PhotoCropr
A first step towards computer-supported automatic generation of photographically interesting cropping suggestions.

by

Evan Golub
Department of Computer Science
Human-Computer Interaction Lab
University of Maryland
egolub@acm.org

ABSTRACT
In the age of digital photography, post-processing (for better or for worse) has become a more common activity. Actions such as red-eye removal, adjusting levels and colors, and cropping are "part of the routine" for many users. When cropping a photograph, a variety of factors can come into play, including basic aesthetics. Tool-based support for a simple guideline such as the Rule of Thirds could provide many photo enthusiasts with a useful tool in their editing arsenal. This work motivates that idea, and explores it both within the context of mostly-manual "assisted" editing as well as the application of automated techniques.

1.0 INTRODUCTION

The motivation for this project came from my own experiences in editing photographs. I often take pictures that have "extra" room on all four sides so that I can catch the shot I want when I see it rather than take the time to try to frame it correctly and risk missing the shot, and then frame it correctly at a later date. I would typically place what I considered to be the image's center-of-interest to be off-center (ironically) [11]. What I found was that in doing this, I was often obeying the Rule of Thirds [14] in composing the cropped version of my photographs, and I began to long for a tool built to support this behavior.

I have built a program I call "PhotoCropr" to explore the idea of semi-automated image cropping. The current version of **PhotoCropr** uses the spirit of the Rule of Thirds and the Golden Mean [16] to generate an array of cropping suggestions for the user. These and other rules will be described later in the paper. The user needs only to click on what they consider to be the point-of-interest in their photograph (often the middle point of the object considered the center-of-interest in the image), and the application then presents either 4 or 6 previews of a cropped version of their original image. These can either be based on the Rule of Thirds, or the Golden Mean. Once presented with the preview images, the user can then change the zoom level of the cropping suggestions, and select one preview to be saved as a full-sized cropping.

The general idea of presenting multiple preview images of the croppings, and then allowing the user to affect them all in parallel (in this case by altering the zoom level) was inspired by the **SideViews** [15] project by Elizabeth Mynatt and Michael Terry at Georgia Tech. In my own

cropping pursuits, I often found myself comparing several different cropped versions on an image side-by-side after the fact, and remembered seeing their video from UIST 2002 showing that you could support this behavior within the selection process itself. The resulting interface appears to bear this out.

The long-term goal would be to eventually use a variety of wider composition rules, along with image processing techniques such as facial recognition [4], torso recognition [17], saliency maps [9], and warping [13] to generate a robust set of possible croppings from which a novice photographer could select.

2.0 THE RULE OF THIRDS

Making an interesting photograph can be a matter of having an interesting subject, composing a scene in an interesting manner, or a combination. For composing a scene in an interesting manner, one suggestion is to follow the Rule of Thirds. To accomplish this, you imagine a pair of vertical lines, each $\frac{1}{3}^{rd}$ away from the sides of the image, and a pair of horizontal lines, each $\frac{1}{3}^{rd}$ away from the top and bottom of the image. You then try to do one of the following:
(a) place your center-of-interest on a point where two of the lines intersect (I refer to these points as the crosshairs of the Rule of Thirds), (b) place a vertical object along a vertical line, (c) place a horizontal object (or border such as the horizon) along a horizontal line.

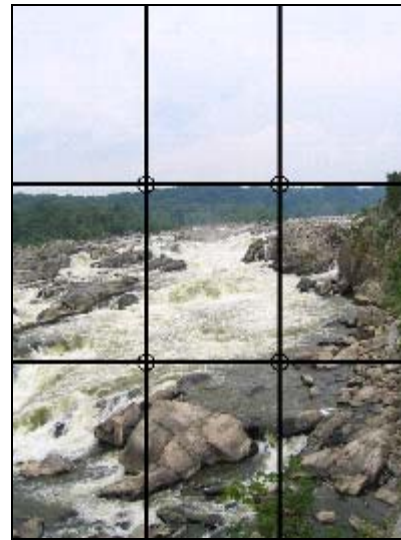Figure 2-1.1: Original that "obeys" the Rule of Thirds by placement of the horizon.

Figure 2-1.2: Rule of Thirds lines superimposed on photograph.

It is important to note that the decision of how to use this rule of thumb is very dependant on the image itself. There might be more than one object that could be considered the center-of-interest. There might be several vertical or horizontal elements. Also, for any given center-of-interest, the context and content might affect upon which crosshair you would want to place that item. For example, if you had an image of a dog, and the dog was looking from left-to-right, an observer of the image might feel that the use of the SW crosshair would be right, while if the dog was looking from right-to-left, the same observer might prefer the use of the SE crosshair.

It is because of issue such as these that this project endeavors to present multiple cropping candidates, shown simultaneously, from which the user may select the one that feels most appropriate within the given context.

3.0 THE GOLDEN MEAN AND GOLDEN SPIRAL

A similar rule of thumb for image composition is the Golden Mean. As was the case with the Rule of Thirds, to use the Golden Mean you imagine a pair of vertical lines and horizontal lines on which to place the center-of-interest. However, the placement of these lines differs. In this case, you partition the image vertically and horizontally using the Golden Ratio [10]. If you divide a line into two sections (a large one of length **a** and the other of length **b**) using the Golden Ratio, the ration between **a** and **b** is that same as between **a+b** and **a**. We can use this ratio to create the vertical and horizontal lines. You can see a comparison between the Rule of Thirds lines and the Golden Mean lines in Figures 3-1 and 3-2. Again, for interesting composition, you can attempt to place the center-of-interest on a point where two of the lines intersect, or place a vertical object along a vertical line, or place a horizontal object along a horizontal line.
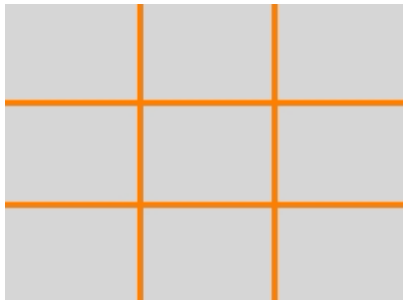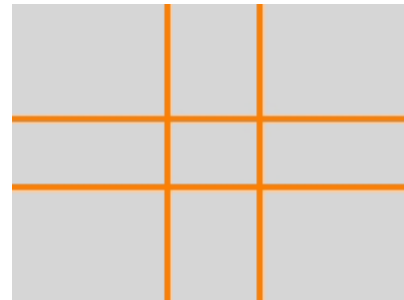

Figure 3-1: Rule of Thirds lines.

Figure 3-2: Golden Mean lines.

The Golden Ratio can also be used recursively to divide a rectangle into smaller pairs of rectangles that follow this ratio (Figure 3-3). Drawing 90° arcs in the larger region of each pair leads to the formation of a spiral pattern known as the "Golden Spiral" (Figure 3-4).
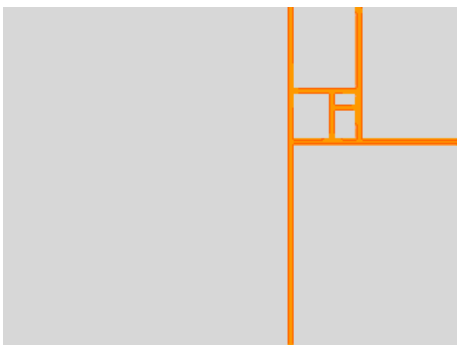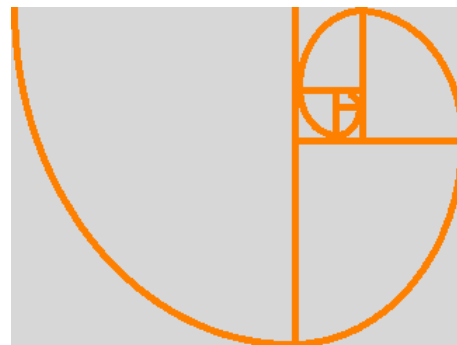

Figure 3-3: Recursive golden ratio divisions.

Figure 3-4: The Golden Spiral filled in.

For the sections that follow in this paper, PhotoCropr will be discussed, and the Rule of Thirds will appear in the various examples. However, in each case, the placement of the point-of-

interest on a crosshair of the Golden Mean, or on the origin of the Golden Spiral is supported as well. The user can select which of the rules to use in their composition via the **Settings** dialog under the **File** menu.

4.0 THE PHOTOCROPR ALIGNMENTS

There are currently two sets of previews available to the user, in each of Rule of Thirds and Golden Mean composition. The first collection is a set of four previews of possible croppings created using the four crosshairs, created by the vertical and horizontal lines, to position the indicated point-of-interest (Figures 4-1 through 4-4). The second collection is the same set as the first, augmented with two central croppings; one where the point-of-interest appears centered horizontally (Figure 4-5), but on the upper horizontal line, and one where the point-of-interest has been fully centered in the frame (Figure 4-6). With both variants, the user is able to adjust the zooming level for the cropping while keeping the point-of-interest in the same relative position in the frame.
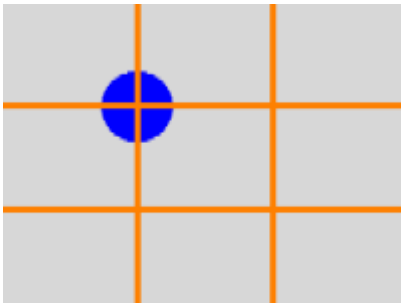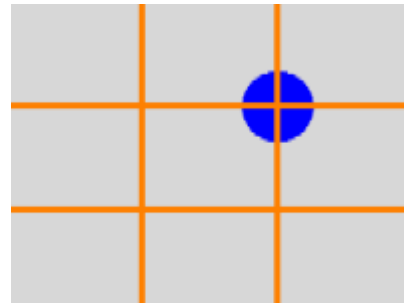
Figure 4-1: On NW Rule of Thirds crosshair.

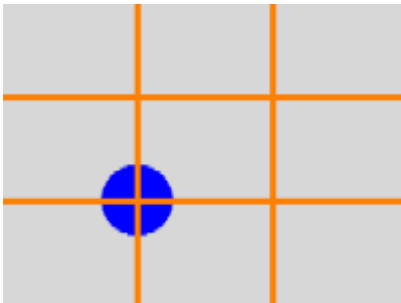Figure 4-2: On NE Rule of Thirds crosshair.
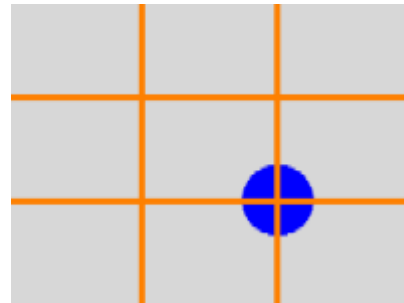
Figure 4-3: On SW Rule of Thirds crosshair.

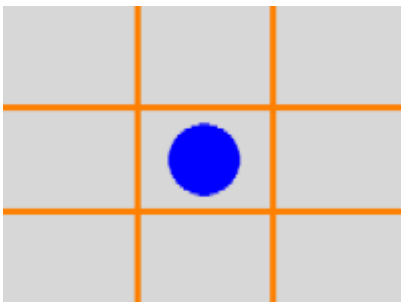Figure 4-4: On SE Rule of Thirds crosshair.
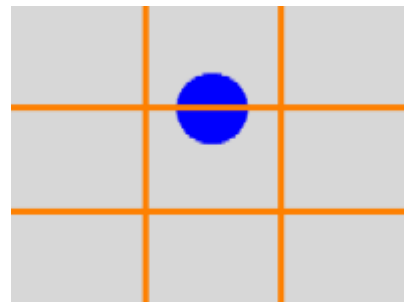
Figure 4-5: Dead Center in frame (CC).

Figure 4-6: Central upper Rule of Thirds line (CT).

The notion for introducing the two central croppings came from a combination of informal user studies performed by Steve Wass, an undergraduate student at Maryland, as part of a general exploration of user croppings and the Rule of Thirds, along with a photographic tip to place things at eye-level. Based on several participant comments, there appeared to be a trend towards desiring the image's center-of-interest to appear in the center of the frame (Figure 4.6). For a photograph with a person whose face is likely to be selected as the center-of-interest by the user, my own exploration with some of my own photographs led me to add the 6th cropping in which the center-of-interest is placed on the upper horizontal line, and is centered horizontally.

5.0 USING PHOTOCROPR

To use **PhotoCropr**, the first step is to load an original (preferably high-resolution) image into the program (Figure 5-1). This can be done in one of two ways; either by dragging an image to the application's canvas, or by opening a file via the **File** menu.

Once the image has been loaded, the user clicks on their point-of-interest within the image with their mouse.

Figure 5-1: An image that the user wants to crop based on the position of a particular point-of-interest loaded into PhotoCropr.

Once the point-of-interest is selected, the application will present previews of possible croppings, one for each of the four possible placements of the point-of-interest on an intersection point (Figure 5-2), and in the case of the second option, one for each of the two center-based placements as well (Figure 5-3).

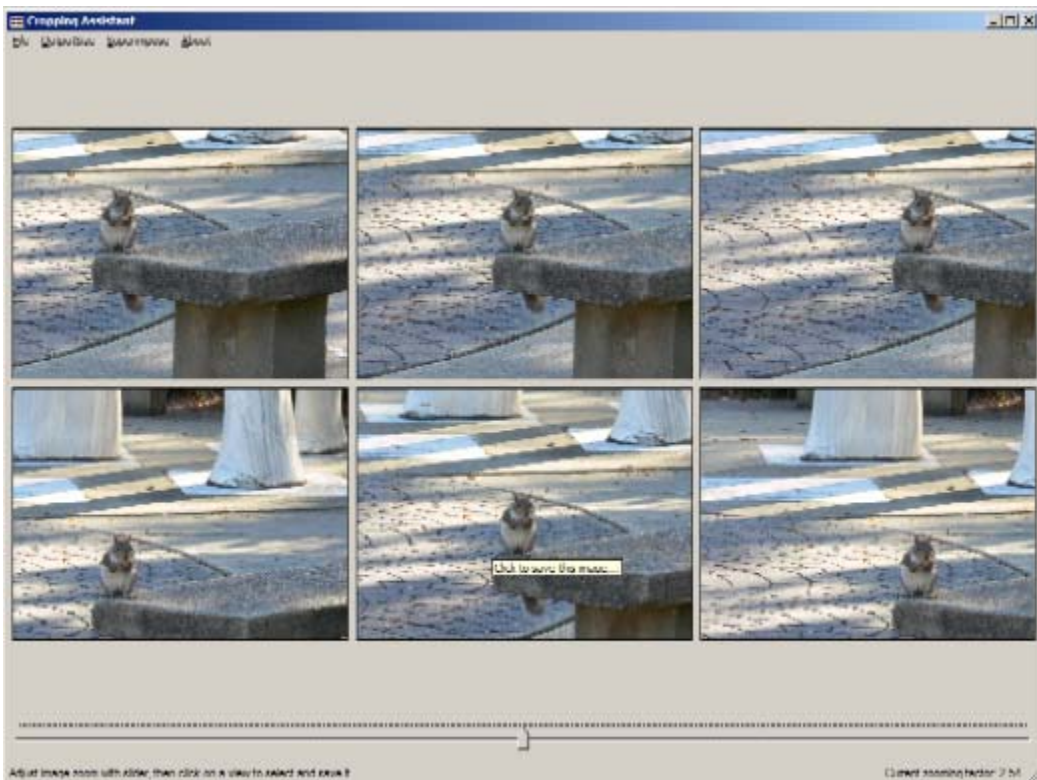Figure 5-2: The four preliminary Rule of Thirds croppings.

Figure 5-3: The Rule of Thirds croppings with the two centralized croppings.

Once the point-of-interest has been selected and the preliminary croppings presented, the user can move the slider below the images to alter the zoom level of the images (Figures 5-4 and 5-5). This will maintain the position of the point-of-interest while changing the framing of the shot.

If after selecting the point-of-interest, the user is not satisfied with the croppings, they can revert to the original image (Control-R) and select a new point.

Figure 5-4: View after zooming in some more relative to the original previews.

While adjusting the zoom level for the previews, the user is able to view the previews with the Rule of Thirds or Golden Mean lines superimposed by going to the **Superimpose** menu, and checking **Show Lines**. If that option is checked, whenever the user does not have the mouse positioned over the zoom adjustment trackbar, the lines will appear on top of the previews. The color and width of those lines can be adjusted under the **Superimpose** menu as well.

Once the user is satisfied with one of the cropping previews, they simply click on that preview to save it at the selected output size. The user can select (via **Settings** under the **File** menu) for a save dialog to appear, pre-populated with the original filename appended with a cropping code (NW, NE, SW, SE, CT, CC) or for the indicated image to be automatically saved using the original filename appended with the two-letter cropping code. The user is also able to "drag-and-drop" from a preview image to specify a destination for the full-size version of that cropping.

Several common output sizes are available in 4:3 and 3:4 aspect ratios, one common output size is available in each of 3:2 and 2:3 aspect ratios (for traditional 4″ by 6″ photo printing), and three output sizes are available relative to the source image's original aspect ratio. The orientation (portrait or landscape) of the original image us used to determine the initial orientation of the cropped version, but the user can switch this orientation (as well as the actual size of the saved file) at any time via the **OutputSize** menu.

Figure 5-5: View after zooming out more relative to the original previews.

Notice that in three of the four previews in Figure 9, there is not "enough" of the original image available to fill the frame of the cropping so a rose-colored edging is displayed. If one of these croppings is selected, the "out-of-bounds" area will actually appear black in the output file. The user may select any color they wish to indicate the "out-of-bounds" area via the **Settings** dialog under the **File** menu.

6.0 AUTOCROPPING

To use the current version of **PhotoCropr**, after loading the image, the user needs to select a point of interest. As has been mentioned, work has been done to use (for example) saliency maps to perform automatic thumbnail cropping. We have begun to explore using the work of the UIST algorithm [9] not to create the full cropping, but rather to automatically select the point of interest. The goal of this would be to automate yet another aspect of cropping for aesthetics. The user could select "auto position" and the UIST algorithm would be executed, and the center point of the UIST cropping would be used as the center-of-interest for **PhotoCropr** to the use to

present cropping options according to the Rule of Thirds, Golden Mean, and Golden Spiral or even (in future versions) the related principles of the Golden Diagonal and Golden Triangle [3].

To explore the feasibility of this option, we took 25 source images, and cropped them each in 5 different ways. In each cropping, what we considered to be the point of interest was placed on one of the 4 rule of thirds crosshairs, or centered on the frame. Using these 100 source images, we then used the UIST algorithm, and other researchers crop them into thumbnails. We then compared the UIST crops to the manual crops. While the widths and heights varied quite a bit, comparing the center points revealed some interesting patterns.

Figure 6-1 shows an image-by-image chart of the (scaled) differences between the center point of the UIST croppings to that of one of the researchers' croppings.
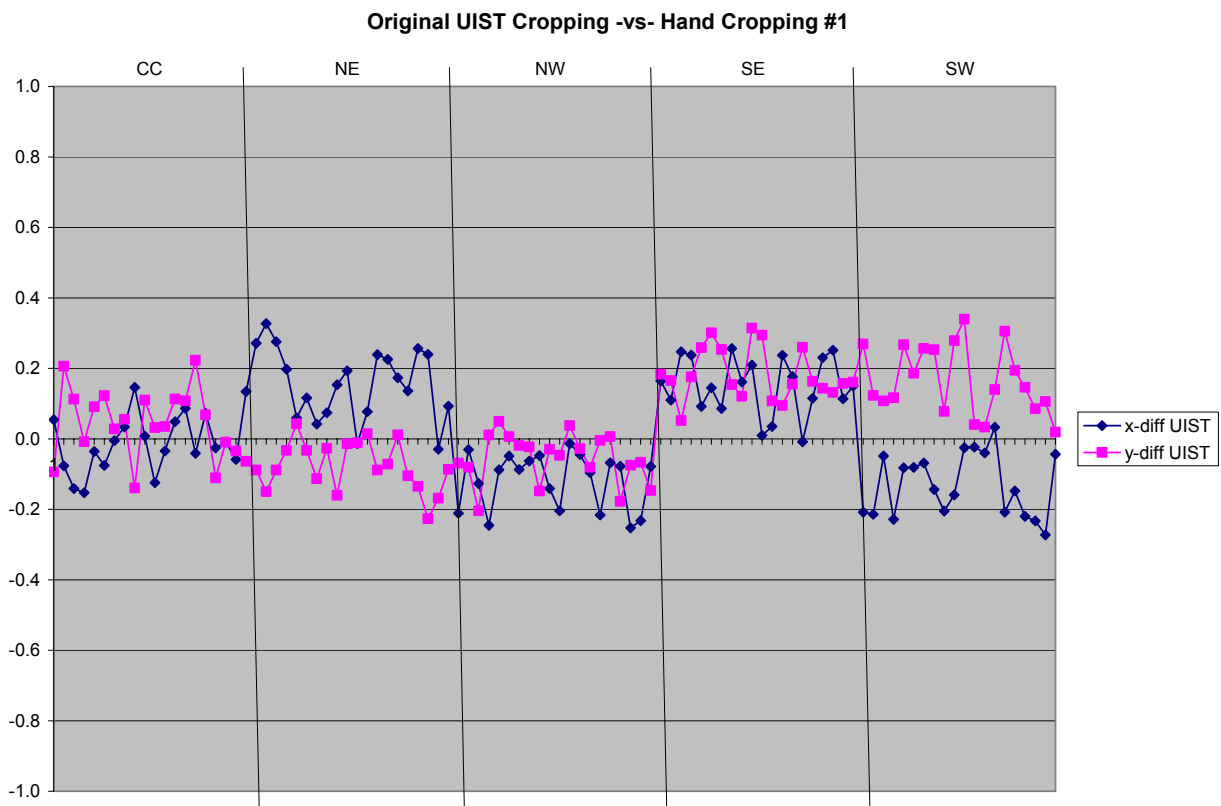

Figure 6-1: Chart of the offset between the (x,y) coordinates of a manually cropped thumbnail and thumbnail generated by the original UIST saliency algorithm, clustered by position of center-of-interest in the original image.

While the worst case cropping had the y-coordinate differ by almost 34% of the image's height and the x-coordinate differ by almost 33% of the image's width, the best cases for both were differences of under 1%, and the average y-coordinate difference was under 12% and the average x-coordinate difference was under 13%. What we found was the although the UIST algorithm tends to favor towards the center of the original image, even when what we perceived as the point of interest was far from the center.

To explore this, an image of the moon was modified so that it was positioned in different locations. Thumbnails were then generated using the original UIST saliency algorithm. The example images and corresponding thumbs created by the UIST algorithm shown in Figure 6-2 partially demonstrated the tendency towards the center of the original image.

Moon Centered

Moon in Center, High

Moon in North-West Quadrant

Moon in North-East Quadrant

Moon in South-West Quadrant
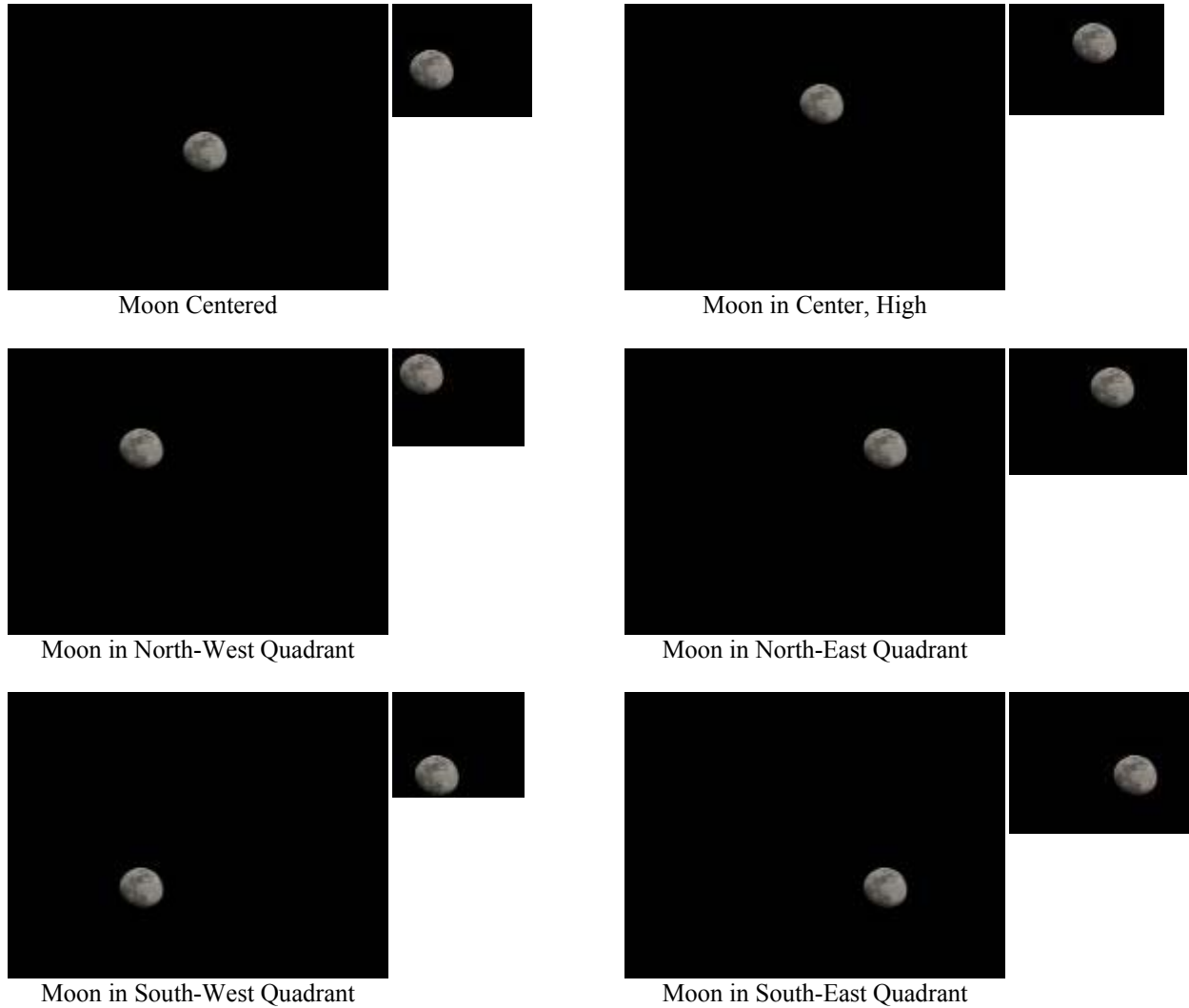

Moon in South-East Quadrant

Figure 6-2: The same image of the moon, positioned in different locations and the corresponding thumbnails generated by the original UIST saliency algorithm.

This was due to the original implementation's [12] assumption that the center of the original image would include something of interest, and the resulting requirement that the center point of the original image must appear in the cropped version.

Without mandating that any particular point had to be kept in the cropped version, some of the results when cropping the same 6 lunar images were better, as can be seen in Figure 6-3. Notice that for the North-West, South-West, and South-East cropped versions, the center is now almost perfect.

Moon Centered

Moon in Center, High

Moon in North-West Quadrant

Moon in North-East Quadrant

Moon in South-West Quadrant
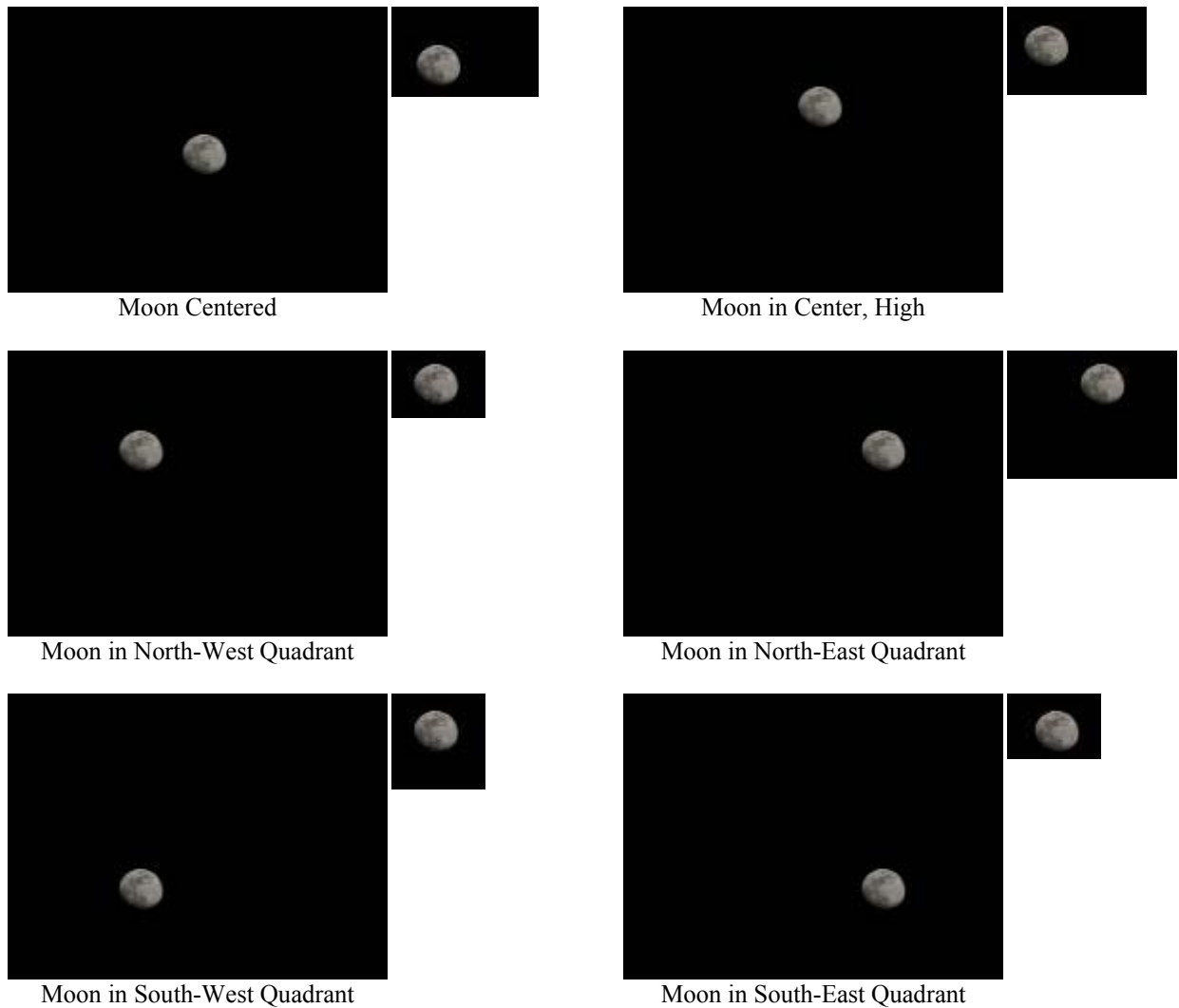
Moon in South-East Quadrant

Figure 6-3: The same images of the moon and thumbnails generated by the modified UIST saliency algorithm.

Running the original set of images through this modified algorithm, we obtain the results shown in Figure 6-4.   With this modified algorithm, which starts at the center of saliency and works its way out, the results were slightly different, but on average very similar. While the maximum difference for an individual image's x-coordinate selection was around 5% less, there was essentially no difference for the maximum error for the y-coordinate, nor for the average of all offset differences.

It is interesting to note that if you use the average of the x-coordinate values and of the y-coordinate values given by the two UIST-based algorithms, the average distances from the manual cropping selections is slightly less, though not significantly. This direction will be revisited shortly.

Figure 6-4: Chart of the offset between the coordinates of a manually cropped thumbnail and thumbnail generated by the first modification of the UIST algorithm, clustered by position of center-of-interest in the original image.

Hazem El-Alfy, who generated the cropping data, noticed that the original UIST algorithm looked to keep a certain amount of the overall saliency within the cropping, and explored what would happen if only half as much overall saliency was required in the cropping since we were only interested in the center point. Figure 6-5 shows these results.
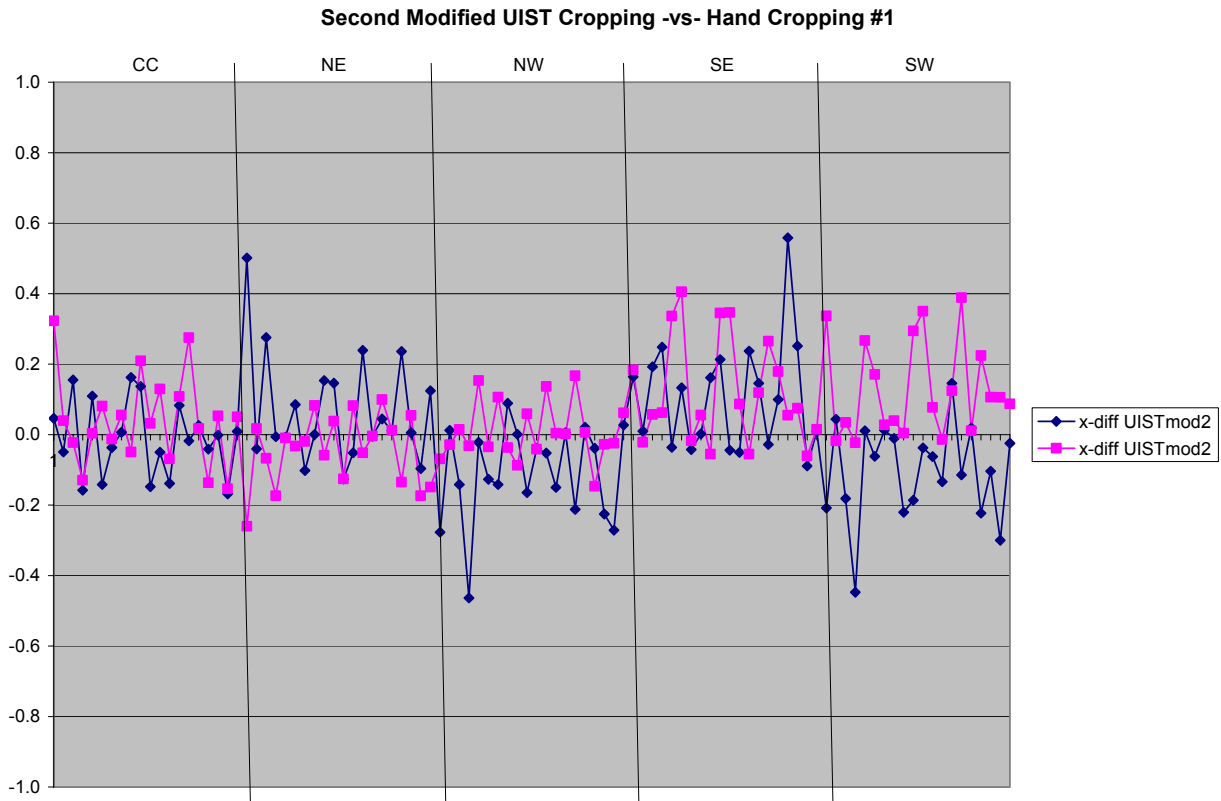
Figure 6-5: Chart of the offset between the coordinates of a manually cropped thumbnail and thumbnail generated by the second modification of the UIST algorithm, clustered by position of center-of-interest in the original image.

In the best cases, things are still very close to the manual cropping, and in the average case the offsets improve by around 1%, but in the worst cases the difference in the x-coordinate increases to almost 56% and the difference in the y-coordinate increases to just over 40%. This can be seen in the high and low spikes.

However, as earlier, some interesting results came from taking the average of multiple automated results. Figure 6-6 shows the results of taking the average of the centers of all **three** UIST-based croppings. The average differences are basically the same as those obtained with El-Alfy's modification, but the maximum differences do not suffer the same penalty. However, the maximum differences on both axis is around 34%, so still problematic. The positive aspect of these last results is that it appears that it might be possible that doing multiple passes over the image with the UIST algorithm could balance out the biases observed above. There is, however, a general limiting factor that might keep this feature out of the PhotoCropr application. The process of generating the UIST cropping (which would have to be done to obtain the new center-of-interest even though we would not use the actual cropping) currently takes 5-10 seconds. This is currently being tested on a duo-core processor machine - on many of the target users' computers, it could take even longer. Considering the ease with which a user can currently select a center-of-interest manually, this delay might be difficult to justify or motivate in **PhotoCropr**. It could, however, still be reasonable as part of the final **AutoCropr** concept.

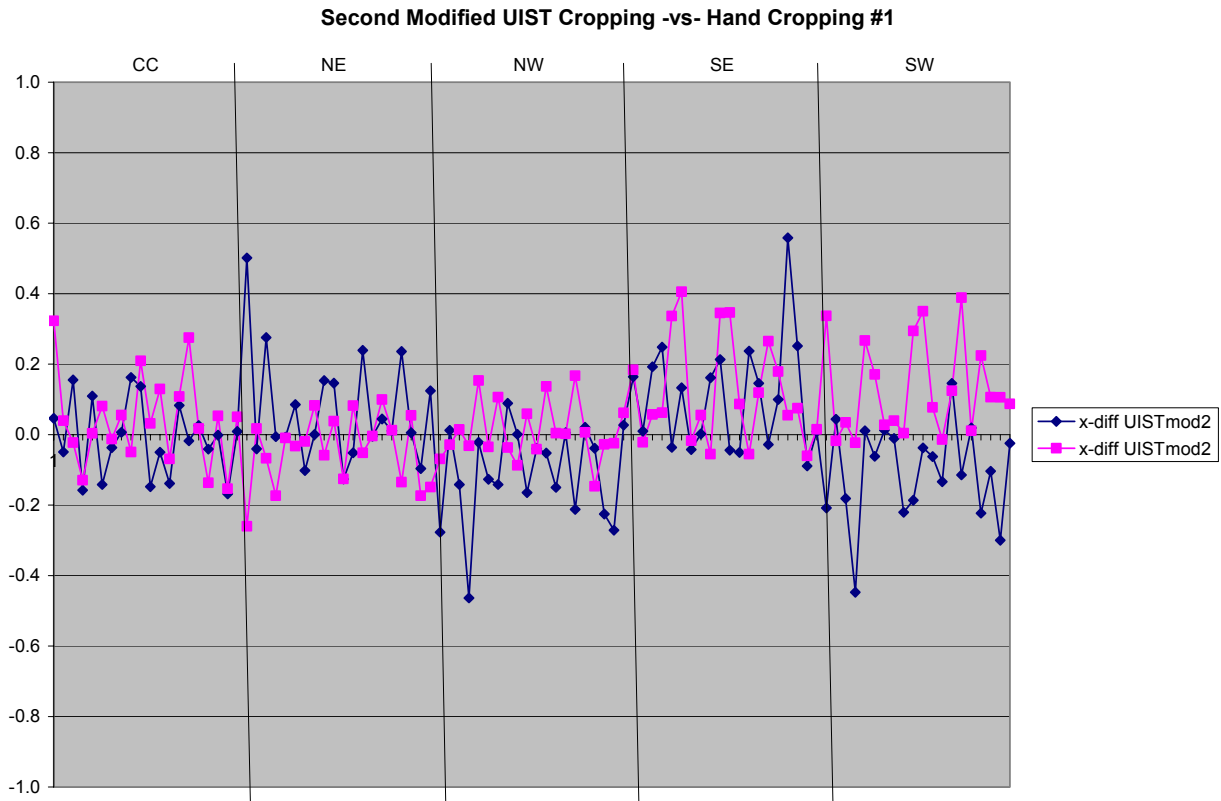**Second Modified UIST Cropping -vs- Hand Cropping #1**

Figure 6-6: Chart of the offset between the coordinates of a manually cropped thumbnail and thumbnail generated by the average of all three UIST-based algorithms, clustered by position of center-of-interest in the original image.

7.0 FUTURE USER STUDY

The next stage of this project will be to conduct a set of user studies primarily to collect data about cropping preferences. Information will be gathered using a set of images that contain various categories of content (e.g.: single person in nature, group of people in nature, posed photograph). The goal will be to determine two things. First, whether there is a particular rule and/or cropping alignment(s) within a rule, which is/are preferred most often for specific types of images. Second, whether there is a pattern to the way in which participants alter the zoom level while composing their cropping.

The results of these studies will be used to inform presentation selections in future projects that involve more automated generation of cropping options. In future iterations of the project, it is anticipated that once the point-of-interest is selected (either by the user, or by automated techniques discussed in the next section), an array of cropping options will be presented to the user automatically.

One possible down-side to having several options presented simultaneously is that the screen's resolution is being partitioned between multiple previews, causing each to appear smaller. On a screen with 1024x768 resolution, each image does appear (to me) rather clearly within the scope of the task at hand. However, the user study will seek to collect broader information about user preferences in this regard. This will be useful in the further design of the current tool, as well as

future versions. As additional composition rules, or a mixture of rules and zoom levels are presented simultaneously in later versions of this tool, the display space might not be sufficient to provide the desired level of detail in a preview.

If the small preview size does appear to be a hindrance to users, there are two possible solutions. One is to incorporate an idea used in **PhotoMesa** [1] (a suggestion attributed to Mark Stefik from Xerox PARC) of temporarily displaying an enlarged preview of any individual cropping by mousing-over that preview. Another is to enlarge one preview while shrinking the remaining preview images (a non-distorting variation on the fisheye effect [5][2]).

8.0 FUTURE RESEARCH

This is hopefully an initial stage in a project to automatically generate croppings without the user being required to manually identify the point-of-interest in an image. Our preliminary results with the UIST algorithm show some promise, but more examples need to be evaluated, both by researchers and with users. There are several different, but related, issues for this future work. Additionally, another direction to explore in an attempt to automate the process of selecting a point-of-interest in a photograph could be to detect any face or faces in the image, and then use the center of those as possible points-of-interest. A related approach would be to detect torsos in the image and align these on one of the Rule of Thirds lines.

Another direction to pursue is to look to identify the category of the content of the image (such as a landscape versus an activity) and then select the most appropriate composition style for that category based on previous user preferences.

One final avenue of exploration would be to attempt to warp an image in order to cause it to "fit" one of these rules. For example, while a nautilus shell generally follows the spirit of the Golden Spiral (Figure 10.1), a particular shell might not fit the curves exactly (Figure 10.2).

Figure 10.1: Image of a nautilus shell from http://tolweb.org/tree?group=Cephalopoda

Figure 10.2: The Golden Spiral superimposed on the image of the nautilus shell.

Liu and Gleicher used non-linear warping to explore displaying images on small devices while minimizing visual distortion. Similar techniques might serve our goals here as well.

REFERENCES
[1]   Bederson, B.B., "PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps." Proceedings of the 14th annual ACM symposium on User Interface Software and Technology, Pages 71-80, 2001.

[2]   Bederson, B.B., Clamage, A., Czerwinski, M.P., and Robertson, G.G., "A Fisheye Calendar Interface for PDAs" ACM Transactions on Computer-Human Interaction, 11(1), Pages 90-119, 2004.

[3]   Elam, K. "Geometry of Design: Studies in Proportion and Composition" New York: Princeton Architectural Press, 2001.

[4]   Frischholz, R., "The Face Detection Homepage" http://www.facedetection.com.

[5]   Furnas, G.W., "Generalized Fisheye Views" Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Pages 16-23, 1986.

[6]   Golub, E., "Composition Rules Visualizer" http://www.cs.umd.edu/~egolub/PhotoApps/CRV.

[7]   Golub, E., "PhotoCropr" http://www.cs.umd.edu/~egolub/PhotoApps/PhotoCropr.

[8]   Hofstadter, D., "Gödel, Escher, Bach: An Eternal Golden Braid" Vintage Books, 1979.

[9]   Itti, L., Koch, C., and Niebur, E., "A Model of Saliency Based Visual Attention for Rapid Scene Analysis." IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11), Pages 1254-1259, 1998.

[10]   Knott, R., "The Golden Section – The Number and its Geometry" http://www.mcs.surrey.ac.uk/Personal/R.Knott/Fibonacci/phi.html.

[11]   Kodak, "Composing Your Picture" http://www.kodak.com/global/en/consumer/film/tips/tips_composing_en.jhtml.

[12] Ling, H., "Smart Thumbnail Clipping."
      http://www.cs.umd.edu/~hbling/Research/ISFall02/ISReport.pdf.

[13] Liu, F. and Gleicher, M., "Automatic Image Retargeting with Fisheye-view Warping."
      Proceedings of the 18th annual ACM symposium on User Interface Software and
      Technology, Pages 153-162, 2005.

[14] Miotke, J., "Rule of Thirds and Photography"
      http://www.betterphoto.com/exploring/tips/thirds.asp.

[15] Mynatt, E. and Micahel, T., "SideViews: Persistent, on-demand previews for open-ended
      tasks." Proceedings of the 15th annual ACM symposium on User Interface Software and
      Technology, Pages 71-80, 2002.

[16] Shutterfreaks Team, "Rule of Thirds and Golden Mean Actions"
      http://www.shutterfreaks.com/Actions/RuleOfThirds.php.

[17] Suh, B. and Bederson, B.B., "Semi-Automatic Image Annotation Using Event and Torso
      Identification." HCIL Technical Report Series, HCIL-2004-15, Computer Science
      Department, University of Maryland, College Park, MD, 2004.

[18] Suh, B., Ling, H., Bederson, B. B., and Jacobs, D.W., "Automatic Thumbnail Cropping and
      Its Effectiveness." Proceedings of the 16th annual ACM symposium on User Interface
      Software and Technology, Pages 95-104, 2003.