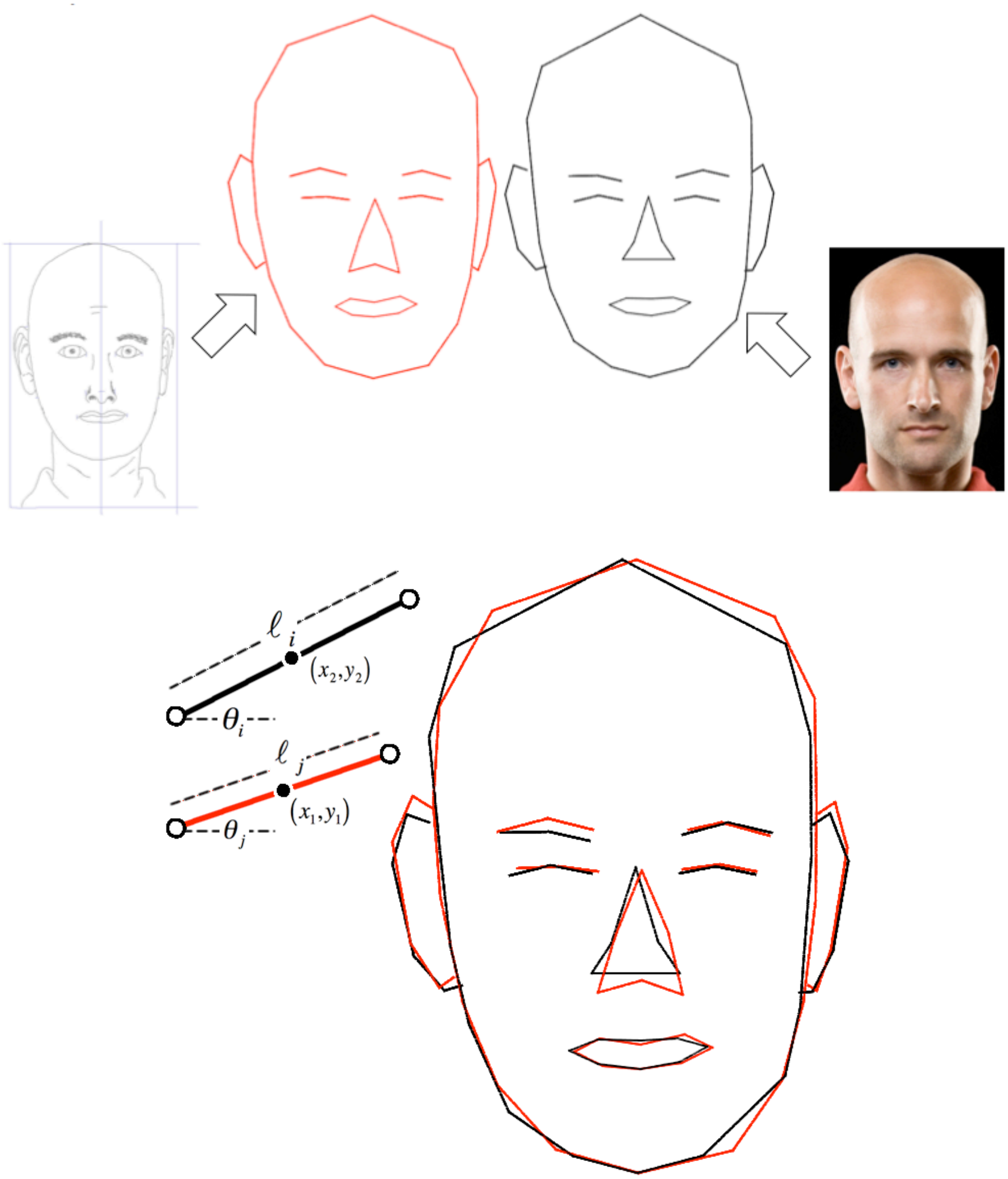


Figure 28. Comparing face recognition templates from a drawing and a photograph. M2's drawing (red) is compared to the reference photograph C (black) in Appendix B. This is done by summing the distance between all corresponding pairs of points of each template. The total error here is 101. The templates pulled from each source as shown above.

All of the error counts shown in Appendix B for the second user study generally reflect what is to be expected. In every case, the error on a participant's drawing for an assisted exercise is lower than when the user is unassisted. When comparing the participants who were assisted to those in the control group who only had the instructions and changing reference image (images B and C of Appendices B and C), the average error was roughly a factor of two lower for the assisted participants and their standard deviation was significantly lower. This strongly suggests that users under this methodology will produce more proportionally correct drawings. However, when looking at the three unassisted freehand exercises for each study (the first, middle, and last columns of Appendixes B and C), the average error can be considered the same for each pair of corresponding exercises. This data might suggest that each group is still capable of drawing the same amount of imperfections when unassisted, but the author believes that a longer study would show different results. While the average error does decrease for each subsequent freehand exercise between the two studies, this fact may be negated as a general sentiment from the participants was that the reference material in images A and B is harder to draw than that of C or D. This is discussed further in the Section 7.3.

7.2 Qualitative Evaluation of User Experience for the First User Study

Six participants performed the first user study: three females (F1, F2, F3) and three males (M1, M2, M3). All participants were employees of a campus department and ranged from college-age to their mid-forties. No person had taken an art class past the early years of grade school and most were quick to admit an inability to draw when asked to participate. The user study consisted of a warm-up exercise and then five exercises with the example template of a famous cartoon character, two of which used assisted sketch recognition. Participants drew on a portable tablet PC using a pen stylus. Results are shown in Appendix A.

For the warm-up exercise, the participant read instructions to adjust the placement of the tablet PC to his/her liking and read suggestions on how to hold the pen and move his/her arm while drawing. The participant then drew two simple sketches (i.e. the Olympic rings and a heart figure) that demonstrated the clean up of unneeded strokes and corrective feedback. F2 liked the pen-holding suggestion of using less wrist movement and more of the full arm. All participants explored the technique and three kept it for the duration of the exercises.

The first exercise required each participant to create a freehand drawing of the full example template at the best of his/her ability. Three participants expressed delight when they recognized the famous cartoon character they would be reproducing. The unconfident drawers commented that recreating the character would be "tough." Both sentiments were expected.

Going piece-by-piece for the second exercise, the participant was guided by the application through steps to create an upside down version of the cartoon character. Some participants did not immediately realize the steps were leading them to recreate the same character, even though its more famous features were present but flipped. M1 expressed his enjoyment of the congratulatory textual feedback after each step. F1 asked if the visual feedback for the shape would not flash, but instead remain present. On one instance, she also said the visual feedback had discouraged her from drawing the shape as smoothly as possible because she focused on making it pass over the visual feedback. This experience and similar ones like it led to a redesign of the corrective feedback. Many participants deemed the recognition too picky about shapes that were smaller (e.g. a dimple, the tongue, a pupil of an eye). Even after the corrective feedback informed M2 to draw his shape larger, M2 repeatedly tried to start his drawing at a 1:1 scale with the corresponding piece on the example template, which was smaller than the drawing area (Figure 8). Requiring the user to create a drawing boundary first alleviated this problem (Figure 23).

Going piece-by-piece for exercise three, the participant was guided through the steps to create the cartoon character. F1 only skimmed the directions and drew faster because, as she explained afterward, she felt she knew “what [she] was making now” after the prior exercise. Other participants drew faster as well. Two participants expressed that drawing the example template upside down was easier, or that its end result looked better. Again, a common sentiment was that the recognition was too strict for smaller features. Requiring the user to pass the constraints, even with thresholds that were gradually lower on each attempt, contributed to a poor experience.

Lastly, each participant created a freehand drawing of the full example template at the best of his/her ability. Half of the participants liked their end result of this exercise better than that of the first. Two of the remaining participants liked their drawing from the first better, while F2 liked her drawing from the second exercise best because of its proportions (though this was not one of the answer choices to the question asked of her).

7.3 Qualitative Evaluation of User Experience for the Second User Study

After the second iteration of development, five participants tested the application in a two-part, two-day user study. The study consisted of performing a blind contour drawing exercise as a warm-up for each session and then creating two to three drawings of human models either freehand or with assistance. All participants were male (M1, M2, M3, M4, M5) and were of college age. Participants were selected based on their unfamiliarity with sketch recognition and computer science and their "confession" to not being able to draw. The first session lasted for three hours and each participant produced three drawings. All participants were told to expect to take 30-60 minutes per sketch and that there was no cutoff or reason to rush. The imagery used was headshots of male babies (see Appendix B) because the thought was that these subjects would be easier to draw with their simpler features, though this was not the case.

After a short demonstration of the straightedge tool and available gestures, each participant was asked to produce a freehand drawing to the best of his ability of a reference image of a male baby. The intent was for the participant to produce a rendition that was exemplary of his drawing ability. Most participants expressed feeling a little overwhelmed given the task to draw a human figure. M5 commented that he did not think he had ever drawn a human face. The author of this thesis viewed this perceived difficulty of the task as validation for a reference test bed that could allow for the most room for improvement in the participants. Participants of the first user study also found their drawing task to be tough.

Results are shown side-by-side in Appendix B. When asked what was wrong with their pictures in a post-questionnaire, most participants could point out incorrect strokes, such as the forehead was not high enough or the mouth was not shaped right. M2 expressed that he did not know how to convey what he wanted in the sketch. Most participants felt that their drawings were not reflective of a male baby and, on an empirical scale of 1-5 with five being best, all chose a rating of 2 ½ or below.

Both sessions contained a blind contour exercise, or "pure contour drawing exercise," as described in Section 4.4. The choice object was the drawer's non-dominant hand positioned to form wrinkles and participants were told to take a full five minutes each time. The objective of this exercise is not to produce an accurate representation of the hand (as the final outcome will be considerably far from it) but to enter the R-mode by training the eye to focus on the details and move the drawer into the mindset for processing them. Participants were initially puzzled by the instructions since they could not see their strokes till the end, but M3 commented that he noticed more detail than he would normally.

For the next drawing exercise, each participant used the step-by-step instruction and assistance of the application to create a drawing of another reference image of a male baby. All participants expressed that they found the step-by-step approach for constructing a face useful. Three of the five participants commented that they did not have much prior knowledge of the symmetry and proportions found in creating the human face. M3 especially liked how the dimming of other facial features on the reference image helped him concentrate on drawing the desired feature, though M1 asked if the dimming could be made slightly lighter so he could see more. All participants had some difficulty adjusting to the gesture for marking reference lines. Originally a flick motion, it was changed to the right-angle gesture between the sessions. M2 also had difficulty adjusting to the straightedge tool, and was the only participant to do so. Some participants initially asked why erased strokes did not completely disappear, but all made use of these faded older strokes for redrawing or placement of new strokes during the exercises. The gradual fade of these strokes was subtle enough that many did notice their disappearance.

The suggestions provided by the application also took adjustment on the part of the participant. If the participant found the suggestion useful, he would make the correction and move on. However, some suggestions were perceived as incorrect when they were not or the application was unprepared for the participant’s conceptual model of how it should work. As an example, M2 verbalized disagreement with where the application suggested putting a reference line for the eyes of his sketch as he felt it should be higher and the forehead smaller. However, the application's suggestion was later verified as correct against the example template. M1 had a similar situation when the application suggested moving the nose lower then he desired. Generally, the participant would respond with two or three additional attempts but then opt to advance to the next step if the application still did not accept his work. There were instances where the application did misinterpret the participant's shape and provide incorrect feedback, causing the participant to simply advance the step. Beyond drawing reference lines, many participants also adopted some scheme for measuring proportions. M5 used his pencil as suggested in the instructions. M2 desired to draw a full grid in the drawing area, which the application unfortunately did not support.

Overall for this exercise, some participants felt that this instance of the exercise was useful but too lengthy and also asked for the application to support shading on their sketch. M5 did take a picture of his drawing with a camera and proudly remarked that this was the first time a drawing of his turned out to be like a face.

For the last exercise of this session, each participant drew another sketch of the same reference image used for the first freehand exercise. The intent was to see what differences and improvements the participant would make within one sitting after having done the prior two exercises. Every participant used a variation of the reference line setup he had learned in the prior exercise. On an empirical scale of 1-5, all the participants (except for M1) rated their new drawing 1-2 points higher than their first when viewing a side-by-side comparison. M1 said that the symmetry and the nose were better, but that he felt he'd repeated the same mistake of drawing the eyes and mouth too large, hence the unchanged rating. A common sentiment was that the new drawing was more symmetric, proportions were better, and it looked more like a human child. M2 expressed that his "drawing was better because of the template." When asked what was wrong with their pictures, participants could still find problems, such as the mouth and eyes being too large. Only M2 attempted to draw the tilt of the head in reference image

A (see Appendix B) adjusting his reference line setup accordingly, but later opted to restart and draw without the tilt.

The second session lasted for two hours and each participant produced two drawings. The imagery used were headshots of bald male adults. The author felt a male adult subject would be harder to draw yet more reflective of traditional facial proportions. Participants began by repeating the blind contour exercise.

Each participant used the step-by-step instruction and assistance of the application to produce a drawing using a reference image with a straight-on view of the model from the camera (see image C in Appendix B). Except for M1, all participants took less time to complete this exercise in comparison to that of the first session and chose a rating between 3-4. The ability to do shading was again requested, and M4 desired to have more control when erasing, more specifically the ability to erase parts of a stroke and instead of removing the whole stroke each time.

As the last exercise of the study, each participant did one last freehand drawing of a different male adult model with a slightly lower camera angle that changed some proportions on the model (e.g. the chin-to-eye and eye-to-top-of-head ratio). The intent was to see how the participant would respond to the different proportions; if he would draw what was there or draw a view that was straight on, repeating the proportions learned prior. Again, every participant used a variation of the reference line setup from the step-by-step exercise. M1 was the only participant to explicitly comment that the perspective was different on this model, noting that the eyes were not in the middle of head anymore. All except M4 appeared to make adjustments for the perspective in their drawings. Again, all participants gave an empirical rating of 3-4 except for M1.

M3 and M4 verbally expressed more confidence to draw after this second session, and M3 said that this exercise was easier after doing the assisted exercises. We asked the participants to compare their last drawing to their first drawing of the baby from the prior session. The common sentiment was that their last drawings had better symmetry and proportions and was more identifiable. M2 stated that, "If you knew this guy, you could identify him." Besides M3, another common sentiment was that the reference image of the baby was actually harder to draw than the one of the adult, pointing out the baby's open mouth and tilt of the head and the adult's more developed features. M1 stated that if redoing the first picture again, he would be able to "line up the facial features better."

Some time later, a supplemental study was performed to gain a control group for the second user study. It was consisted of four males and one female participant, all of the same characteristics and qualifications listed for the original participants. Reults from this study are shown in Appendix C and comparison of the data from these two studies is discussed in Section 7.1.

7.4 An Exploration in Quantifying the Aesthetics of a Drawing

The metrics presented do not address the aesthetics of a drawing but only the proportions and placement of key facial features. What is more desirable is a way to quantify the aesthetic qualities of the sketch. This will require an approach that uses vision-based recognition and this discussion of this subsection covers possible techniques for doing so. One approach might be to take advantage of the symmetry found in the human face (i.e. facial features generally mirror on a vertical axis that divides the head). After normalizing to only major facial features, the user's drawing could essentially be folded over on itself (Figure 29) and a distance metric generated between the features. This simple approach makes many assumptions though. While there is currently a stipulation that the reference image must have a reasonably straight-on point of view and the reference line that the user creates for aforementioned vertical axis is known, there is still not a guarantee that the drawing will contain symmetry. This is seen in the user's drawing (Figure 29) and the reference image itself (Figure 22) as the face is symmetric but the perspective has its placement favoring one side. This approach would also require normalizing the image by removing or masking strokes that were meant as details and accessories such as hair, jewelry, and clothing. A more robust approach would be needed.

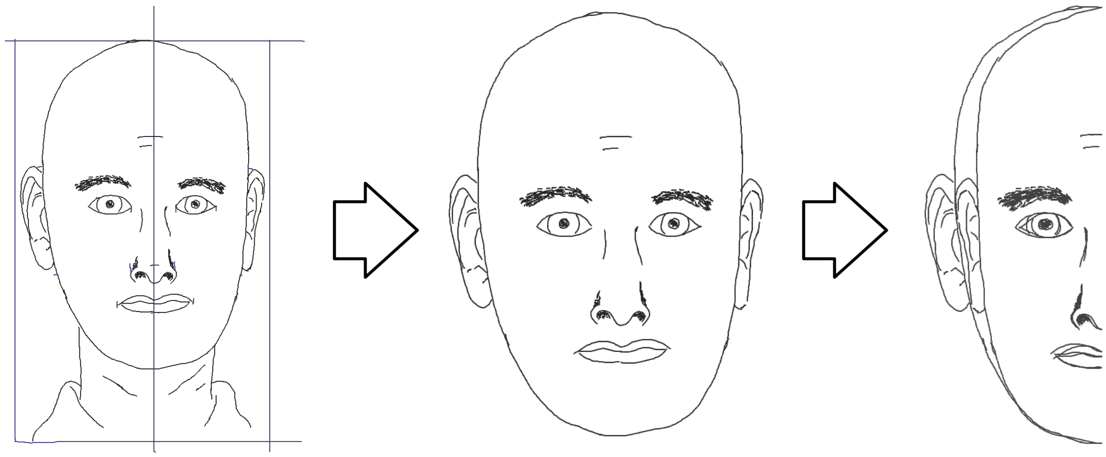


Figure 29. Example of using symmetry as an error metric. The drawing on the left is from participant M2 of the second user study. It shown in the middle as normalized for the major facial features. The right shows a check of the symmetry.

A more holistic approach could compare the user's drawing with a generated "expert" sketch of the reference image to calculate an error metric. There are multiple ways to obtain this expert sketch. An actual experienced artist could provide his/her rendition of the image under the instruction to stay true to the form and refrain from personal flair. This technique has drawbacks, though, as it would be time-consuming, could require a specific drawing style from an unknown artist, and would likely eliminate the possibility of this methodology and application ever being standalone. The absence of any "personal touch" on a work by an artist could also be considered hard to guarantee. Another source could be the example template generated from the face recognition software. Though the template is void of everything but the major facial features, a normalized version of the user's drawing could be used for comparison (Figure 30.a). This option was explored in tandem with an error metric calculated using the low level image processing technique of optical flow [36], which is discussed next. The example template works well for proportions and alignment of facial features, but this vision-based technique will obviously need a better representation of features (e.g. the eyes) to generate a fair comparison (Figure 30.d).

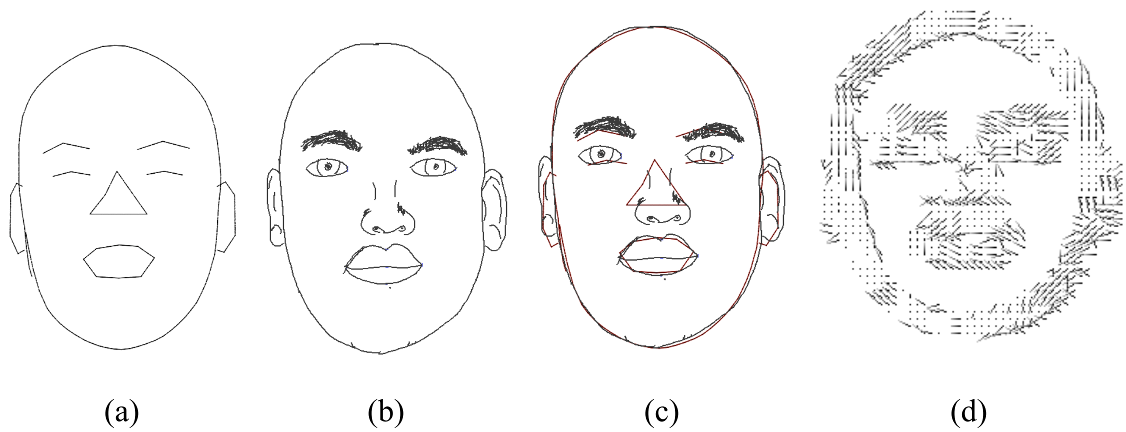
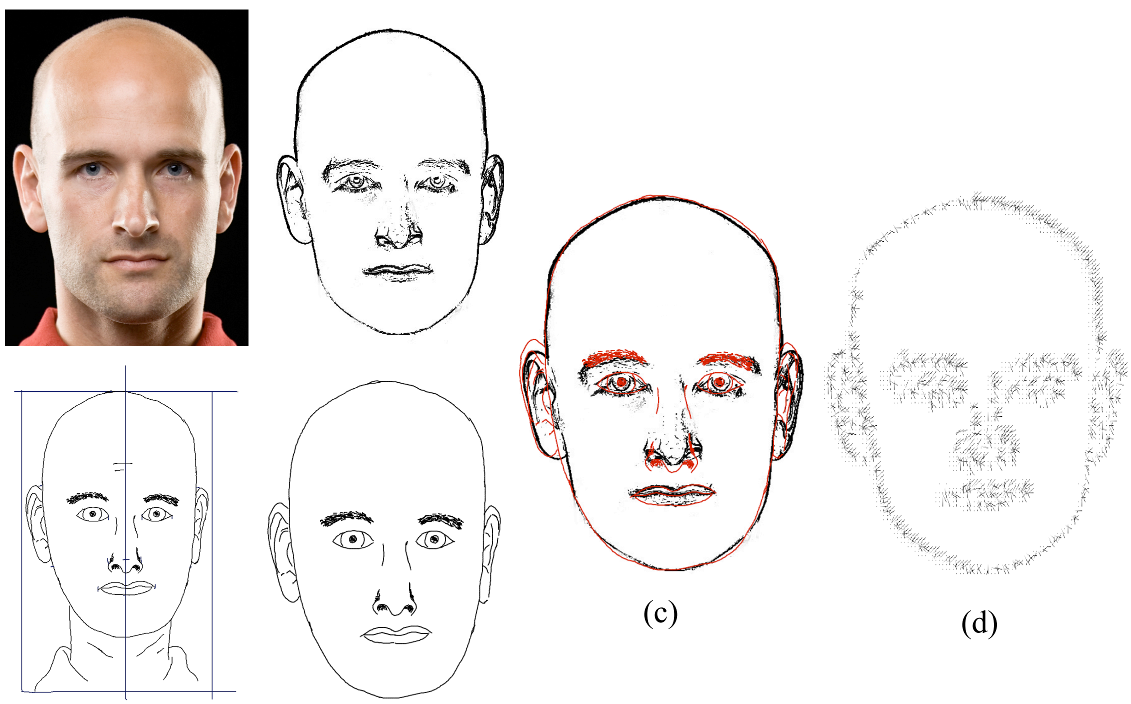


Figure 30. Comparing user's drawing to the example template using optical flow. (a) The example template. (b) The user's drawing done without assistance and then normalized. (c) The drawing with the template superimposed in red for visualization purposes. (d) The vector field between (a) and (b) calculated using optical flow.

In order to generate an expert sketch of the reference image, the image itself could be converted to the appearance of a sketch. Synthesizing a sketch from a photo of a face is a research problem found in the face recognition and image-based search domain. Hong [37] formed a technique to create a simplistic line sketch of a photo by mimicking artists' traced and hand-labeled renditions of example photos. Lui [38] mapped the nonlinear relationship between example photos and their drawn sketches by computing and synthesizing patches from this training set and then combining all the patches to form a pseudo-sketch of a given image. Xinbo [39] created a more holistic approach by using multi-state, two-level Hidden Markov models to generate full pseudo-sketches that were then merged into one pseudo-sketch using a “selective ensemble strategy.” All of these approaches, however, are example based and do require a large corpus of photos and their corresponding sketches to train on in order to do face sketch synthesis. Xu [35, 40] created a slightly different approach, though it still requires a large corpus, by producing a “sketch graph” for a photo (similar to a caricature) through a decision tree that matched facial features with their closest representation in a library of existing drawn elements. Jin [41] did develop a non-example based approach that required little user interaction by reducing an image to gray scale, using local edge detection to discover curves and generate a vector field, and then synthesizing pencil strokes by interpolating this field. This later approach could also be mimicked offline by hand using any one of many graphics editing applications. Any one of these techniques would suffice for obtaining and expert sketch of the reference image given the preexisting conditions.

Looking at the image processing technique of optical flow (Figure 31), the author proposed using it as an error metric, more specifically area correlation optical flow [36]. Traditionally used for video and image sequence analysis, optical flow is a way of describing the motion depicted by two images. Area correlation optical flow works by matching the location of a point in one image to its new location in the second image and then returning a velocity as the distance that the point moved. The intensity of the point is used to find the latter location within a given window that is moved to search around the original. Limiting the search movement of the window to within a displacement factor from the original reduces computation. Two locations are paired for a point when they have the smallest difference in intensity. Two assumptions are also made for this technique: that the intensity of the point remains constant and that the neighboring points move with a similar velocity.



(a) (b) (c) (d)

Figure 31. Comparing user's drawing to an "expert sketch" using optical flow. (a) The original photo and drawing. (b) The expert sketch created using a third-party graphics application, and the user's assisted drawing shown normalized. (c) The expert sketch with the drawing superimposed in red for visualization purposes. (d) The vector field between (a) and (b) calculated using optical flow.

Though promising, optical flow had weaknesses when comparing a drawing and an expert sketch. First, it is uncertain how to quantify an error from the vector field generated by optical flow as the technique is traditionally used to describe motion depicted by two images. One could start by eliminating all vectors of magnitude less than the displacement factor, leaving only those of significance to consider, but the problem of translating these to the user still exists. Second, this technique is dependent on the window size and displacement factor chosen, which [39] similarly pointed out in his summary of the patch-based face sketch synthesis technique. Besides contributing to the inability to quantify an overall error, this also means that corresponding pieces between the expert sketch and the drawing may not be matched if the window size and displacement factor are not set large enough. Increasing these two factors can increase the amount of false positives, not to mention computation time.

As one last consideration, Yee [42, 43] developed a tool for performing regression testing on imagery produced from a multimedia studio's rendering pipeline. Modeled after how human visual system, the tool can compare two images to verify if they are perceptually different based on spatial frequency, luminance, color, etc. However, this approach is meant for imagery that can contain variations but both originate from the same source and not with different hand-drawn styles.

8. SUMMARY

Teaching a person to understand his/her perception and successfully sketch an object is a difficult task to attain, especially via a computer application. This thesis presents the first methodology for using computer-aided instruction to assist a student in learning to draw human faces. This methodology uses face recognition and sketch recognition to understand the reference photograph of a human model and a user's drawing of it. The subject of the human head was chosen because it has many advantages. First, it is a drawing task that is enjoyable to achieve. Second, improvement is quickly noticed as many make common mistakes with drawing the human head. Third, a body of knowledge already exists for drawing the human head and face. Fourth, a wealth of prior image-based techniques can be built upon to implement effective sketch recognition. The final application built under this methodology uses principles obtained from art literature in concert with principles gained from user interactions with the two versions of the application themselves. The iterations of development and lessons learned were also recounted and user experiences shared.

8.1 Expanding on This Methodology

As requested in the latter user study and detailed in Principle 9 (Table 2), this work would benefit greatly by the addition of shading. The rendering of the stroke could be manipulated by contextual information such as pen pressure, tilt, and drawing style and/or the user could be allowed to select "pencils" with different ratings (e.g. 5B, 2B, H) that are commonly used in the art world. This would open new areas of instruction for such an application, such as teaching about hatching, crosshatching, and using these to create perspective.

Enhancements could continue with the sketch recognition by automatically determining reference tick marks without requiring the user to use an explicit gesture to mark them. It might be possible to find distinguishing features of strokes meant as reference tick marks, such as their relatively short length. A developer might also employ the conceptual model of semi-transparent tracing paper where the reference lines and the main drawing are on two separate sheets and the user can move one on top of the other yet still see the strokes of both. Sketch recognition could also be enhanced by a more natural, freeform erasing paradigm, not one that is all-or-nothing of the stroke, as requested in the latter user study.

It would also be advantageous to use artificial intelligence algorithms to explore how the application might generate its own text remarks and tips based on the example template instead of these being hard coded and hand crafted by the developer. For example, the application could process the shape of a facial feature, given enough data from the image analysis, and comment on special characteristics of it in the instructions for that step using the format, *"[call to attention] [facial feature] [verb] [description]"* (e.g. "*Notice that the jaw forms three sharp turns.*" "*Take note that the head is slightly tilted.*").

This same methodology could also be applied to variants of the application presented. One such variant could work with profile pictures of human heads, which have similar rules-of-thumb for proportions like those on the face (e.g. placement of the ear in relation to the nose and chin) and can still use face recognition technology. This version could show "negative space" (another technique for finding the contour of a shape) on the reference photo based on where the user moves his/her pen or draws a line. For the creation of caricatures, another variant might modify the underlying template as the user purposefully exaggerates proportions. To allow the user to interactively play with the subject to be drawn and also obtain an example template for sketch recognition, another variation could have the user manipulate a mannequin and then practice full-body figure drawing. This 2D or 3D mannequin (Figure 32) could be implemented and moveable using inverse kinematics and would allow the application to know what the user is trying to draw and how to provide corrective feedback. Given the amount of existing 3D models for animation and games that can be found online, this approach might take one of them as input, allow the user to manipulate the view, and then help the user draw it.

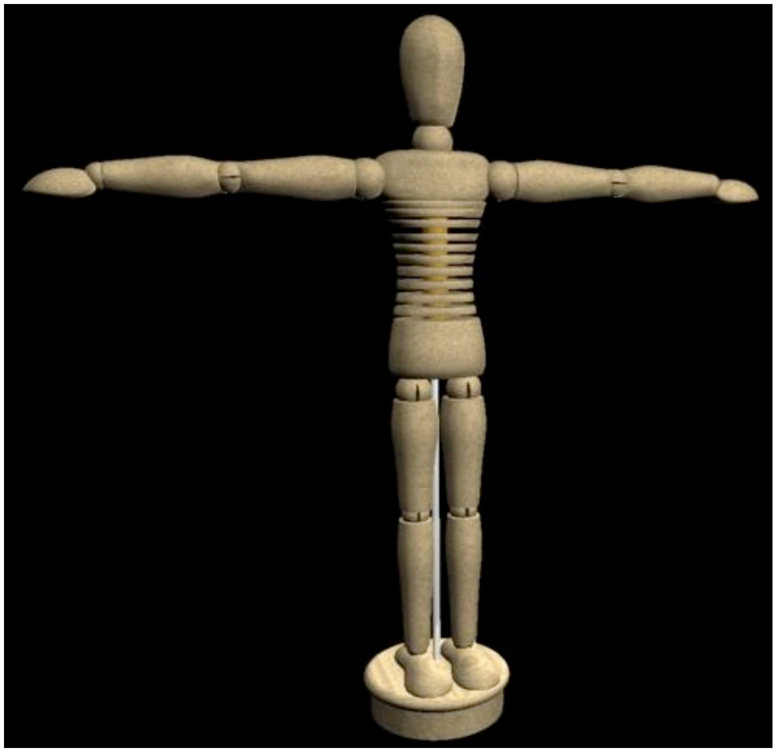


Figure 32. Example of an interactive mannequin for full body drawing. This view could be manipulated by the user for a figure drawing application.

REFERENCES

1. D. Roam, *The Back of the Napkin – Solving Problems and Selling Ideas with Pictures*. New York, New York, USA: Penquin, 2008.

2. M.C. Yang, and J.G. Cham, “An analysis of sketching skill and its role in early stage engineering design,” in *Journal of Mechanical Design*, vol. 129, no. 5, pp. 476-482, May 2007.

3. B. Edwards, *The New Drawing on the Right Side of the Brain*, New York, New York, USA: Penguin, 1999.

4. M.C. Corballis, and I.L. Beale, *The Psychology of Left and Right*, Hillsdale, New Jersey, USA: Erlbaum, 1967.

5. J.T. Bruer, “In search of...brain-based education,” 1999 [Online]. Available: http://www.pdkintl.org/kappan/kbru9905.htm. [Accessed: January 2009].

6. J. Willats, “How children learn to draw realistic pictures,” in *The Quarterly Journal of Experimental Psychology*, vol. 29, no. 3, pp. 367-382, August 1977.

7. M. Bartel, “How to teach drawing to children,” 2007 [Online]. Available: http://www.goshen.edu/art/ed/draw.html. [Accessed: January 2009].

8. B. Hoddinott, “DrawSpace.com,” 2009 [Online]. Available: http://www.drawspace.com. [Accessed: March 2009].

9. C. Neumann, “Teaching drawing in engineering courses in Santa Catarina State – Brazil,” in *International Conference on Engineering and Education*, Coimbra, Portugal, September 3-7 2007.

10. D.A. Pariser, “Two methods of teaching drawing skills,” in *Studies in Art Education*, vol. 20, no. 3, pp. 30-42, 1979.

11. M. Bartel, “Teaching creativity,” 2008 [Online]. Available: http://www.goshen.edu/~marvinpb/arted/tc.html. [Accessed: January 2009].

12. G. Clark, E. Zimmerman, and M. Zurmuehlen, *Understanding Art Testing: Past Influences, Norman C. Meier's Contributions, Present Concerns, and Future Possibilities*, Reston, Virginia, USA: National Art Education Association, 1987.

13. B.J. Orde, “Drawing as visual-perceptual and spatial ability training,” in *Proc. of National Convention of the Association for Educational Communications and Technology*, Albuquerque, New Mexico, USA, February 1997, pp. 271-278.

14. D. Rubine, "Specifying gestures by example," in ACM *SIGGRAPH Computer Graphics*, vol. 25, no. 4, pp. 329-337, July 1991.

15. M. Oltmans, "Envisioning sketch recognition: a local feature based approach to recognizing informal sketches," Ph.D. dissertation, Massachusetts Institute of Technology, Boston, Massachusetts, USA, 2007.

16. J.O. Wobbrock, A.D. Wilson, and Y. Li, "Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes," in *Proc. of the 20th Annual ACM Symposium on User Interface Software and Technology (UIST)*, Newport, Rhode Island, USA, October 7-10, 2007, pp. 159-168.

17. T. Hammond, and R. Davis, "LADDER: a language to describe drawing, display, and editing in sketch recognition," in *Proc. of 2003 International Joint Conference on Artificial Intelligence (IJCAI)*, Acapulco, Mexico, August 9-15, 2003, pp. 461-467.

18. B. Paulson, and T. Hammond, "PaleoSketch: accurate primitive sketch recognition and beautification," in *Proc. of 13th International Conference on Intelligent User Interfaces (IUI)*, Gran Canaria, Spain, January 13-16, 2008, pp. 1-10.

19. T.M. Sezgin, T. Stahovich, and R. Davis, "Sketch based interfaces: early processing for sketch understanding," in *ACM SIGGRAPH 2007 Courses*, San Diego, California, USA, August 5-9, 2007, Article No. 37.

20. A.C. Long, J.A. Landay, L.A. Rowe, and J. Michiels, "Visual similarity of pen gestures," in *Proc. of SIGCHI Conference on Human Factors in Computing Systems*, The Hague, The Netherlands, April 1-6, 2000, pp. 360-367.

21. C. Alvarado, and R. Davis, "SketchREAD: a multi-domain sketch recognition engine," in *Proc. of 17th Annual ACM Symposium on User Interface Software and Technology (UIST)*, Santa Fe, New Mexico, USA, October 24–27, 2004, pp. 23-32.

22. C. Alvarado, "Sketch recognition user interfaces: guidelines for design and development," in *AAAI 2004 Symposium on Making Pen-Based Interaction Intelligent and Natural*, Arlington, Virginia, USA, October 22-24, 2004, pp. 8-14.

23. J. Arvo, and K. Novins, "Fluid sketches: continuous recognition and morphing of simple hand-drawn shapes," in *Proc. of 13th Annual ACM Symposium on User Interface Software and Technology (UIST)*, San Diego, California, USA, November 6-8, 2000, pp. 73-80.

24. S.-H. Bae, R. Balakrishnan, and K. Singh, "ILoveSketch: as-natural-as-possible sketching system for creating 3d curve models," in *Proc. of 21st Annual ACM Symposium on User Interface Software and Technology (UIST)*, Monterey, California, USA, October 19-22, 2008, pp. 151-160.

25. S. Tsang, R. Balakrishnan, K. Singh, and A. Ranjan, "A suggestive interface for image guided 3D sketching," in *Proc. of SIGCHI Conference on Human Factors in Computing Systems*, Vienna, Austria, April 24-29, 2004, pp. 591-598.

26. S.B. Fan, "Helping teachers automate student sketch assessment," in *Proc. of IEEE Symposium on Visual Languages and Human-Centric Computing (VLHCC)*, September 23-27, 2007, pp. 256-257.

27. Y. Oh, E.Y.-L. Do, and M.D. Gross, "Intelligent critiquing of design sketches" in *AAAl Fall Symposium: Making Pen-Based Interaction Intelligent and Natural*, Arlington, Virginia, October 21-24, 2004, pp. 127-133.

28. Y. Oh, E.Y.-L. Do, and M.D. Gross, "Critiquing freehand sketching - a computational tool for design evaluation," in *Visual and Spatial Reasoning in Design (VR)*, Cambridge, Massachusetts, USA, 2004, pp. 105-120.

29. P. Taele, and T. Hammond, "Hashigo: a next-generation sketch interactive system for japanese kanji," in *Twenty-First Innovative Applications Artificial Intelligence Conference (IAAI)*, Pasadena, California, July 14-16, 2009.

30. G. Clark, and E. Zimmerman, *Teaching Talented Art Students - Principles and Practices*, New York, New York, USA: Teachers College Press, 2004.

31. Luxand Inc., Luxand FaceSDK, version 1.81, 2009.

32. C.D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*, New York, New York, USA: Cambridge University Press, 2008.

33. S.M. Watt, and X. Xie, "Prototype pruning by feature extraction for handwritten mathematical symbol recognition," in *Proc. of Maple Conference 2005*, Maplesoft, 2005, pp. 423-437.

34. A. Wolin, B. Eoff, and T. Hammond, "ShortStraw: a simple and effective corner finder for polylines," in *Eurographics 5th Annual Workshop on Sketch-Based Interfaces and Modeling*, 2008, pp. 33-40.

35. Z. Xu, and J. Luo, "Face recognition by expression-driven sketch graph matching," in *Proc. of the 18th International Conference on Pattern Recognition (ICPR)*, Hong Kong, August 20-24, 2006, pp. 1119-1122.

36. M.S. Nixon, and A.S. Aguado, *Feature Extraction and Image Processing*, St. Louis, Missouri, USA: Academic Press, 2008.

37. C. Hong, X. Ying-Qing, S. Heung-Yeung, Z. Song-Chun, and Z. Nan-Ning, "Example-based facial sketch generation with non-parametric sampling," in *Proc. of Eighth IEEE International Conference on Computer Vision (ICCV)*, Vancouver, BC, Canada, July 7-14, 2001, pp. 433-438.

38. Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, "A nonlinear approach for face sketch synthesis and recognition," in *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, San Diego, California, USA, June 20-25, 2005, pp. 1005-1010.

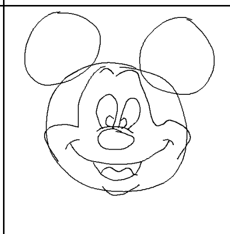
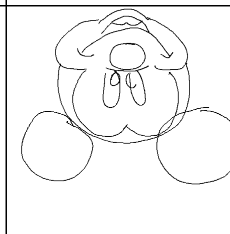
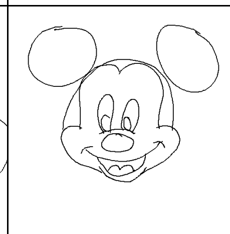
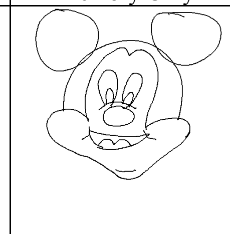
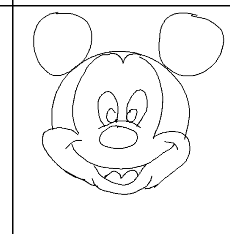
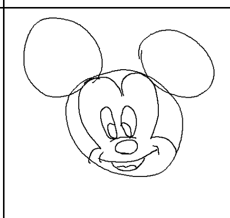
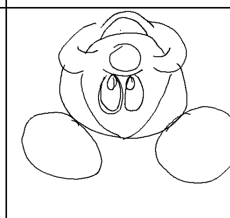
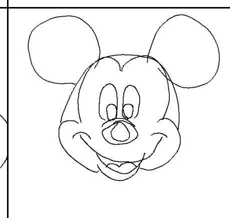
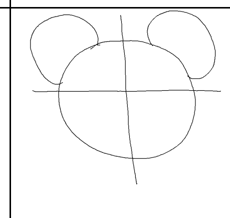
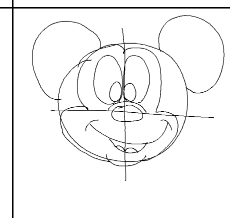
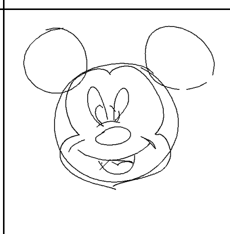
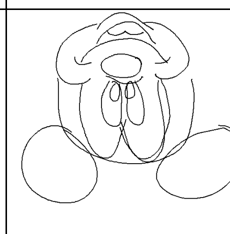
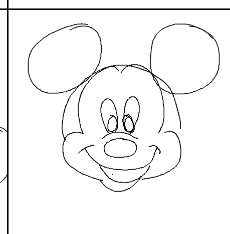
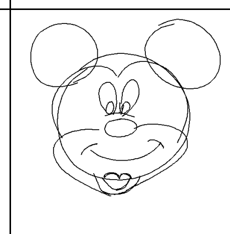
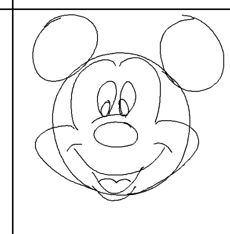
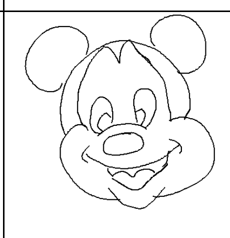
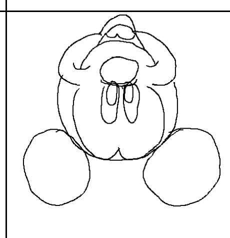
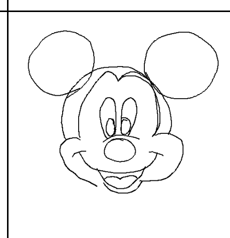
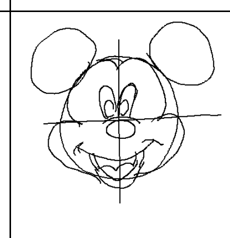
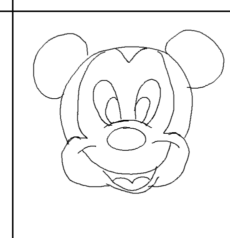
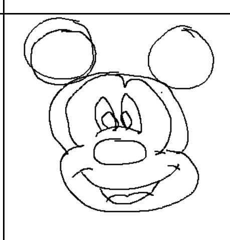
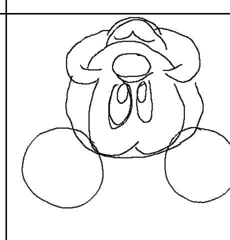
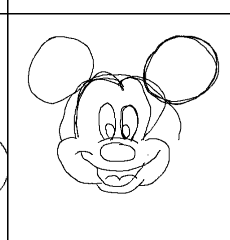
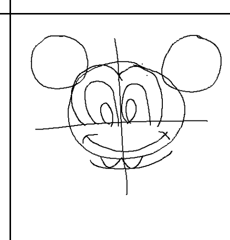
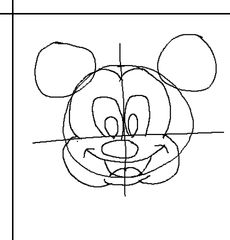
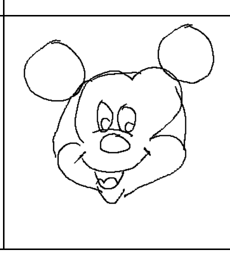
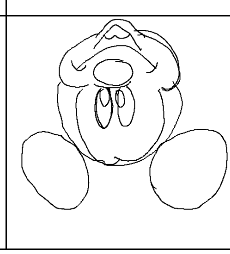
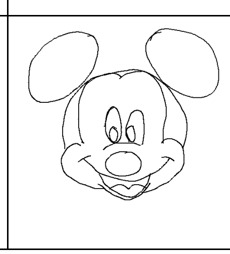
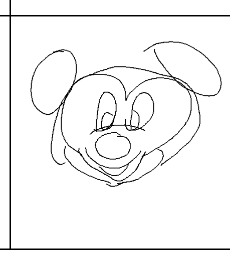
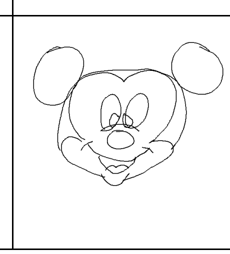
39. G. Xinbo, Z. Juanjuan, L. Jie, and T. Chunna, "Face sketch synthesis algorithm based on E-HMM and selective ensemble," in *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 18, pp. 487-496, 2008.

40. Z. Xu, H. Chen, S.-C. Zhu, and J. Luo, "A hierarchical compositional model for face representation and sketching," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp. 955-969, 2008.

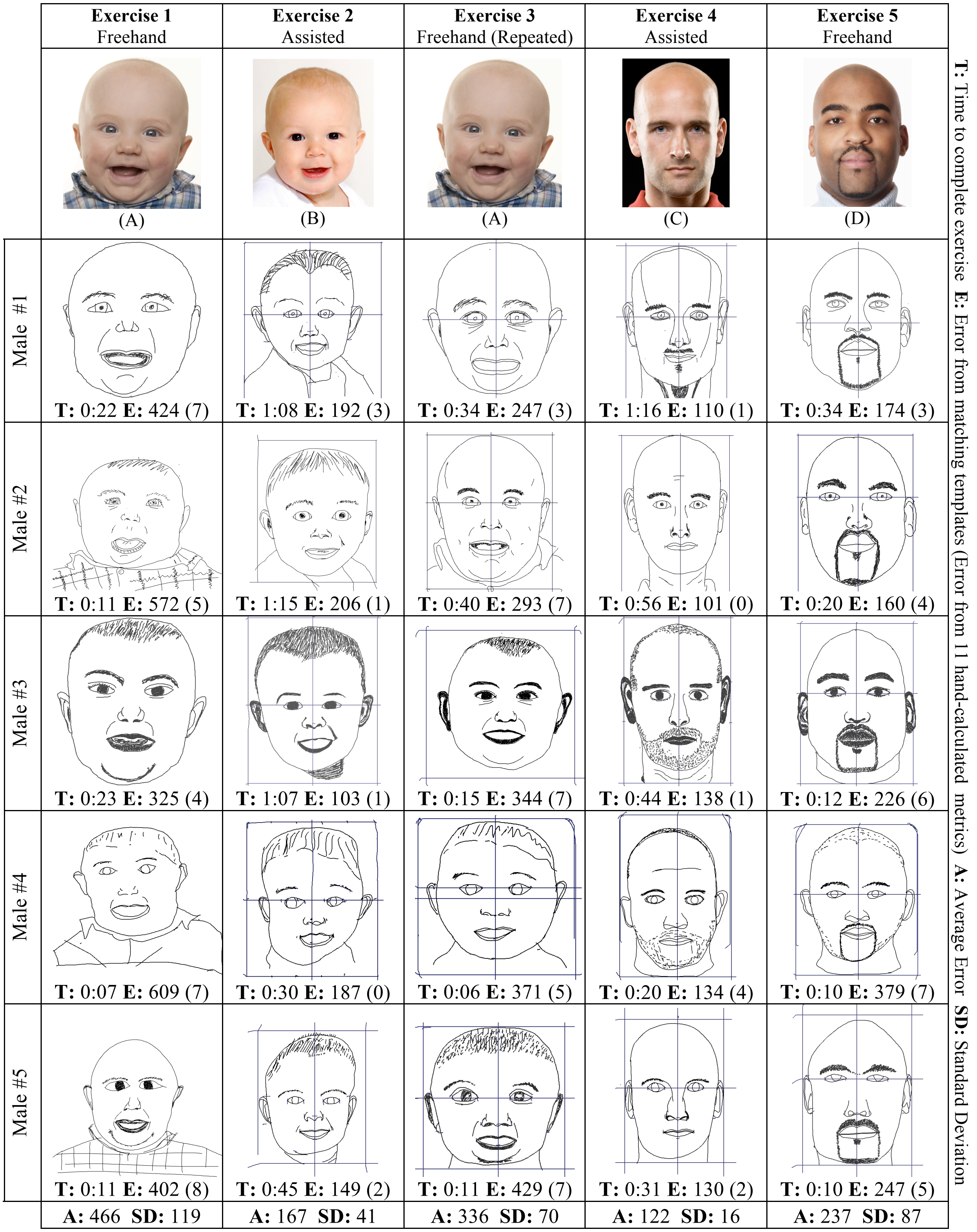
41. W. Jin, B. Hujun, Z. Weihua, P. Qunsheng, and X. Yingqing, "Automatic image-based pencil sketch rendering," in *Journal of Computer Science and Technology*, vol. 17, no. 3, pp. 347-355, May 2002.

42. H. Yee and A. Newman, "A perceptual metric for production testing," in *ACM SIGGRAPH Sketches*, Los Angeles, California, USA, August 8-12, 2004, p. 121.

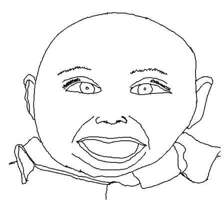
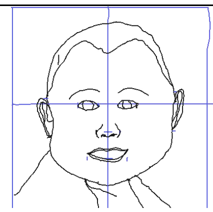
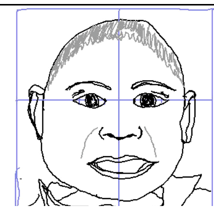
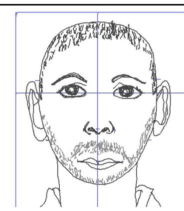
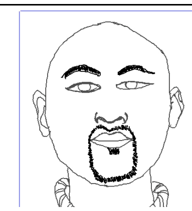
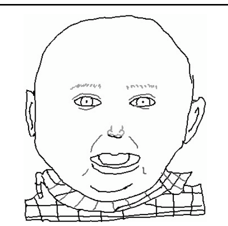
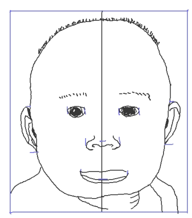
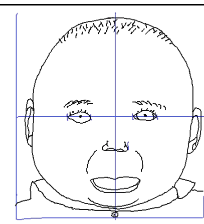
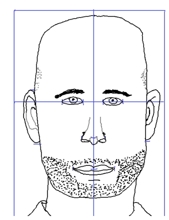
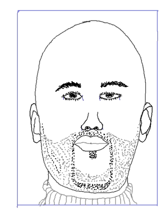
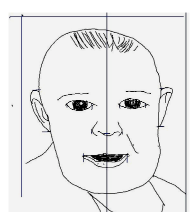
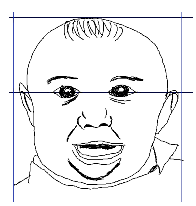
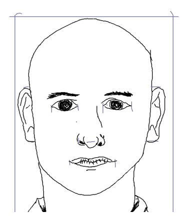
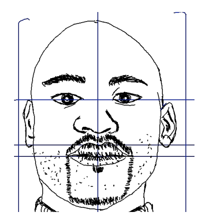
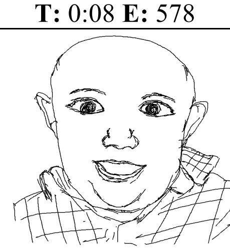
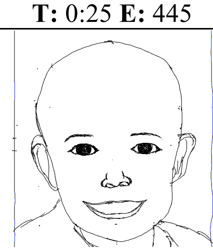
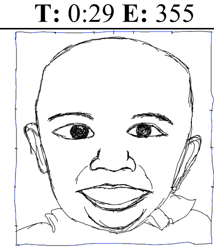
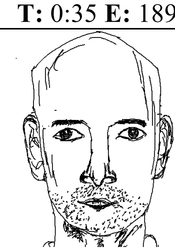
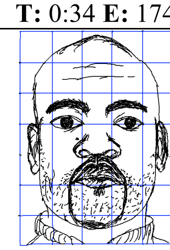
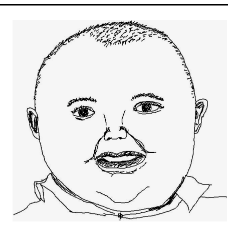
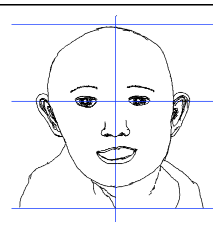
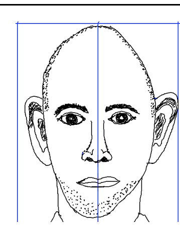
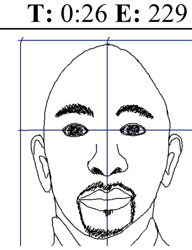
APPENDIX A – FIRST USER STUDY RESULTS

	Exercise 1 Freehand	**Exercise 2** Assisted	**Exercise 3** Assisted	**Exercise 4** Freehand from Memory Only	**Exercise 5** Freehand
Female #1					
Female #2					
Female #3					
Male #1					
Male #2					
Male #3					

APPENDIX B – SECOND USER STUDY RESULTS



APPENDIX C – SUPPLEMENTAL USER STUDY RESULTS

	Exercise 1 Freehand	**Exercise 2** w/ Instructions Only	**Exercise 3** Freehand (Repeated)	**Exercise 4** w/ Instructions Only	**Exercise 5** Freehand
	 (A)	 (B)	 (A)	 (C)	 (D)
Male #1	 **T:** 0:08 **E:** 579	 **T:** 0:28 **E:** 289	 **T:** 0:12 **E:** 473	 **T:** 0:20 **E:** 343	 **T:** 0:12 **E:** 305
Male #2	 **T:** 0:12 **E:** 359	 **T:** 0:18 **E:** 225	 **T:** 0:13 **E:** 374	 **T:** 0:17 **E:** 182	 **T:** 0:20 **E:** 248
Male #3	 **T:** 0:08 **E:** 578	 **T:** 0:25 **E:** 445	 **T:** 0:29 **E:** 355	 **T:** 0:35 **E:** 189	 **T:** 0:34 **E:** 174
Male #4	 **T:** 0:09 **E:** 518	 **T:** 0:44 **E:** 172	 **T:** 0:35 **E:** 261	 **T:** 0:33 **E:** 141	 **T:** 0:26 **E:** 229
Female #1	 **T:** 0:19 **E:** 415	 **T:** 0:34 **E:** 370	 **T:** 0:18 **E:** 322	 **T:** 0:18 **E:** 259	 **T:** 0:16 **E:** 264
	A: 489 **SD:** 99	**A:** 300 **SD:** 109	**A:** 357 **SD:** 77	**A:** 222 **SD:** 79	**A:** 244 **SD:** 48

T: Time to complete exercise **E:** Error from matching templates **A:** Average Error **SD:** Standard Deviation of Error

APPENDIX D – STEPS AND INSTRUCTIONS OF FINAL APPLICATION

Welcome

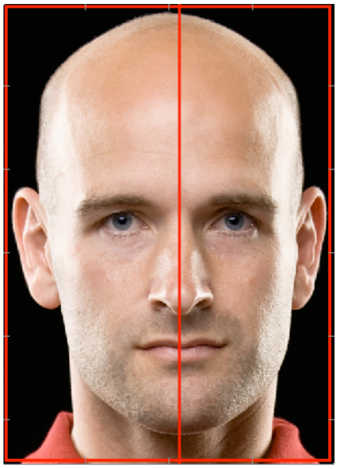
Welcome to iCanDraw?, an application to assist you step-by-step in drawing a human head and face. For each step, the instructions will be here. When you've completed the task, click "Check My Work" for text feedback (go ahead and try now). Sometimes the application will have visual feedback, too. You can see this by using the "Show Suggestions" button, if it's available. Anytime that you feel you're done with a step, click the "Next" button to continue on.

Setup: Format

Before you start on the figure, you will want to draw a "format", or a box, to setup your drawing area and establish proportions. Using the straight edge tools, draw a box of similar proportions (height and width) to the example photo but scaled to the size you want on the paper. Click and drag on a straight edge to place it, then run the pen along it.

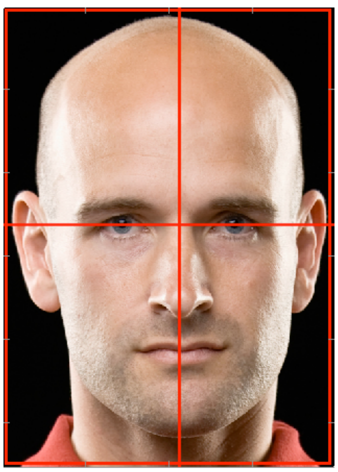
Step 1: Draw the Head

Draw the outline of the head without the ears and inside the format. You might start with a simple oval or egg shape, but remember the contour line exercise and continually refine the shape to look like what is actually there. The line for your shape should turn and move the same as the edge of the head shown in the photo.



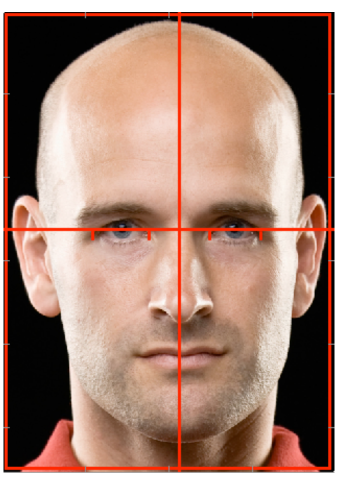
Step 1: Draw the Head

Using a straight edge, draw the central axis (a vertical reference line) down the center of the head. The human head is very symmetrical and the central axis will help establish this on your drawing. Note that it may not be perfectly in the center as the photo may not have been taken straight on. You can use your pencil as a measuring device on the photo to get proportions. Slide your thumb along the pencil to mark a length from your finger to its end. Use this to measure where the central axis on the photo is placed.



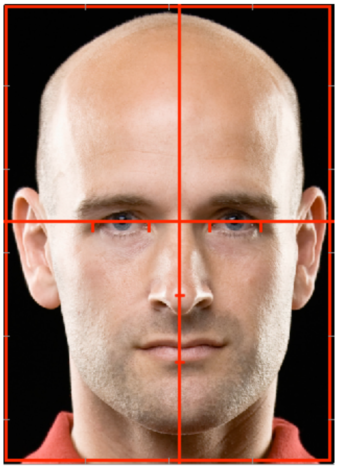
Step 2: Setup Eyes & Nose

Again using a straight edge, draw the eye line, or a horizontal reference line, to divide the head into upper and lower halves. The line will pass through the center of the pupils. The eyes are generally in the middle of head, though a common mistake is to draw them too high and make the forehead too small. Use your pencil to measure this on the photo, or put it against your own face, and measure from eyes to chin and then eyes to the top of the head.



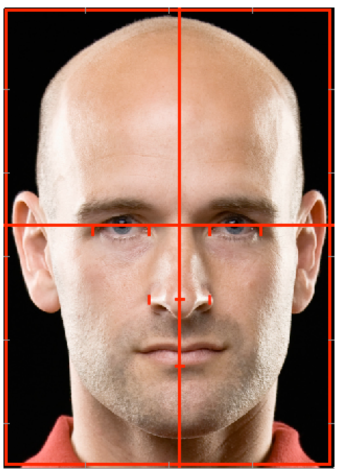
Step 2: Setup Eyes & Nose

To determine where to place the eyes, look at the eyes on the photo. If you measure with your pencil, you'll notice that the distance between the eyes is about the width of an eye itself. When looking straight on, there is usually an eye width between each side of the head and its corresponding eye. Along the horizontal eye line, draw four short reference tick marks for the corners of the two eyes. These should roughly divide the horizontal eye line into fifths. You can also tap on the photo show and hide the guide lines.



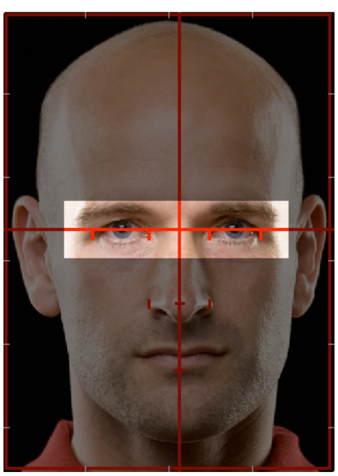
Step 2: Setup Eyes & Nose

Look at where the nose and mouth fall along the central axis in the photo. If you divide the lower portion of the central axis into thirds, the end or the tip of the nose will fall at the end of the first third. The middle or bottom of the mouth should be at the end of the second third. Divide the lower half of your head into thirds and draw two short reference tick marks.



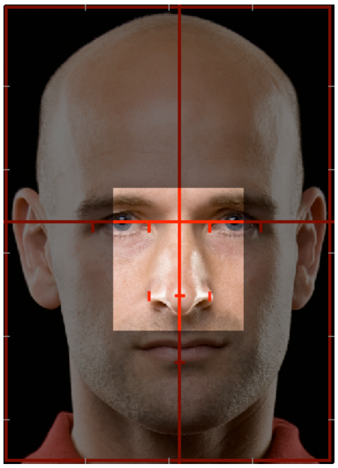
Step 2: Setup Eyes & Nose

Follow the inside corner of an eye straight down on the photo and you should cross the edge of a nostril. Use this same guide to put two short reference tick marks on your drawing for the width of the nose.



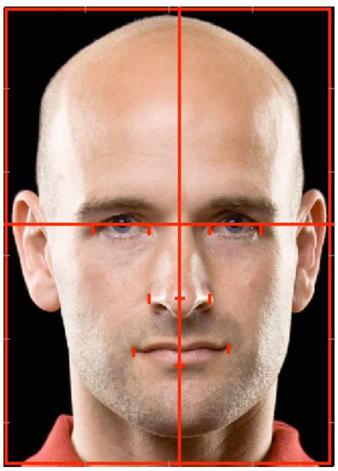
Step 3: Draw the Eyes & Nose

Using the reference lines, draw the eyes just like you see in the photo. What bends and moves does the curve of an eyelid make? What the shape is made by the white space around the pupil? Is the pupil partially hidden by an eyelid? How big are the eyes in comparison to other features? Are the eyes symmetric? Feel free to practice drawing an eye in the space around your drawing.



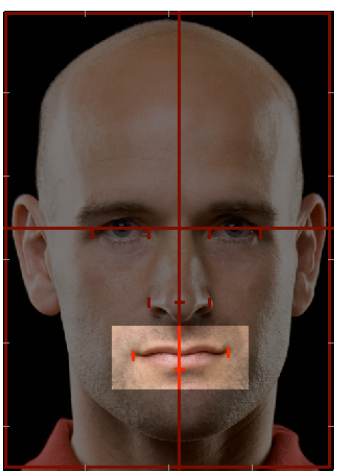
Step 3: Draw the Eyes & Nose

Using the reference lines, draw the nose like you see in the photo. What shape is made by a nostril? Is the nose symmetric? Can you convey what you need without drawing an entire nose, perhaps just the lower third with the tip and nostrils? Again, you can practice drawing a nose in the space around your sketch. The nose is often wider then most people think.



Step 4: Setup & Draw the Mouth

Look at the photo and determine where the left and right edge of the mouth. These edges will probably line up with the pupils of the eyes. Use your pencil as a straight edge on the photo to determine where the mouth starts and ends in relation to the eyes. Draw two tickmarks down from the same points of the eyes for the outside edges of the mouth on our sketch. Make sure these are above the previous tickmark for the mouth. Many people will tend to the draw the mouth too low and too compact.

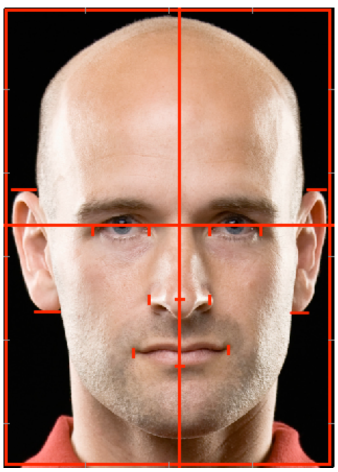


Step 4: Setup & Draw the Mouth

Using the reference lines, draw the mouth just as you see it in the photo. What turns and moves does the curve of a lid make? What is the height of the mouth in comparison to other features? Again, use the extra space to practice drawing the mouth.

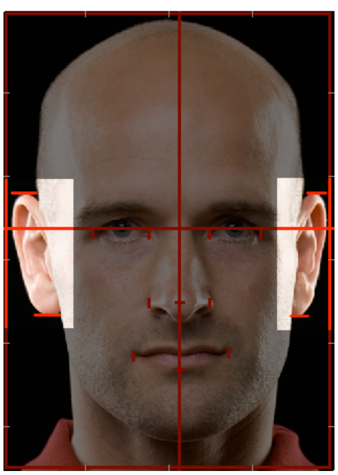
Step 5: Draw the Eye Brows

Draw the eyebrows just as they are in the photo. Where do the eyebrow start and end? How thick is the eyebrow? Are the eye brows symmetric?



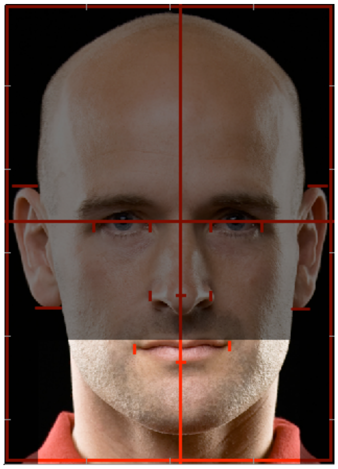
Step 6: Setup & Draw the Ears

Now setup for the ears. If you follow along a horizontal line from the top of the eyes, you should come close to the top of the ears. Mark the tops of the ears with short reference tick marks. Starting at the bottom of an ear, follow a similar horizontal line and you should end up in the area between the nose and the mouth. Mark the bottom of the ears with short referernce tick marks. People often draw ears too small.



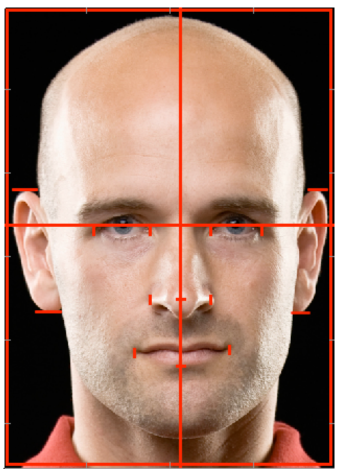
Step 6: Setup & Draw the Ears

Using the reference lines, draw the ears on each side just as you see them in the photo. How far does the ear protrude? What lines do the inner folds make? Are the ears symmetric? Use extra space to practice if need be.



Step 7: Draw the Neck

Draw the neck just as it is on the photo. How wide is the neck? How long is it? What other features does it line up with? You might feel along the side of your own neck and face to get an idea, too. Many people will draw the neck too thin when it should be almost as wide as the jaw or even wider for males.



Step 8: Add Details

For the last step, add what details that you would like to. Does the person have hair in the photo? Unique facial features? Could more detail be added to existing features?

VITA

Daniel Meyer Dixon received his Bachelor of Science in software engineering from Auburn University in December 2006. He received his Master of Science in computer science from Texas A&M University in December 2009. His plans for the future are to continue to learn what it means to love God and be obedient to Him.

Mr. Dixon may be reached at Richardson Building, Room 911, 3112 Texas A&M University, College Station, TX 77843. His email is danieldixon@acm.org.