

Grocery Shopping Assistance

Continuity Report

Grozi

Faculty Advisor: Serge Belongie T.A.:
Brent Lee

Jae Shin
Alrik Firl
David Ing
Chris Lam
Robert Tran
Chung Chen
Yu-Tseng Chou
Brandon Lehrman
Donnie Douangchak
Jose Melendez Castro
Mehran Takapoomanesh Baghaei

UCSD ENG 100L Global Ties
Spring 2011

Introduction

Goal and Quarter Approach

GroZi's goal is to create a system to aid the visually impaired in shopping for groceries. The scope of GroZi is currently limited to the shelved section of the store. This project has been developing for some quarters now, but in this quarter it was decided that we would start with a new approach that would allow the project to improve the communication and assistance for the blind user.

This quarter we focused on user interface development using sonification. Sonification is the process of transforming information or data into sounds. For this quarter, there was a division from within the team in order to present better and faster results. The team was divided in 3 sub-teams that were mainly identified by the kind of information they processed and received. These were the 3 teams:

- Image Team
- Sonification Team
- Interface Team

All of the sub-teams needed to have constant communication and feedback in order to process efficiently the information that is being extracted from the user and its environment. Each one of the teams used different platforms and programs to process the information.

As an additional work, a member of our team worked during the quarter with gestural interfaces to provide the general team a better understanding of better ways for a person to interact with a device, like the one we attempt to create. This work is resumed in the last part of this continuity report to provide an insight of what will be studied in depth during the next quarter.

Image Team

Brandon Lehrman, Mehran Takapoomanesh Baghaei, Alrik Firl, Jose Melendez Castro

Goals:

- Modularizing the code.
- Integrating with the sonification and UI

For the Spring quarter of 2011, the image team built on the progress made in the Winter; specifically, we integrated the distance transform code with label-Me, a website that allows for manual segmentation and labeling, while concurrently making the distance transform algorithm into a module, which allowed us to run the algorithm on different images much easier. That is, prior to this quarter, our distance transform algorithm was hard-coded for each image, meaning that every time the distance transform was run on an image, input parameters specific to that image had to be manually input. Now, the distance transform algorithm has been changed into a function that takes generic inputs of an image, a segmented image, an output file, and RGB color. This new change in our distance transform algorithm makes segmenting a set of images much easier and faster, instead of manually running the distance transform code on each image ourselves. For example, if we wanted to run the distance transform algorithm for a particular object on an image, we used to do the following:

```
A = createMaskFromXML('cereal.xml', 'miniwheats');
```

```
A = ~A;  
%Ain = IMRESIZE(A,.3);  
%Ain = ~Ain;  
res= bwdist(A, 'euclidean');  
res= uint8(res);  
subplot (2,2,2), imshow(res);  
subplot(2,2,2), title('miniwheats');
```

```
A = createMaskFromXML('cereal.xml', 'specialk');
```

```
A = ~A;  
%Ain = IMRESIZE(A,.3);  
%Ain = ~Ain;  
res= bwdist(A, 'euclidean');  
res= uint8(res);  
subplot (2,2,3), imshow(res);  
subplot(2,2,3), title('specialk');
```

The above code illustrates how every time we wanted to produce a segmented image for a particular object in the image, we had to hard code the parameters. Now, we reduced the above code to the following:

```
LabelCreator('shelf_2.jpg','cereal.xml', 'miniwheats','miniWheats.jpg');  
LabelCreator('shelf_2.jpg','cereal.xml', 'specialk','specialk.jpg')
```

Every time we want to run the distance transform algorithm on an image, we call the function LabelCreator, which takes an image, an XML file, the object we are looking for, and an output file. Upon being called, the LabelCreator function will call the distance transform algorithm to create the desired segmented image. Since our goal was to integrate the distance transform algorithm with Label-Me, our function LabelCreator allows us to do just that. The purpose of having this code is to offer a mock-up of the image processing pipeline, in order to facilitate the sonification team's development. Namely, their code relies upon receiving the distance transform of the image, along with the identifying metadata. The actual segmentation and object classification is planned to be done by Serge Belongie's graduate students. As such, our use of Label-Me and our distance transform code are the current stand-ins.

Future work in the image team's domain could include further integrating with Label-Me to utilize its Matlab toolbox to dynamically obtain the XML file from Label-Me, in hopes of automating the process of segmenting images from Label-Me.

Figure A - Last Quarter Distance Transform Algorithm

```
Ain = IMRESIZE(Ain,.5);  
ImSize = size(Ain);  
for i = 1:ImSize(1,1)  
    for j = 1:ImSize(1,2,1)  
        if((R == Ain(i,j,1))&& (G == Ain(i,j,2))&&(B == Ain(i,j,3)))  
            Ain(i,j,1) = 0;  
Ain(i,j,2) = 0;  
Ain(i,j,3) = 0;  
        else  
            Ain(i,j,1) = 255;  
Ain(i,j,2) = 255;  
Ain(i,j,3) = 255;  
        end  
    end  
end  
  
Bin = Ain(:,:,1);  
Object = bwdist(Bin, 'euclidean');  
Object= uint8(Object);
```

```

imwrite(Object,outFile);
output = imread(outFile);
subplot(2,2,3),imshow(output);
subplot(2,2,3),title('Distance Transform')

```

Figure B - Improved Distance Transform Algorithm

```

function DistanceTransforms_Ver1_2(Image,segImage,outFile,R,G,B)
%performs distance transform of 6 images and returns the images of the
%distances for a set of objects classes

%inputs
%Reads in the image from the MSRC dataset. The images need to be in the
%same directory as the m-file. imread is used.
%

A = imread(Image);
Ain = imread(segImage);
ins = figure;
subplot(2,2,1), imshow(A)
subplot(2,2,1), title('Original')
subplot(2,2,2), imshow(Ain)
subplot(2,2,2), title('Segmented')
orient Landscape

Ain = IMRESIZE(Ain,.5);
ImSize = size(Ain);
for i = 1:ImSize(1,1,1)
    for j = 1:ImSize(1,2,1)
        if((R == Ain(i,j,1))&& (G == Ain(i,j,2))&&(B == Ain(i,j,3)))
            Ain(i,j,1) = 0;
            Ain(i,j,2) = 0;
            Ain(i,j,3) = 0;
        else
            Ain(i,j,1) = 255;
            Ain(i,j,2) = 255;
            Ain(i,j,3) = 255;
        end
    end
end

Bin = Ain(:, :, 1);

```

```

Object = bwdist(Bin, 'euclidean');
Object= uint8(Object);
imwrite(Object,outFile);
output = imread(outFile);
subplot(2,2,3),imshow(output);
subplot(2,2,3),title('Distance Transform');
end

```

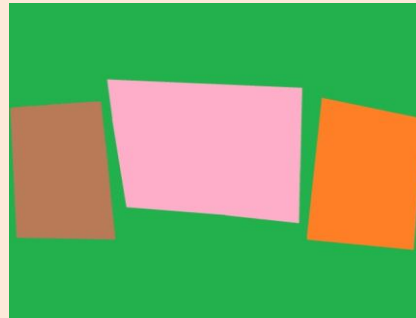


Figure 1 – Real life image of bag of chips Figure 2 – Segmented image of different brand of chips

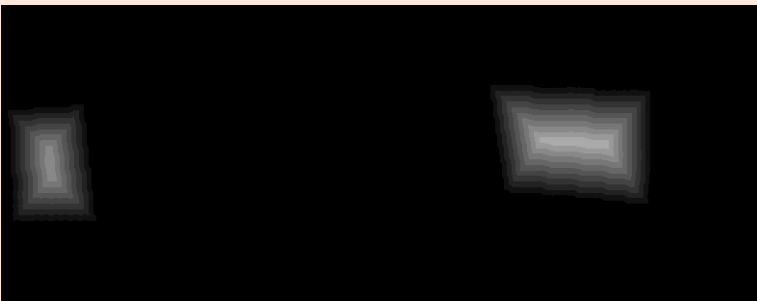


Figure 3 - left chips

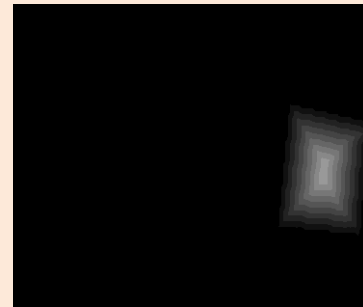


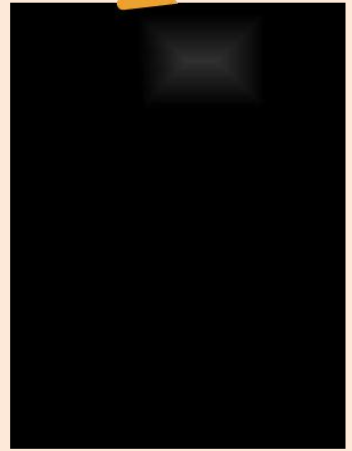
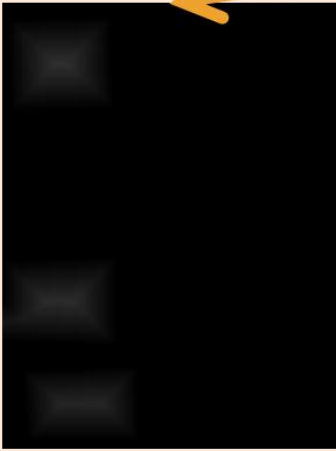
Figure 4 – middle chips

Figure 5 – right chips

Figure 1 displays the actual image of the bag of chips within a grocery store. For the winter quarter of 2011, the groceries that are being focused on are the ones that are shelved notice that the segmentation is done by hand. The following shows our interrogation with labelme.



```
<?xml version="1.0"?>
- <annotation>
  <filename>img_436781078.jpg</filename>
  <folder>static_submitted_3d</folder>
  - <source>
    <sourceImage>submitted</sourceImage>
    <sourceAnnotation>LabelMe Webtool</sourceAnnotation>
    <submittedBy>grozi</submittedBy>
  </source>
  - <object>
    <name>Cheerios</name>
    <deleted>0</deleted>
    <verified>0</verified>
    <date>04-Mar-2011 01:43:18</date>
    <id>0</id>
    - <polygon>
      <username>blehrman</username>
      - <pt>
        <x>5</x>
        <y>17</y>
      </pt>
    </polygon>
  </object>
</annotation>
```



Sonification Team

David Ing , Jae Shin

Objective:

- Receive information & data from Image and User Interface Team.
- Create representing sound or sound effects for data

This quarter, our primary goals were to expand sound database, develop code and finally integrate with Image team and User Interface team. Overall, we are satisfied with what we achieved this quarter. We are now one step closer to full integration with other sub-teams.

Sound Database

In the previous quarter, we built a massive sound database for MSRC and grocery related txt-to-speech mp3 audio files. This quarter, we extended our sound database with set of high-quality earcon and auditory icon mp3, wav files. Auditory icons are physical sounds from the everyday environment that helps users understand what kind of information they are dealing with. We believed that one way to manage the representation of grocery related items is via representing sounds with auditory icons (i.e. sound of biting apple and sound of pouring cereal) and earcon (alert tone). Earcon is a brief, distinctive sound used to represent a specific event.

Working with Code

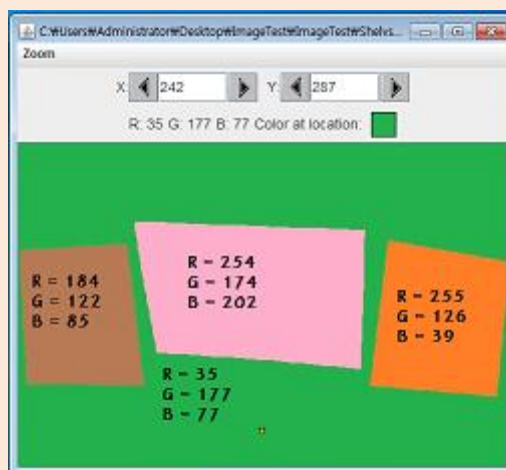
For the previous two quarters, Sonification team used ChuckK which is a powerful audio synthesis programming but in the beginning of this quarter, we decided not to use ChuckK due to its CPU intensiveness. Instead, we decided to use Java to create and develop the program. As we mentioned in previous report, there are two different processes of making sounds based on data and information we receive from other sub-teams. (For further information of description, refer to Continuity report of Fall 2010)

- **XY Object**
- **Distance Transform**

Following steps are to explain how two methods are working as of today.

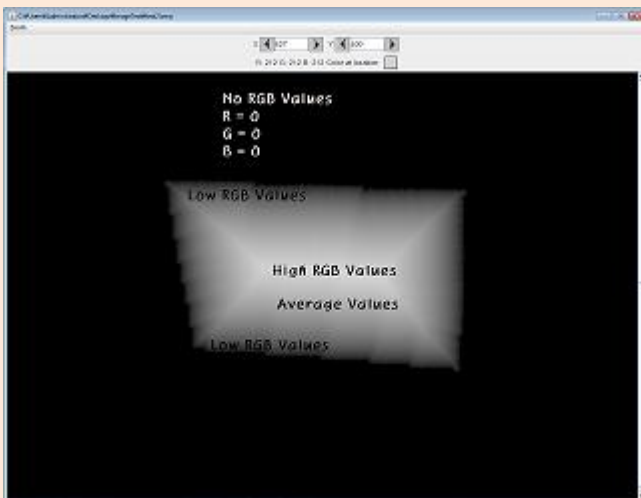
XY Object:

1. Compile and Run the program
2. Load the Segmented Image with RGB values that we received from Image Team.
 - A. Each item/object is assigned to one unique RGB values.
3. Select any point from Segmented Image
4. Selected point will have (x,y) coordinates
5. Each pair of (x,y) coordinates is assigned with specific RGB values.
6. Item representing audio file will be played when RGB values of selected point verifies the item/object in the Segmented Image item. For XY Object, Auditory Icon will be used.



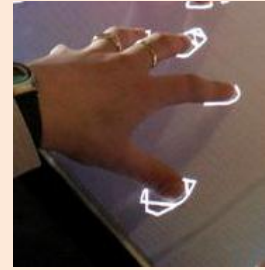
Distance Transform:

1. Compile and Run the program.
2. Load Distance Transform Image with RGB values that we received from Image Team.
 - A. Unlike XY Object segmented image, only the desired item will be assigned with RGB values.
 - B. The desired item will have various RGB values depending on the region.
 - C. The edge of the item will have RGB value of 1. As the point gets closer to the center of the item, RGB values will go up. The center of the item will have RGB values of 215.
3. Select any point from Segmented Image
4. Selected point will have (x,y) coordinates
5. Each pair of (x,y) coordinates is assigned with specific RGB values.
6. Item representing audio file will be played when RGB values of selected point is greater than 1. As the RGB values increase, the sound will increase in loudness/volume. For Distance Transform, Earcon will be used.



Future Plans:

Now that we are almost done with our part of coding, one of our future goals would be to integrate with Image Team and User Interface Team. We were able to successfully integrate with Image Team but we did not have any time to integrate with User Interface Team. I strongly believe that we will be able to integrate with User Interface Team soon.



Interface Team

Last quarter, we had a multi-touch screen that consists of a webcam, a box and a frame and a program in C# that outputs the name of the object each finger is touching on an image (Figure 1). Building off from last quarter, we implemented the code with a new gesture using the idea of a convex hull this quarter.

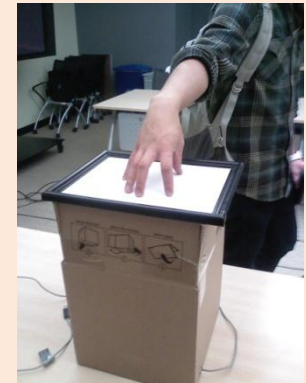


Figure 1. The multi-touch device

When 3-5 fingers are detected, we assume that the user is attempting to locate or search for some object. We then implemented a “searching mode” which follows the three main steps below: approximates a polygon by the positions of the finger touches, checks the color of each pixel within the polygon and lastly outputs the object with the highest percent area. The idea of a convex hull, the boundary of a point set, helps drawing a polygon efficiently since it includes every single point while using the least amount of “string” possible. Two figures on the right illustrate an example of a convex hull (left) and a concave hull (right) and how drawing a convex hull is more efficient (Figure 2). The next possible approach follows a scan line algorithm to mark/fill the region within the polygon and to reduce comparison time.

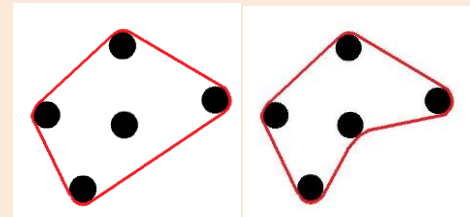


Figure 2. Example of convex and concave hull

As our project progresses, we decided to build a new multi-touch prototype closer to the final product. The new prototype should ideally produce more accurate finger blobs, be more durable while still remain affordable. We will utilize FTIR technology (Frustrated Total Internal Reflection) to track finger touches (Figure 3). This approach involves feeding infrared LED lights through a sheet of acrylic.

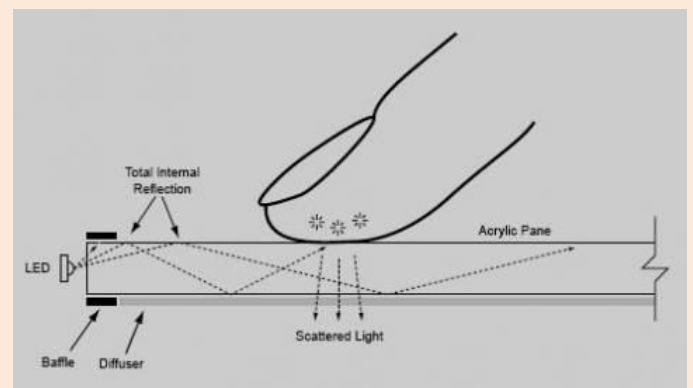


Figure 3. A diagram that illustrates the FTIR technology

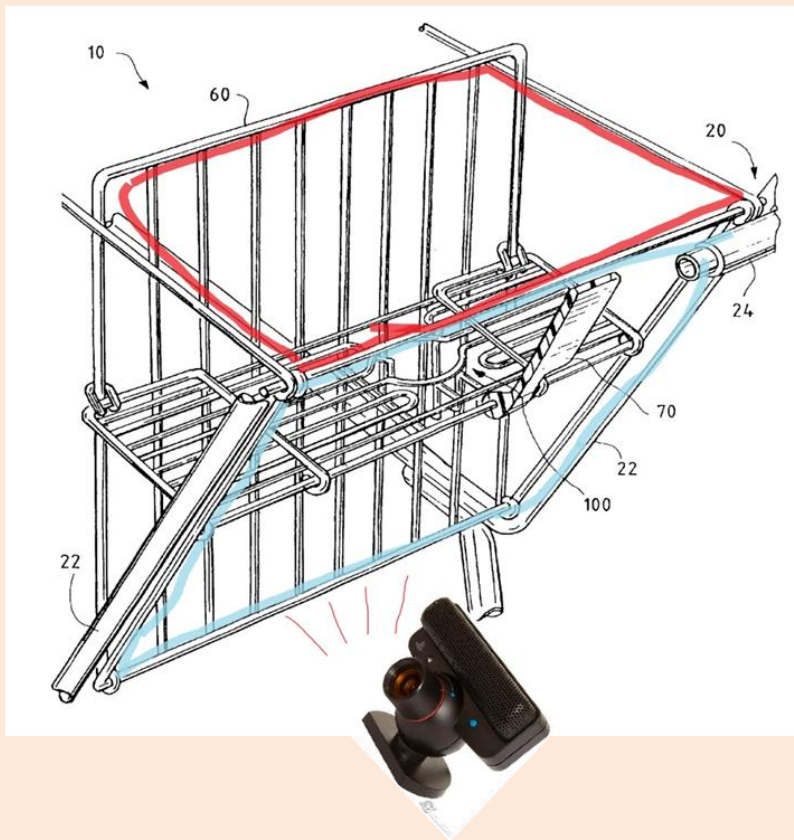
When the user touches the surface of the acrylic, the IR light will scatter, thus indicating the instance of contact. Furthermore, we will utilize a PS3 eye camera equipped with a 850nm bandpass filter to specifically track IR light and ignore all else. This will essentially yield finger blobs with higher contrast (Figure 4).

The code for the project can be accessed at <http://code.google.com/p/grozi/>.

This quarter we have worked on the new multi touch device that has the following benefits:

- PS3 camera equipped with band-pass filter
- LED Infrared lights
- 60% more accurate than our previous prototype.
- Cost-effective

INTEGRATION WITH SHOPPING CART



- Borrow or purchase shopping cart.
- Mount tablet on child's seat.
- Propped for use at red.
- Collapse at blue

Our future plan:

Future plans:

- Sub-team integration.
- Moving toward the evolving Android market and create the GROZI app.
- Exploring other options.

Acknowledgements

- Dr. Serge Belongie
 - Our Faculty advisor

- Dr. Mandy Bratton
 - Program Director

- Barb Donovan
 - Program Specialist

- Brent Lee
 - Teacher Assistant

- Zach Salin
 - Global TIES Fellow