Sonorific Toaster: Image Processing
COS436/ELE436: HCI Technology
Professor Perry Cook
January 13, 2009
SchGaDY's Musical Toaster

Malia Douglas, Dmitri Garbuzov, Gordon Scharf, Jennifer Yu

I. Introduction:

Sight and attention are valuable in the kitchen. Using tones to alert the cook is a common convention in appliance interfaces. A timer alert, however, is not a very informative form of feedback to receive from an appliance. Color is a better indicator of whether or not the food in question is ready. For the purposes of this project, we restricted the problem to only one class of baked goods: white bread toast. We then used a webcam to acquire images of the inside of the toaster and used image processing techniques to analyze the change in color. The ultimate objective of the image processing is to detect when the bread has reached its desired level of brown and thus, presumably, its desired toastiness. However, this question can also be seen as determining the point in time when the toast has reached the optimum color change from its original toast color. Thus, instead of dealing with absolute color measures, the image processing could merely be concerned with sensing relative color differences. By sonifying this change in color, we hope to help the user be more productive and consistently achieve optimal brownness of their baked goods.

II. Equipment:

- Logitech QuickCam for Notebooks
- Laptop
- Toaster
- Bread

III. Image Acquisition:

JPG images were taken at 30 seconds intervals using a Logitech QuickCam for Notebooks. Originally we took a set of data manually holding the camera directly against the glass. This produced sharp images, and the color differences were readily apparent. However, holding the camera against the glass of the toaster oven turned out to be too hot for the camera to withstand for long periods of time. Thus, we mounted the camera by clipping it onto a piece of curved plastic that was taped to the handle of the toaster oven. The distance from the glass caused the quality of the images to decrease slightly.

To further enhance the differences between the images, we removed the IR filter from the camera. This involved disassembling the camera and breaking off the small red filter that was in place. We then took new sets of images with the modified camera. Unfortunately, removing the IR filter seemed to wash out the image with white light. Luckily, the datasets still seem to be usable provided that the right heuristic was used to detect changes in the colors of the bread as it toasts.

Another large problem is inherent to toaster ovens. On the “toast” setting of a toaster oven, the toaster oven starts off cold and with its heating coils turned off. As the toaster oven heats up, the heating coils begin to glow brighter and brighter red; as their brightness increases, the light inside the toaster oven increases which dramatically alters the quality of the images.

IV. Color Spaces:

Although to the human eye the color differences between white bread and toast are readily apparent, it is more difficult to sense these differences using a computer. One of the difficulties with detecting color changes as the bread toasts is that the bread will be a shade of brown at all stages of toasting and thus, under some color schemes, very similar in appearance. There are a multitude of color models that exist for image processing. The picture information produced by the camera is transmitted in terms of RGB values. However, in order to track the browning of the bread, we converted our photo information from the RGB to the CIELAB color space.

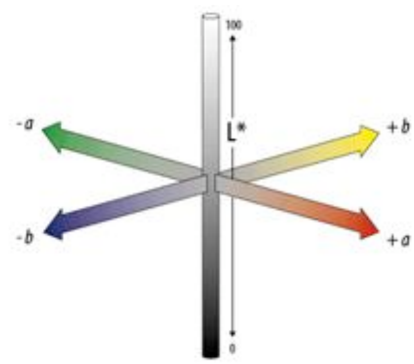
IV.1: RGB

In an RGB color scheme, colors are decomposed into different amounts of red, green, and blue. This is an additive color model; thus white is the additive combination of equal amounts of all three colors while black is the absence of colored light. When an image is read into MATLAB, it is read using the RGB color scheme. In MATLAB, images stored in this format are stored as a set of three arrays (one per color) each of which is 240x320 (the total number of pixels in the image). In an RGB color scheme, all of the shades of brown look remarkably similar. Since what distinguishes the browns from one another is primarily the amount of white present in the image, a different color scheme is more suited to the problem at hand.

IV.2: CIELAB

CIELAB is an opponent color system in which colors are decomposed into values that are represented by the L*, a*, and b* axes. Color opposition models supposedly most closely approximate a human interpretation of color. According to Adobe, "[c]olor opposition correlates with discoveries in the mid-1960s that somewhere between the optical nerve and the brain, retinal color stimuli are translated into distinctions between light and dark, red and green, and blue and yellow." Additionally, unlike other color schemes, CIELAB aspires to be perceptually uniform.

In the CIELAB model, the L* axis is the central vertical axis in the space and represents lightness; its values range from 0 to 100 where 0 represents black and 100 represents white. Both a* and b* are color axes whose values range from negative to positive values. The a* axis represents the opposition between red and green, and the b* axis represents the opposition between yellow and blue. The origin, which is located at the intersection of the three axes, is a neutral gray color. The diagram below illustrates the CIELAB color model.



V. Final Implementation

V.1 Edge Detection: Jen Yu

Originally we planned on using edge detection for two potential purposes. The first was to detect the location of the toast so that we could distinguish between toast areas and toaster areas. This would be useful in the processing of the toast's color information. The second was to track the change in volume of the toast. After we toasted our first piece of bread, we realized that, in addition to color, volume changes significantly as bread toasts. Thus, if we could track the movement of the edge of a piece of toast, we would have another indication of toastiness.

Ultimately, edge detection was not incorporated into our final project. Edge detection of edges in the images was not difficult to do (using Canny edge detection algorithm), but figuring out the appropriate threshold to use and figuring out which edges corresponded to the sides of the toast became a harder problem. For the initial piece of bread, thickening the detected edges slightly allowed the disjoint edges to connect and for continuous edges to be detected. Using the number of pixels in the thickened edge and the pixel span in the y-direction, the vertical edges of the toast could be detected in the first frame; thin, long vertical lines would be detected as the toast edge. For subsequent frames, this method was unreliable as the toast got darker because the edges would not be detected so clearly, leaving too-short edges and increasingly more noise. The next problem became tracking the toast edge. Even though the toast is a stationary object, tracking the edge of the toast was still difficult. If this endeavor were successful, we could use the change in volume of the toast as an additional indicator. However, tracking normalized color changes proved to work pretty well and Operation Edge Detection was shut down.

V.2 Color Detection: Malia Douglas

Before arriving at our final solution, we attempted several different methods using various other color schemes such as HSV and gray scale. Unfortunately, although using a scheme such as gray scale seemed as though it would be simple and convenient, it ultimately turned out that there was no discernable trend that held true for multiple sets of toast images. Thus, after several unsuccessful trials, we ultimately decided to use a CIELAB-based scheme.

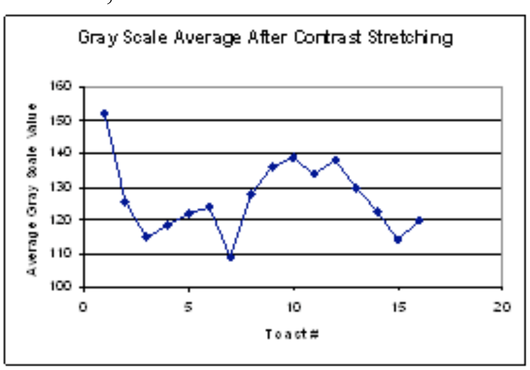
V.2a: First Approach

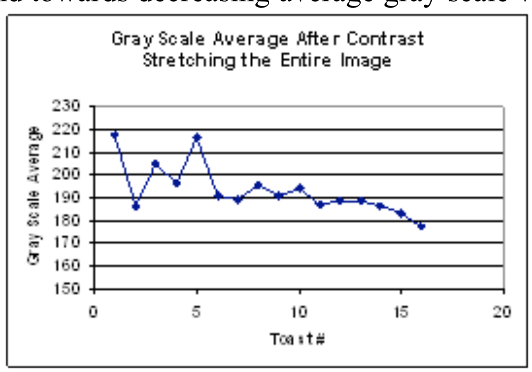
One of the simplest ways to detect the change in color of toast is to watch the change in intensity of the toast. Presumably as the bread gets toasted, the majority of the pixels will go to increasingly dark shades of gray in a gray-scale representation. This means that if you were to plot a smoothed version of (gray-scale value, number of pixels), the peak of the curve should move slowly from its initial value to a value of darker gray as time passes.

Although this scheme would be convenient, it turns out that there is no discernible trend that the peaks of the histograms follow. The peak quickly changes and then stays more or less stationary for the rest of the toasting process. I suspect that the initial sudden change in the peak correlates with the point at which the heating coils began to glow red.

V.2b: Second Approach

In order to compensate for the changing light conditions in the toaster oven, I tried taking a small section that only represented the surface of the toast. I then converted this section to gray scale and found the minimum and maximum values for which the histogram had a nonzero value. I then used these values to contrast stretch the image so that the minimum value was mapped to the 0 gray scale value and the maximum value was mapped to the 255. I then calculated the average value for the gray scale array of the contrast-stretched section of the original toast image. Unfortunately, when I plotted these results, I got the chart seen at the bottom left; there was no discernible pattern to the increase and decrease of the gray scale averages over time. I then realized that contrast stretching before taking a small section of the toast image might make more sense in terms of normalizing the image across the different lighting conditions. When I first contrast stretched and then took the average of the gray scale values of the same small section of toast, I got the results displayed in the chart on the bottom right. Although there is not a large decrease after the initial rises and falls, there does seem to be a slow trend towards decreasing average gray scale values.

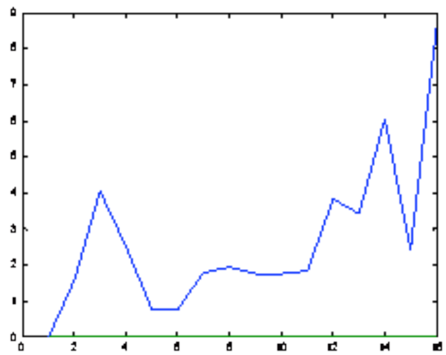




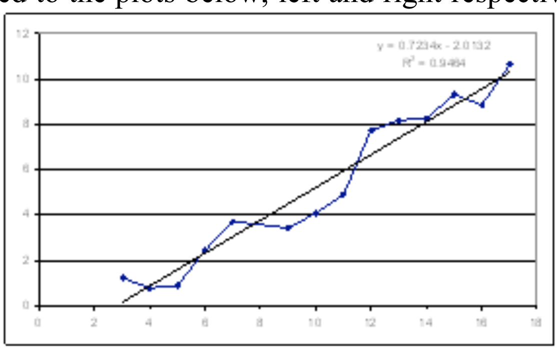
V.2c: Third Approach

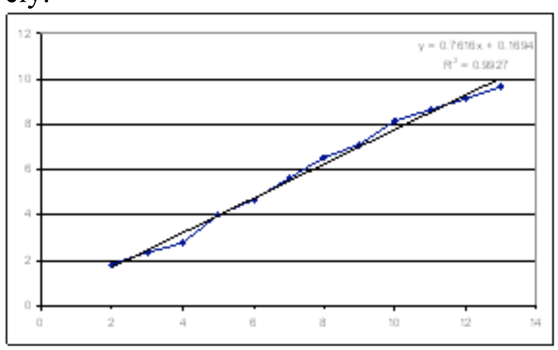
In several papers from the Journal of Food Engineering, a popular way of analyzing the change in color as food items browned was to convert the RGB information to CIELAB data. In a paper by Purlis et. al., the difference between time points was calculated as the Euclidean distance between two points in CIELAB space. In order to convert from RGB color space to CIELAB color space, we used a function called "colorspace" in MATLAB that can convert from RGB to XYZ to CIELAB color spaces. However, one of the significant differences between our project and this paper is that the baking done in the Purlis paper was done in a pre-heated oven; thus, the lighting remained essentially constant throughout. Since the lighting in the toaster oven changed over time, we wanted to normalize the color of the toast against some reference color.

In order to normalize the illumination, we found a square of image that was associated with the background color of the toaster, took the average, and then subtracted this average from the image of the square of toast. As the light changed, the average of the background color presumably changed as well. If this average were not subtracted out, the plot of the distances between consecutive coordinates would look like the plot below for the toast set "UVless 3." The x-axis represents the number of the picture and the y-axis represents the distance in CIELAB space between the current toast and the initial toast.



However, when the average is subtracted out, the plots for toast sets "UVless 1" and "UVless 3" are transformed to the plots below, left and right respectively.

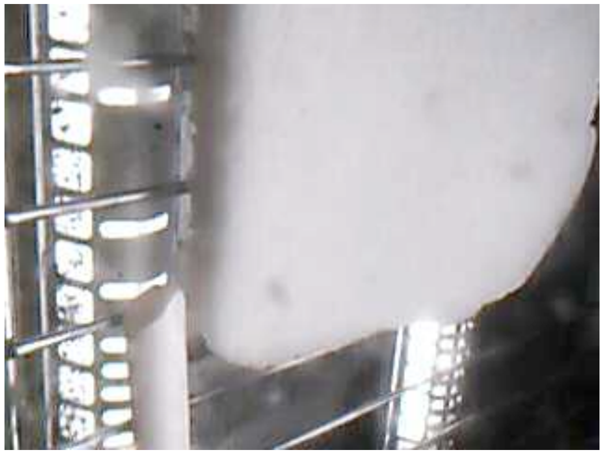




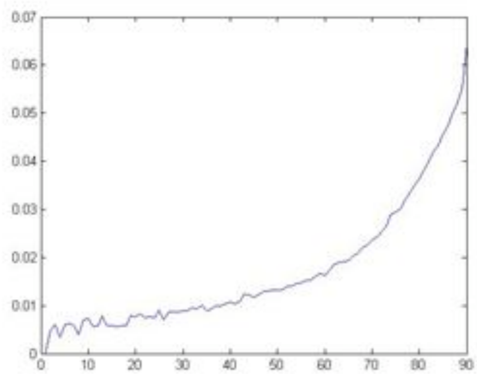
Only the most linear parts of the plot were preserved, and a linear regression was run on each line. As can be seen above, the fit was reasonably strong, and the slopes of the two plots were comparable. Thus it seemed as though this scheme could be used to produce consistent results across trials.

V.3d: Modified Third Approach

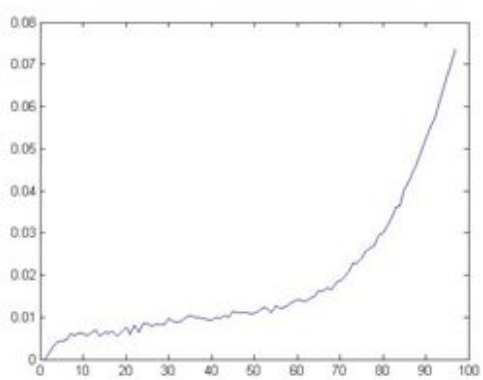
Although it seemed as though the approach described above produced results that followed a general trend, it turned out that using a square from the toaster as a reference was not very reliable. Even small changes in the square selected as a reference made huge differences in the output. Thus, we stuck a piece of pink chalk into the toaster to use as a reference color instead. In the picture below, the chalk is seen in the lower left corner.



We then toasted a new piece of bread and used the same algorithm to quantify the browning of the bread; however, we used the chalk as the reference. This produced the following plot in which the x-axis represents the number of the picture and the y-axis represents the distance in CIELAB space between the current toast and the initial toast:



This plot seemed much more plausible than the linear plots generated previously since toasting bread is not a linear process. In the first 5-6 minutes (using the toaster in the Brown Co-op), the bread toasts slowly and at a relatively uniform rate. However, as the bread gets more brown, it suddenly starts moving towards becoming burnt very quickly. In our new plot, this is reflected in the rapidly changing slope of the graph as time passes. In order to verify that these results were reproducible, we toasted another piece of bread using the chalk as the reference point and got the following plot in which the x-axis represents the number of the picture and the y-axis represents the distance in CIELAB space between the current toast and the initial toast:



This plot agrees with the previous plot confirming that this color scheme seems to be a reliable means of detecting color change.

V.4 Switching from MATLAB to Java: Dmitri Garbuzov

After developing and testing our algorithm in MATLAB, we chose to implement the final version in Java. We used Java primarily because our application was not particularly performance intensive and all of the members of our group had some previous experience with Java. This allowed us to split up our final system into a number of independent classes for sonification, image acquisition, and analysis. The final implementation could analyze roughly an image every second, which was sufficient for our application.

V.5 Sonification: Gordon Scharf

Once the toast's status was reduced to a single dimension, the final remaining task is was to create an effective metaphor in the auditory space. In keeping with my research into mappings from images into auditory displays, I created two mapping schema to test on end users:

- Pitch-based sonification: For each new image, a single interval (two notes) is played. The first note of the interval is always the same, and the second is based on the difference between the current image and the first image taken.
- Tempo-based sonification: For each new image, five tones are played, the pace of which (the length of the tones and spaces between them) is based on the difference between the current image and the first image taken.

:

Based on early feedback from members of the Brown Co-op and research performed by Dmitri into the sonification of single-dimension plots, we decided to create another pitch-based scheme. We store all the historical data for the piece of toast, and then play a series of 15 pitches in the space of 1-2 seconds that give a condensed history of the toast's darkness. This begins with a low murmur, and builds up to something that sounds like a slide whistle. This auditory mapping gave much better qualitative results from test subjects.

VI. Experimental Design:

VI.1 Qualitative:

Test Subjects: For testing, 8 members of the Brown Co-op were asked to experience the Sonorific Toaster and then fill out the following survey. Responses were given on a scale of 1 to 5 where 1 is strongly disagree and 5 is strongly agree.

Pre-Sonorific Toaster Experience Statements:

(1) Setting a fixed time on the toaster oven (without visually checking the toast until time is up) is a sufficient means of making toast that is optimally toasty.
(2) It would be useful not to stand by the toaster oven to know how brown your toast is.
(3) Burning toast is a huge problem in my life.

Post-Sonorific Toaster Experience Statements:

(1) It was easy to correlate the sonification to the browning of the bread.
(2) I feel the quality of my toast was improved by using the Sonorific Toaster.
(3) Getting sonified updates on the progress of the browning of your toast was useful.
(4) If your response was 3 or less, rate the following statements:
(4a) Sonifying the change in other baked goods as they bake (e.g. cookies, muffins, etc.) would be useful.
(4b) Sonfication information would be useful some of the time.
(4c) I would like this product better if the sonification feature could be turned off and on manually.
(5) The individual sounds were spaced at good intervals.
(1 on the scale correlates with "The sounds should have been spaced much closer together" while 5 correlates with "The sounds should have been spaced much farther apart.")
(6) The quality of the sound was aesthetically pleasing. (1 correlates with "I hated the sounds; they were really annoying" while 5 correlates with "I loved the sounds; I would put them on my Ipod.")

Fill-in-the-blank: I would pay $____ for the Sonorific Toaster.

After using the Sonorific Toaster a second time, please re-rate statement (1) of the Post-Sonorific Toaster Experiment Statements and reply (on the same scale of 1 to 5) to the statement below:

(1) My toast was perfectly toasted.

Additional Comments (if any):

VI.2 Quantitative:

Given the nature of this project, it is difficult to quantitatively evaluate the performance of the Sonorific Toaster. We primarily quantitatively evaluated the consistency of our output. This included both consistency within a single run and across multiple runs. Within a single run, we were interested to see how dependent the outputted curve from the color-processing code was on the region of the bread that was used as input to the code. Across runs, we were interested to see how much variation in range and curvature there was in the output plots. Regions of interest in the curves were the point at which the slope began to dramatically increase and the range of output values.

VII. Results and Evaluation:

VII.1 Survey Results:

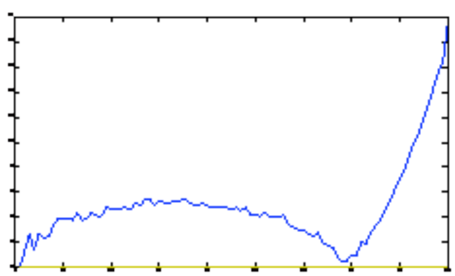
The majority of respondents were not particularly interested in sonifying toast (ones and twos), though many avid bakers were excited about the prospect of using it for other applications (fours mostly, with one five). Toasters are consistent enough that setting a timer for that application is sufficient most of the time. Once given the opportunity of experiencing the sonorific toaster, respondents gave slightly higher evaluations, saying that it was moderately useful to have sonified updates of the toast's brownness. One trial, however, was not sufficient for them to guess when their toast was done based purely on sound. This is most likely due to the fact that any such mapping requires a period of adjustment that we cannot afford to observe, coupled with the difficulty in achieving consistency in our system.

VII.2 Quantitative Results:

VII.2.a Variation within a single run of toast based on input

(1) Changes in the selection of the reference image

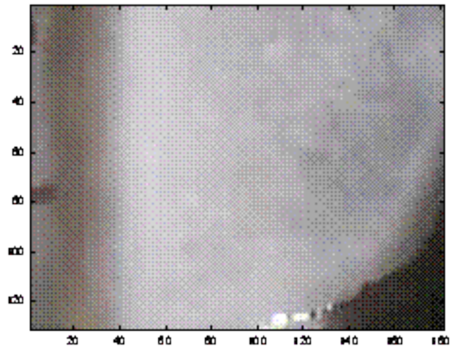
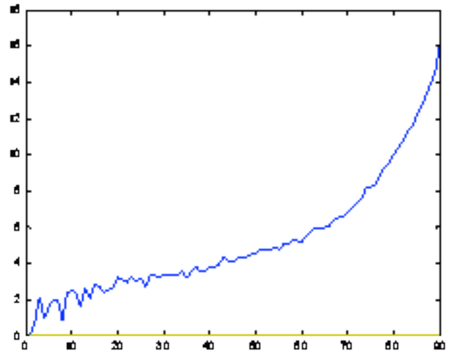
The reference image had to be vey carefully selected. Through trial-and-error we realized that unless chalk, and only chalk, was included as part of the reference image, our results would be dramatically different. When we used a chalk input that was approximately half-chalk and half-background as the reference image, we got variations on the following plot:



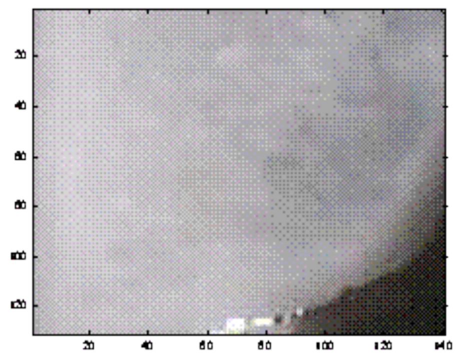
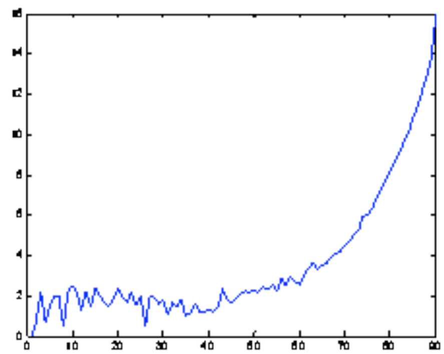
(2) Changes in the selection of the toast image

In all of the plots below, the x-axis represents the number of the picture and the y-axis represents the distance in CIELAB space between the current toast and the initial toast. To illustrate the differences between the toast inputs that were used, the final input image (i.e. the 90th toast picture) is displayed on the left of each plot.

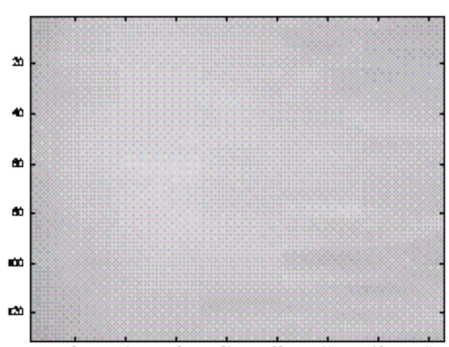
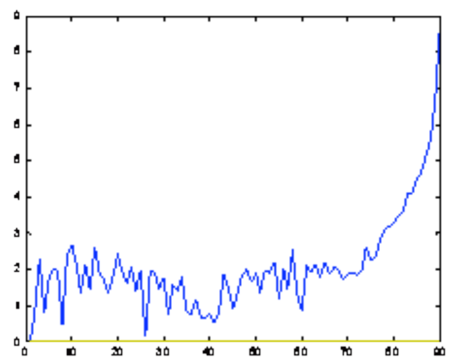
This first plot represents the “best” output. The parameters for the size and location of the input, which define the image seen on the left, were selected to make the output as smooth and as exponential as possible. As can be seen, the input included non-toast elements such as some pieces of the background; additionally the crust of the bread was a prominent feature.

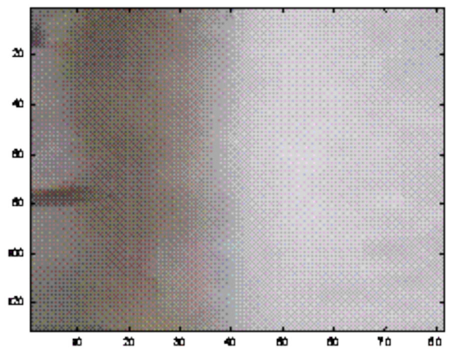
The plot below is the output when the toast input from before is cropped so that the crust is no longer part of the input image. The plots are very similar in range and curvature. The range is approximately 0-16 in both cases and both look roughly exponential. The biggest change is that the portion of the graph that is reasonably flat has become much noisier.

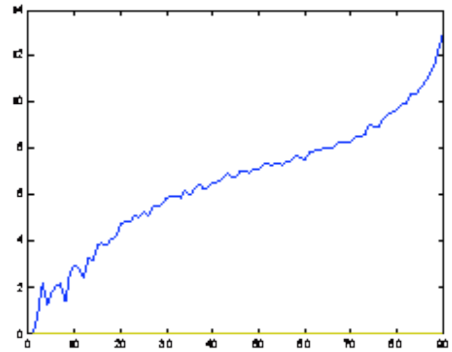



The next plot represents the output when the image above is cropped further so that the lower right corner is eliminated. Now the input is solely the surface of the toast. Now the range of the output is significantly reduced and the initial part of the graph is extremely noisy. However, the plot still follows the same general trend.

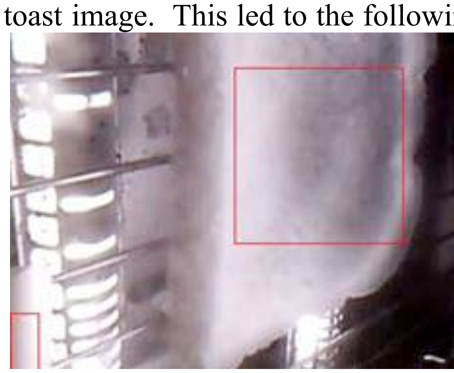
The last plot illustrates the output when the original “best” input was cropped to remove the lower right hand corner, but not the crust. This output seems to follow a different trend from the original three, but is also not that noisy.





These results are largely inconclusive as to what the best manner for selecting the region of toast for input is. The last image may have had a different trend because the ratio of toast to non-toast is relatively low compared to the initial three images. If only the first three images are taken into consideration, it seems that the choice of toast merely affects how noisy the original signal is, but that the general trend is the same regardless of toast selection. Thus, given that we have no algorithm for determining the best input, we need to have sonification scheme that can handle noisy input.

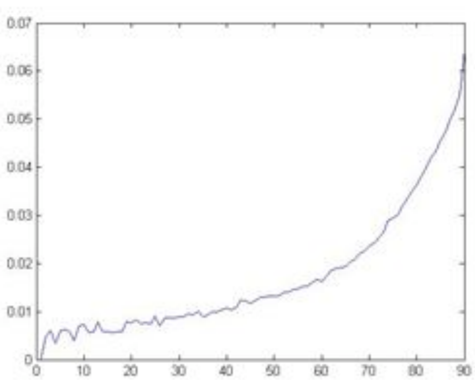
Based on these results, we realized that we needed some sort of targeting system so that we could precisely line up the reference image and the toast image. This led to the following:

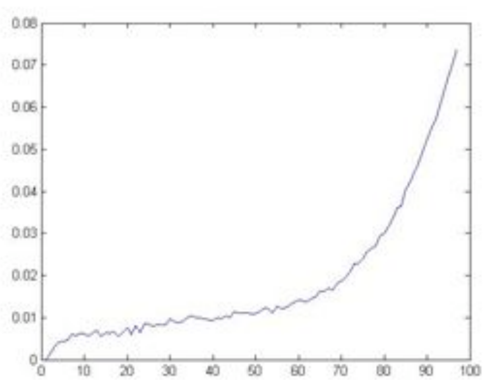


The red boxes are targets that appear on top of the image generated by the webcam. Thus, before beginning to toast the bread, you can properly align both the chalk (reference) and the toast. This greatly improved the quality of our sonification.

VII.2.b Variation across runs of toast

The following two charts are the results from our first two runs of toast and represent the "best" output from the algorithm. The "best" output was primarily determined based on the smoothness of the output curve. It should be noted that these plots were normalized in a manner that the four shown in the previous section were not. However, the color algorithm was the same.





The two plots are largely similar. Although it appears that the plot on the right has a larger range, it also had a larger number of samples; the plot on the right had 98 while the plot on the left only had 90. Thus our color algorithm seems to give consistent results across different trials as long as the reference input is selected carefully.

VIII. Future Work

Our current project was very limited in scope. We only used bread/toast as the food item for testing. It was also required that the toast be placed in approximately the right area. Future work would expand the Sonorific Toaster to be used on a larger variety of food items, such as cookies or hot wings or wheat bread. Our scale for sonification is based on the range of intensities that toast reaches. Other items toast differently and span a different color range. It would be great if we could either provide an interface to make it easy to add new food items to be used by the Sonorific Toaster. Also, moving from a toaster to an oven would be another avenue to pursue, further widening its range of applications.

Since determining the level of toastiness is currently based on color, it is limited in its ability to detect changes. With more time and a broader focus than toast, we'd like to reassess the usefulness of volume change and other characteristics. Adding a learning algorithm and trying out various classifiers, both to detect what type of food item is being toasted and to determine its level of toastiness, would be an extension that would greatly increase the flexibility of the Sonorific Toaster. We may also consider advancing our sonification scheme to use MIDI instead of our current less aesthetically pleasing tone-based system.

IX. References

P. Getreuer. Color Space Converter. MATLAB Central, 28 May 2005.
E. Purlis and V.O. Salvadori. Bread browning kinetics during backing *Journal of Food Engineering*, 80:4,pp. 1107-1115, 2006.
S. Jabri, Z. Duric, H. Wechsler, and A. Rosenfeld, Detection and Location of People in Video Images Using Adaptive Fusion of Color and Edge Information *Proc. Int'l Conf. Pattern Recognition,* pp. 627-630, 2000.
L. M. Brown, S. A. Brewster, "Drawing By Ear: Interpreting Sonified Line Graphs", In *Proceedings of ICAD*
2003, University of Boston, MA, 2003

X. Acknowledgments

We'd like to thank the fine people of the Brown Co-op for letting us mount a webcam on the front of their toaster and for donating countless slices of white bread to our cause.