

AN ABSTRACT OF THE THESIS OF

Jennifer A. Inouye for the degree of Master of Science of Computer Science presented on May 25, 2012.

Title: Analysis of Bio-based Composites for Image Segmentation with the Aid of Games.

Abstract approved: _____
Sinisa Todorovic

A fundamental problem in computer vision is to partition an image into meaningful segments. While image segmentation is required by many applications, the thesis focuses on segmentation of computed tomography (CT) images for analysis and quality control of composite materials. The key research contribution of this thesis is a novel image segmentation framework for including end-users in computation. This represents a departure from the traditional methods, which segment images without considering domain knowledge, and access to user feedback. Given a set of CT images of three different composite materials, we would like to create a database of annotated images for all the regions of interest. The annotated images can be used to check the accuracy of segmentation algorithms. Because of how time consuming and mundane image annotation is for a person to do, we propose to turn this task into a game. The game is aimed at making the annotation task easier, because it engages imagination, creativity, fellowship of all subjects involved. In particular, we are interested in games that can be played on the internet by many people like those in Amazon Turk, so that the broader public can get involved. We create a Game with a Purpose (GWAP) called ESP 2.0 for creating image annotations, and thus enable benchmarking of existing segmentation algorithms on our database.

©Copyright by Jennifer A. Inouye
May 25, 2012
All Rights Reserved

Analysis of Bio-based Composites for Image Segmentation with the Aid of Games

by
Jennifer A. Inouye

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

May 25, 2012
Commencement June 2012

Master of Science thesis of Jennifer A. Inouye presented on May 25, 2012

APPROVED:

Major Professor, representing Electrical and Computer Engineering

Director of the School of Electrical Engineering and Computer Science

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Jennifer A. Inouye, Author

ACKNOWLEDGEMENTS

I would like to thank my advisor, Sinisa Todorovic, for giving me the opportunity to work on this project and for his help and support. In addition, Lech Muszyński, for letting me work alongside him during the summer. Last of all, I would like to thank the two high school students that were part of the Apprenticeships in Science and Engineering (ASE) program, Casey Schafer and Anita Chow, for helping to create our database of annotated images and coming up with a basis for the design of this game.

TABLE OF CONTENTS

	<u>Page</u>
1 Introduction.....	1
1.1 Objective	5
2 Literature Review.....	6
3 Our Approach.....	12
3.1 Supervised Segmentation.....	12
3.2 Unsupervised Segmentation.....	14
3.3 Error Calculation.....	15
4 Method.....	18
4.1 What is a Game?	18
4.2 Study of End-Users	19
5 Results.....	26
5.1 Game Outline	26
5.1 Basic Game Play	27
5.2 Break Down Task.....	28
5.3 Game Points	29
5.4 The Single Player Game.....	33
5.5 The Two Player Game	34
5.6 Cheating	34
5.7 Implementation	36
6 Discussion/Evaluation.....	40

TABLE OF CONTENTS (Continued)

	<u>Page</u>
6.1 Collecting Image Data.....	40
6.2 Generalizing Our Approach	41
6.3 Ethical Considerations	41
6.4 Outcomes of this thesis	42
7 Conclusion.....	43
Bibliography.....	45

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
Figure 1 Block diagram of problem statement.....	1
Figure 2 CT scans of the materials.....	4
Figure 3 Region boundary detection.....	6
Figure 4 LabelMe.....	8
Figure 5 VATIC.....	9
Figure 6 Object detection.....	10
Figure 7 ESP and Peekaboom.....	11
Figure 8 Our approach block diagram.....	12
Figure 9 PB images.....	13
Figure 10 NCuts.....	14
Figure 11 Berkeley Segmentor.....	15
Figure 12 Error calculation.....	16
Figure 13 Surveys.....	20
Figure 14 Graph of the centroids of annotations.....	21
Figure 15 Graph of the areas of annotations.....	22
Figure 16 PB surveys.....	23
Figure 17 OSB surveys.....	23
Figure 18 PB image.....	28
Figure 19 Break down of the image.....	29
Figure 20 Learning stage block diagram.....	30
Figure 21 Cheating example.....	31
Figure 22 Levels of the game.....	32
Figure 23 Consecutive PB images.....	38

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure 24 3D structure of the PB.....	39

Chapter 1 –Introduction

This is an interdisciplinary research in which the goal was to analyze images that would be used for image segmentation, and would be used by the Wood Science department. In addition, we had high school students, involved in the Apprenticeships in Science and Engineering (ASE) program, annotate these images. Figure 1 is a diagram of the problem statement that will be described in more detail below.

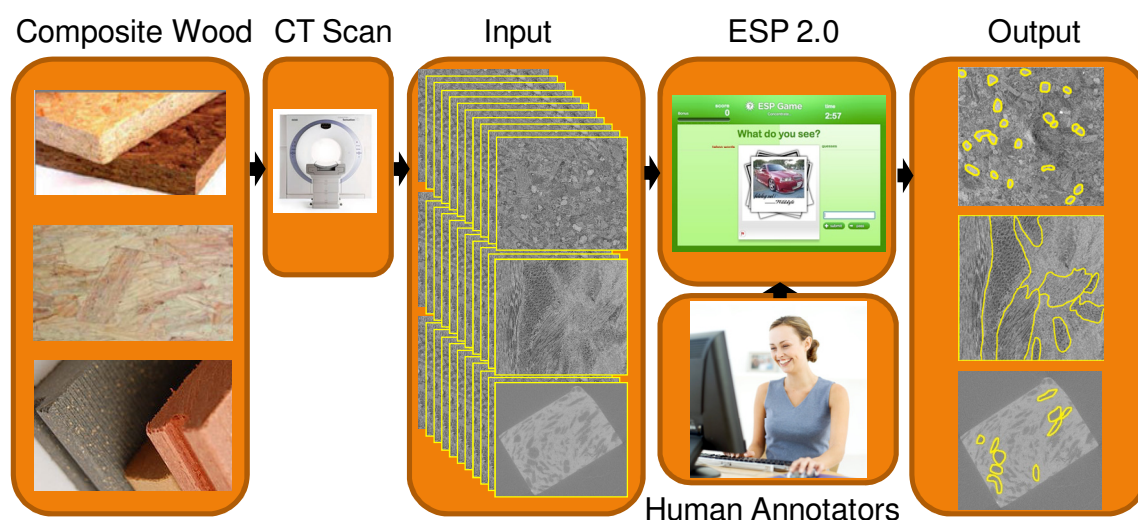


Figure 1. Block diagram of the problem statement. The Wood Science department would like to study the internal structure of three different wood composite materials via CT scanning. Thousands of these images are then annotated manually, which we would like to use as an input to a game to get other people involved with these annotations. Ultimately, we would compare the output of the game, the annotated images, with the output of an automatic image segmentation program to test the accuracy of its detection of the regions of interest.

Particulate bio-based composites (e.g., wood plastic composites WPC) consist of

tightly packed irregular particles of various sizes, bonded with adhesives, or suspended in a polymer matrix (shown in the first block of the diagram in Figure 1). They are widely recognized as a key to healthier and more sustainable economy, since they can be manufactured from naturally renewable resources. Specifically, they have great potential for utilizing low-grade woody biomass generated by forest-fire prevention operations, prescribed forest thinning, or urban waste. Bio-based composites are already common in many applications of limited structural responsibility. However, the progress in the industry is currently being made by way of trial and error, because the existing theories and models do not (or poorly) account for the complex, anisotropic, and heterogeneous nature of particulate. The poor understanding of how the microscopic internal structure affects the performance and durability of the bulk composites hinders their application in many products.

Digital imaging and image analysis – in particular, various computed tomography (CT) techniques (shown in the second block of the diagram in Figure 1), based on magnetic resonance, ultrasound, or X-ray radiation – are becoming important tools in morphological characterization of complex materials. They facilitate the discovery of many new aspects of the material's surface and internal structure, and assists in their statistical quantification. CT images conveniently record the spatial layout and structure of particles, and properties of adhesives comprising the composite material. However, CT images (shown in the third block of the diagram in Figure 1) are typically analyzed by visual inspection, where particles of the imaged material are measured manually. This becomes increasingly tedious, time-consuming, and expensive.

Two high school students from the ASE program were asked to annotate these images and think of a way to turn this task into something fun so that others would be more willing to get involved. Our approach to this problem was to come up with a design for a game for this task like the ESP Game, which gets other people involved to accomplish a meaningful task, and are willing to do so because it is fun (shown in the fourth section of the block diagram in Figure 1). With the help of others, we can then

obtain a larger database of annotated images (shown in the last block of the diagram in Figure 1).

It has been shown that tools for automated image analysis can aid in image segmentation, and, in some cases, replace the expert eye, and thus make the way to discoveries faster, more accurate, and more cost effective. A key computational step of these tools is image segmentation. Its goal is to delineate boundaries of objects present in the image, and thus eliminate the need for their manual annotation. The difficulty of this task depends largely on the number, placement, and brightness of object occurrences, as well as on the complexity of their boundaries[5]. Ultimately, we would like to compare the output of the algorithm with the output of the game (shown in the last block of the diagram in Figure 1) in order to determine how accurate the algorithm is at detecting and segmenting these regions correctly. No image is like the other and each material consist of thousands upon thousands of images, therefore, we need to run several tests in order to determine with high confidence whether the algorithm is working properly or not. And for this, we need a very large database of annotated images to compare against.

The three different types of composite materials to be annotated would be the – Particle board (PB), Oriented strand board (OSB), and Wood plastic composites (WPC). PB consists of pressurized wood particles with adhesive, which we identify as the light colored ‘blobs’ within the image. OSB consists of different strands of wood layered in different directions. WPC is made of wood fiber and plastics. In the CT scans of the PB material, the adhesives, or ‘blobs’ as we will start referring them as, are to be annotated. In the OSB images, the strands aligned in the same direction and clustered in the same region are to be annotated. Last of all, in the WPC images, the wood particles are to be annotated, which are identified by the dark grey areas within the light grey rectangular region shown in Figure 2 below.

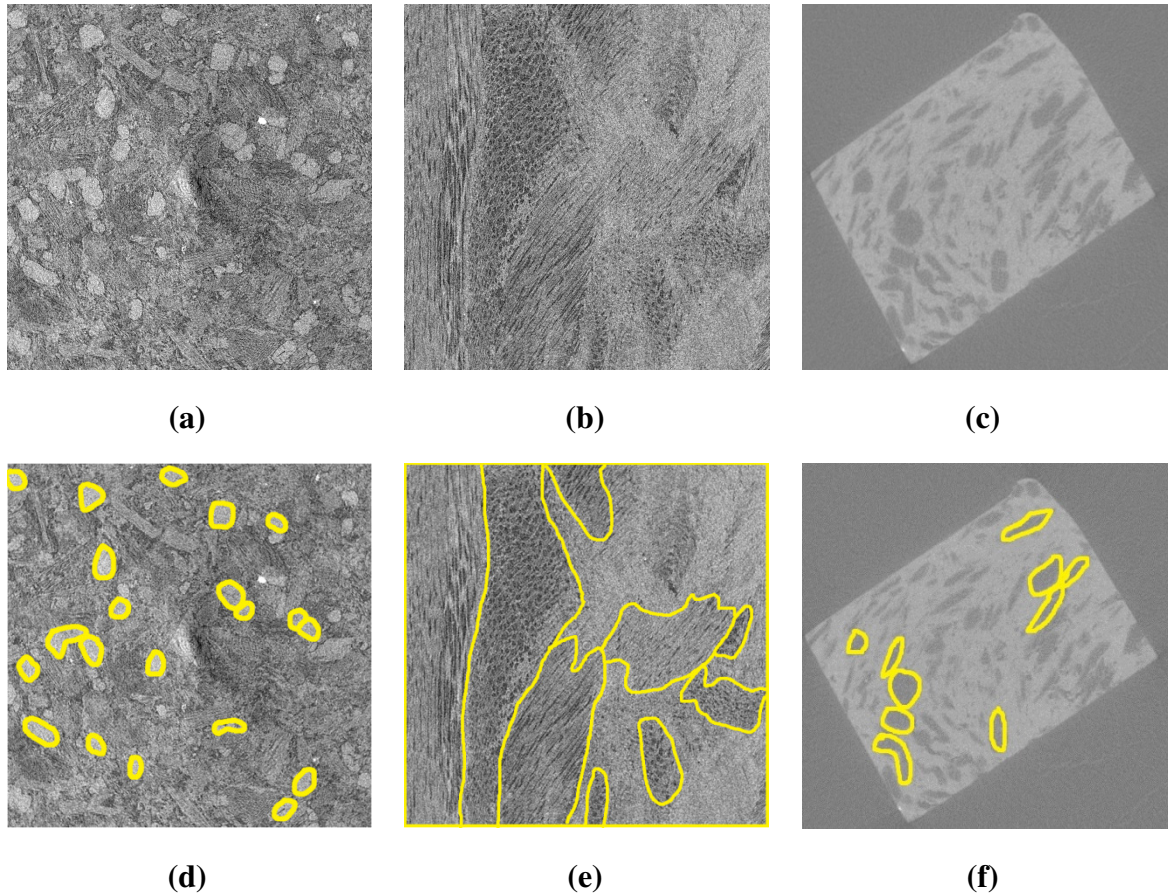


Figure 2. (a) This is a single layer of the PB composite material (1 x 1 mm). (d) This is the same image, shown in (a), where the adhesive composites are outlined in yellow. (b) This is a single layer of the OSB composite material (2 x 2 mm). (e) This is the same image, shown in (b), where each region is delineated by a yellow line. Regions to be annotated in this composite material are defined as strands going in the same direction located in the same area. (c) This is a single layer of the WPC composite material (2 x 1 mm in the light grey rectangular area within the image). (f) This is the same image, as in (c), where the wood particles are delineated in yellow. The wood particles are identified by the darker grey areas within the light grey rectangular area within the image. Note that it may appear that not all ‘blobs’ or wood particles are annotated because there was a set threshold that only those that were approximately 10 x 10 pixels in size were to be annotated, but ideally we would like all of them to be annotated eventually.

The challenge of detecting these region of interest, for example the adhesives in the PB, is that known features for these ‘objects’ do not exist currently. These ‘objects’ have no meaning like cars for example. The ‘blobs’ are not distinct and random in its shape and location within the images. In addition, these ‘blobs’ do not differ much from its surroundings. Like that of a car in an image, if the algorithm were to detect a car in the sky, we would know based on the surroundings that a car cannot exist in the sky and therefore must be wrong. But in the PB images, a ‘blob’ can exist anywhere within the image. There is not enough information on the study of the internal structure of these images to say otherwise. Therefore, we are relying on human annotations to help determine the accuracy of the algorithm. But as mentioned previously, getting others involved in this task of annotating these images is a challenge because it is dull, time consuming, and prone to human error based on different perceptions of what these ‘objects’ are.

1.1 Objective

The overall objective of this proposed study is to create a database of annotated images and to develop a system to obtain a larger collection in which would be used to test the quality of automatic image segmentation algorithms for these images. We need to find a way to get the public involved in such a way that they find it fun and will voluntarily do this without supervision and be able to produce high quality results. This analysis will include: (1) Study of image segmentation algorithms; (2) What is a game and how can we turn this task into one; (3) The behavior of end-users; and (4) How to calculate the percent accuracy of our algorithm using the results we obtain from annotations done by end-users.

Chapter 2 - Literature Review

This section covers a brief summary of various studies in image segmentation, and games and techniques being used in computer science that include end-users.

A number of approaches use segmented image regions as the basic vocabulary of object categories [11]–[18]. These methods argue that regions are in general richer descriptors, more discriminative, and more noise-tolerant than local features. Since, typically, there are fewer regions than local features, using regions leads to great computational savings, which is oftentimes conducive to an improved performance (e.g., because the number of potential outliers are significantly reduced). Region boundaries coincide with the boundaries of objects and their subparts, allowing for simultaneous object detection and segmentation. Regions make various constraints, frequently used in object recognition, such as those dealing with continuation, smoothness, and containment and adjacency. In addition, they are implicit and easier to incorporate than local features (refer to Figure 3 below).

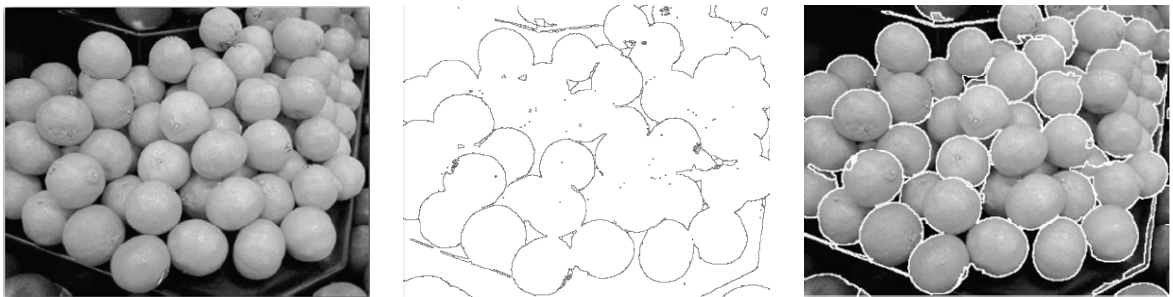


Figure 3. Example of an input image (on the far left) and its output (in the middle). This algorithm takes the input image and finds its region boundaries. This is an example of how region boundaries are similar to object boundaries (better viewed in the far right image, which is the output image overlapping the input image).

For extracting regions as features, the following segmentation algorithms are frequently used: energy minimization [19], active contours [20], scale-space [21], normalized-cuts [22], and mean-shift [23]. The shortcomings of these algorithms induce the deficiencies of higher-level object recognition algorithms that are presented in the sequel. (1) The descriptive power of regions is typically reduced by assumptions about general shape, size, gray-level contrast, and topological context of regions. For example, the energy minimization [19] and active contours [20] approaches define prior constraints on curvature and shape of regions, and the number of regions present in the image. Also, the vocabulary of regions is impoverished by those segmentation algorithms that implement scale as parameter, which controls the degree of image decimation, using different blurring and subsampling methods (e.g., scale-space [21], and pyramid-based methods [24]). (2) The vocabulary of regions describes object categories only at a specific scale, in cases when image partitioning is not performed at a range of homogeneity values, saliencies, or scales (e.g., when using normalized-cuts [22]). (3) Most segmentation algorithms use an oversimplified model of image regions, as being homogeneous and are surrounded by step discontinuities, instead of the more realistic ramp (non-step) discontinuities. Therefore, many regions in real images with small intensity gradients are not segmented, thus reducing the accuracy of object recognition.

In addition, there are other studies being done in different techniques for annotating images. LabelMe is a database and an online annotation tool that allows the sharing of images and annotations. The online tool provides functionalities such as drawing polygons, querying images, and browsing the database[1]. The following steps describe how to label an object using LabelMe[1]:

1. Start by pressing the left mouse button at some point along the boundary of the object.
2. Continue clicking along the boundary of the object to create a polygon.
3. Once you have finished clicking along the boundary of the object, either click on the first point or press the right mouse button to complete the polygon.

4. Click the "Done" button

Figure 4 below is an example of the result created using LabelMe.

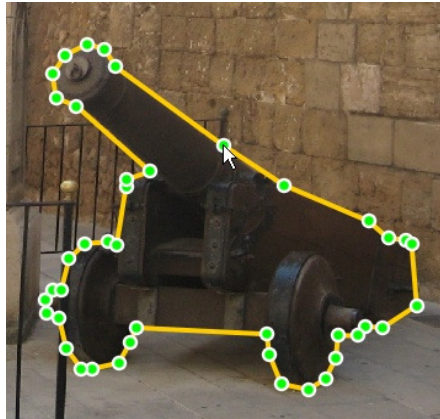


Figure 4. Example output of LabelMe using points to create a polygon around object[1].

This technique is also applied to videos as videos are basically a sequence of images put together. The advantages, of annotating objects via video, is that it lessens the amount of work required by annotators because not all images need to be annotated, but just a selected set of them at different time intervals. VATIC is another tool for annotating objects in videos.

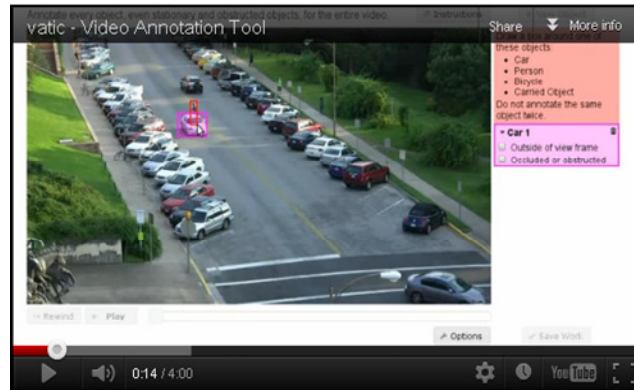


Figure 5. Screen shot of the VATIC tool[2].

VATIC (Video Annotation Tool from Irvine, California) is an interactive video annotation tool for computer vision research that crowdsources work to Amazon's Mechanical Turk[2]. To use this tool, a bounding box is created around the object as it moves throughout the video. The shortcoming of this technique is that for our images, the size and shape are constantly changing and we would like to be as accurate as possible in the annotation of the objects as it will affect the information required by the wood science department such as the area and volume of the objects within the composite materials that need to be identified.

The idea of how object detection works is shown briefly in Figure 6 below.

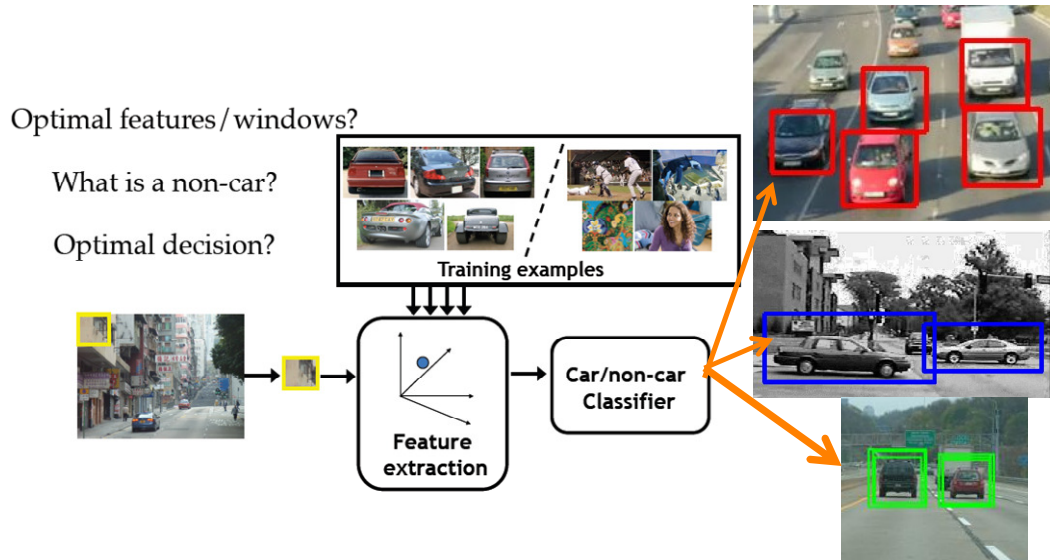


Figure 6. Object detection.

Given an input image, a scanning window, consisting of $n \times n$ pixels, is used to scan across the image. From each scanning window, features are extracted. These features may be based on interest points (ex. corners, patches, or junctions), edges, contours, shapes, boundaries, curves, regions, and textures. The features are then compared against a known database of images and their features, referred to as training examples. Based on the comparison between features in the scanning window and features in the training examples, it is then determined if a car exists or not. If a car does exist, then it will output a bounding box around the car (indicating that a car was detected). An example of this output is shown by the orange arrows. There are, as mentioned previously, shortcomings to these approaches, which can be observed in the first image of the results that show some of the cars not being detected (without bounding boxes).

In addition, there are various studies being done to turn a task into a game otherwise known as games with a purpose (GWAP's). These include the ESP Game and Peekaboom. The ESP Game is a two-player game that collects word labels for arbitrary images[26]. Peekaboom is an extension to the ESP Game in which two players take turns

being “Peek” and “Boom”. The goal of the game is for Boom to reveal parts of the image to Peek, so that Peek can guess the associated word. Therefore, Peekaboom is capable of finding the location of the object in the image rather than just associating the label with the whole image as in the ESP Game[7]. Figure 7 shows an example of what each game looks like.



Figure 7. (a) ESP Game[26] (b) Peekaboom showing the screen viewed by “Peek”. (c) Peekaboom showing the screen viewed by “Boom”.

The shortcoming of these approaches, in relation to this study, is that we do not require labels for our ‘objects’. What we need are the annotations of these ‘objects’ and their location within the images. The location of each of these ‘objects’ in the images are pertinent to this study for the testing of the computer vision algorithm, which we can use to track the area and volume of these ‘objects’ that can help to determine the internal structure of the composite materials.

Chapter 3 – Our Approach

Based on previous studies and research in this area, we need to come up with another approach that incorporates detailed annotations, collected by end-users that we will use to compare to the output of our algorithm. This comparison, will allow us to determine how accurately the algorithm is detecting the regions of interest. The general idea of this approach is displayed in Figure 8 below.

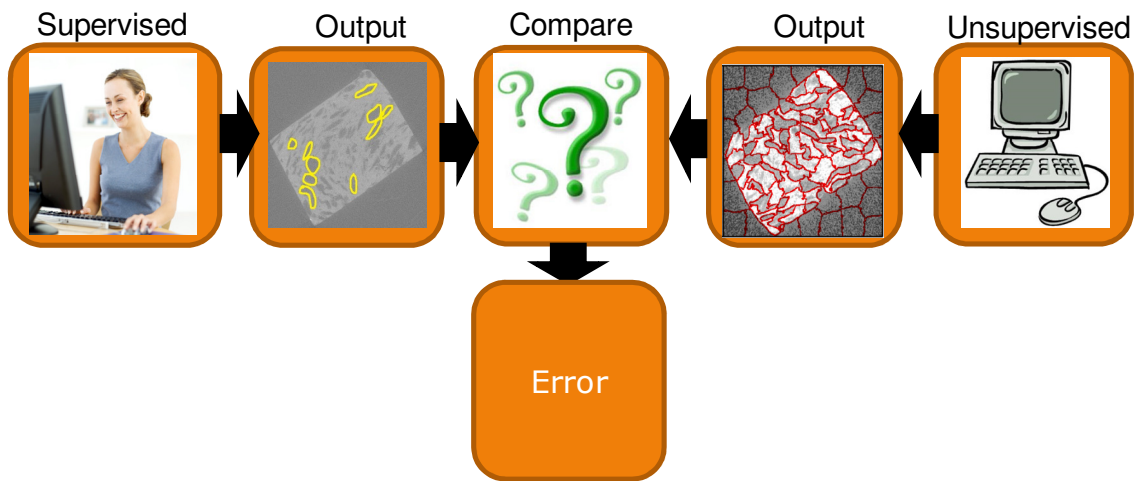


Figure 8. Block diagram which displays how we will test our algorithm for accuracy.

3.1 Supervised Segmentation

Our approach consists of end-users that we want to annotate the CT scanned images of the three different composite materials. The challenging part of this is that the objects we are looking for are not obvious like that of a car. The ‘blobs’ in the PB images do not consist of any distinct features that allow us to easily detect these. Therefore, this

makes these types of annotations difficult. This can be observed in the example images in Figure 9 below that displays two different images from the same PB composite material that was scanned, but from two different layers within the material.

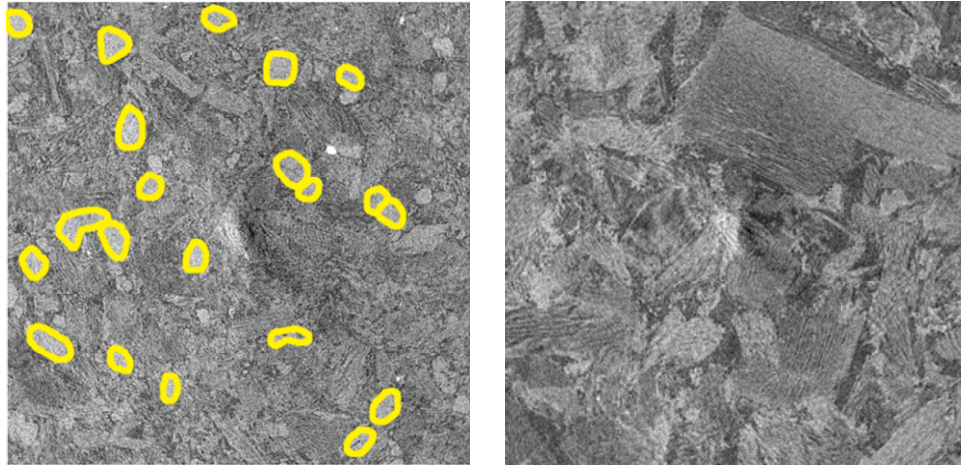


Figure 9. Two images within the same PB composite material.

Based on the two images in Figure 9, it can be seen how difficult it can be to detect these ‘blobs’ in the images as the level of difficulty varies from image to image. The image on the left appears easier to detect these ‘blobs’, but it is not quite clear if a ‘blob’ exists in the image on the left or not. In addition to the difficulty of detecting these ‘objects’, annotating them can become very dull quickly. These images are not stimulating or as pleasing to the eye as seeing a picture of cars. In images with cars there is a lot going on in the image that we can identify with compared to these black and white images containing particles that are unknown and meaningless to us. Therefore, it is necessary that we motivate our end-users to annotate these images and so to address this problem, our approach is to design a game like the ESP Game.

3.2 Unsupervised Segmentation

Now we will discuss the right side of the diagram in Figure 8, the unsupervised approach to this problem. An algorithm to segment these images automatically has been proposed and can be found in ‘Advancing Bio-Based Composites by Automated Image Analysis’ by Sinisa Todorovic and Lech Muszyński[5]. In addition, two other image segmentation algorithms have been looked into. One of which is the Normalized Cuts algorithm (NCuts), and the other is the Berkeley Segmentor. The NCuts algorithm takes as input an image and the number of regions the image is to be segmented into. But because we cannot always predict correctly how many regions there may be in the image, this is the downfall of the algorithm, which in the case that too many regions are set as the input, the image becomes over segmented. An example of this is shown in Figure 10.

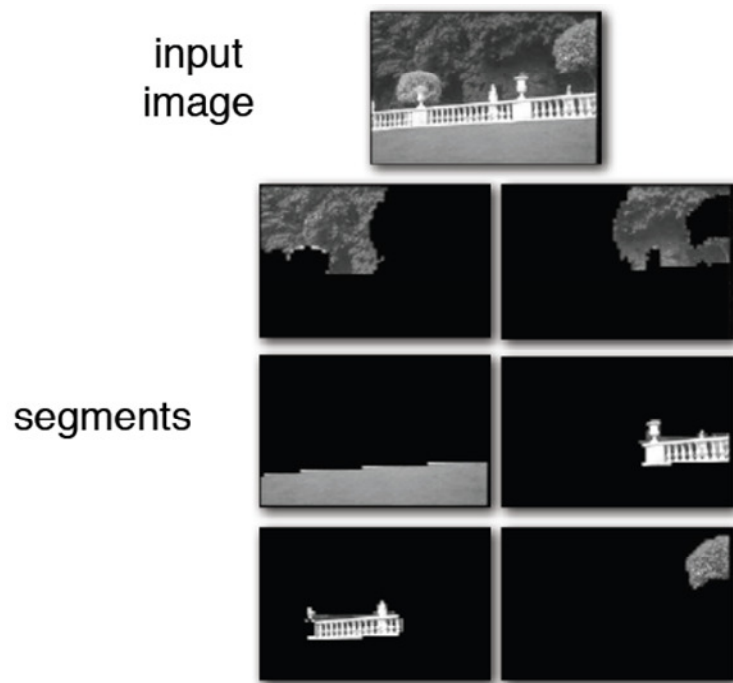


Figure 10. Example of NCuts on the input image. The number of regions set as the input was 6 and therefore resulted in the fence being split into two objects rather than one[22].

The Berkley Segmentor is the other algorithm that was studied, which could be used as a foundation starter for the design of an algorithm for the segmentation of these types of images. What is good about this algorithm is that it incorporates human annotators to improve the segmentation result. Here is an example below of the input image and what types of results it produces.



Figure 11. Example of the Berkeley Segmentor[25].

3.3 Error Calculation

Utilizing these two algorithms, we can benchmark their results and compute which performs better. Basically, we will have two identical images, one from the output of the game, and one from the output of the algorithm. We will then overlap the two

images to compute the amount that each region overlaps each other in the two images, which we will define as the amount of accuracy. We can calculate the amount of accuracy of the algorithm in comparison with the output of the human annotators using a loss function.

$$Loss\ function = \frac{|D \cap GT|}{|D \cup GT|} \times 100\% = accuracy\ \% \quad (1)$$

Where D is the area of the region produced by the algorithm and GT is the ground truth, which is the area of the region produced by the end-users. Refer to Figure 12 for an example of what this may look like.

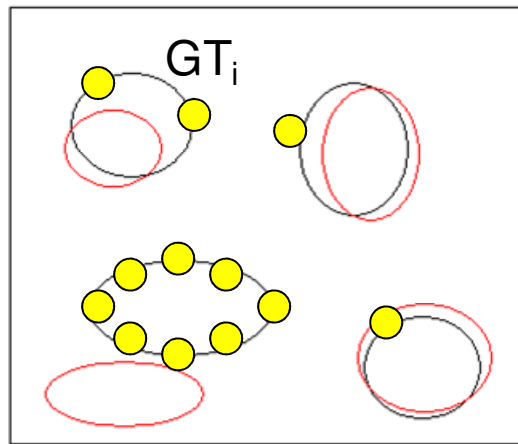


Figure 12. Example of two images overlapping each other where the circles in black are the annotations done by the end-users and the circles in red are the segmented regions produced by the automatic image segmentation algorithm. Each yellow circle represents a click of the mouse done by an end-user when annotating the regions.

Based on the loss function, we would like to determine the amount of overlapping regions over the total area of the two regions, D and GT. Using a simple matlab code we can identify the overlapping regions by their pixel value. For example, the region D

might be identified with pixels of value 1 and the region GT might be identified with pixels of value 2. We can add the two images together to find the overlapping regions that consist of pixels of value 3. If the amount of overlapping region, $D \cap GT$, is equal to the amount of total region, $D \cup GT$, then we can say that the algorithm is approximately 100% accurate. What we would also like to consider is the cost of the annotations, and balance that out with the accuracy produced. Ideally, what we want is for the cost of the annotation, the amount of clicks required by the user to define the boundary of the regions using say LabelMe, to be as minimal as possible and have the algorithm be as accurate as possible (large total overlapping area over combined area of both annotations). We can calculate this using the equation below.

$$\{D_i\} = \max_{\forall i, D_i} \left[\sum_i \frac{|D_i \cap GT_i|}{|D_i \cup GT_i|} - \sum_i R_i(GT_i) \right] \quad (2)$$

Where R_i represents the cost of the annotation for some region i .

Using this approach to test the accuracy of the algorithm, we will now discuss how we can obtain these annotations from end-users.

Chapter 4 - Method

In order to obtain annotations from end-users we need to figure out a way to make it fun, so we will go over what is a game, how we can turn our task into a game such that we are able to get high quality results, and a study that was done to understand the behavior of end-users.

4.1 What is a Game?

To help understand what makes games fun we look at the definition of a game: “A game is an interactive structure of endogenous meaning that requires players to struggle toward goals.” In addition, a game consists of pleasures such as sensation, fantasy, narrative, challenge, fellowship, discovery, expression, and masochism[4]. Sensation is any type of visual, sound, tactile, or muscle pleasure whether it be from a Wii controller vibrating in your hand, music, animation, the landscape, graphics, or the characters in the game. Fantasy is analogous to that “sense of disbelief” that helps with the endogenous meaning. For example the fantasy could consist of the characters in a game that does not exist in real life such as wizards or monsters, etc. The narrative could be the story of the game that gives the game some structure. The challenge is the struggle within the game that keeps the player interested in continuing to play. It is a balance between not too hard and not too easy. The fellowship is the community within the game whether you are playing with others or just yourself. Discovery is another aspect that keeps the game interesting because it gives the player different things to explore and find, which, for example, can be new areas in the game that the player have not yet seen. Expression is basically allowing a player to express themselves uniquely such as being able to personalize their character within the game with different items to wear. Masochism means ‘submission to structure’. Games abide by a certain structure that must be followed in the game in order to win. With the understanding of what a game is, we can now study how to turn this game into a meaningful task.

Crowdsourcing is a technique that is applicable to this situation, which would help us create a larger database of annotated images in which the public can get involved in voluntarily. Amazon Mechanical Turk (MTurk), a popular crowdsourcing platform, has proven useful in large-scale image label collection. An MTurk task creation needs to take into consideration (1) what motivates workers to complete tasks; (2) what is the best way to design a task; (3) what is the best way to ensure high quality results; and (4) how can one use crowdsourcing in a cost effective manner[8]. An example of a successful MTurk is the ESP Game, mentioned previously.

A common method for this task is games with a purpose (GWAP), which we will study what they are and consists of to be successful. GWAP is a way of turning a mundane task into something fun that people are willing to participate in to produce meaningful output that can help us improve computer vision algorithms. In addition, it can help to reduce the load of annotations on people as well as reduce the cost of annotations. A basic game template consists of: (1) Networked individuals accomplishing work, where collaborative effort by a large number of people work together to accomplish a task that would otherwise be too difficult and time consuming for a single individual or a small group of people; (2) Open mind initiative, where the task aims to leverage human effort via “intelligent” software. This involves humans helping the computer to learn what is commonsense; and (3) Making work fun[6]. People are more willing to participate when the task is fun. Many researchers have suggested that incorporating game-like elements into user interfaces could increase user motivation and the playfulness of work activities.

4.2 Study of End-Users

In order to understand the outputs to expect from end-users surveys were created using a randomly chosen PB image and an OSB image. From this, it can be observed how different people annotate the same ‘blob’ in the PB image or wood particle in the OSB

image. Approximately 50 surveys for each material were created in which transparencies were placed over the two images and random people were asked to annotate a specified region (shown by a red cross 'X' in Figure 13) using a sharpie.

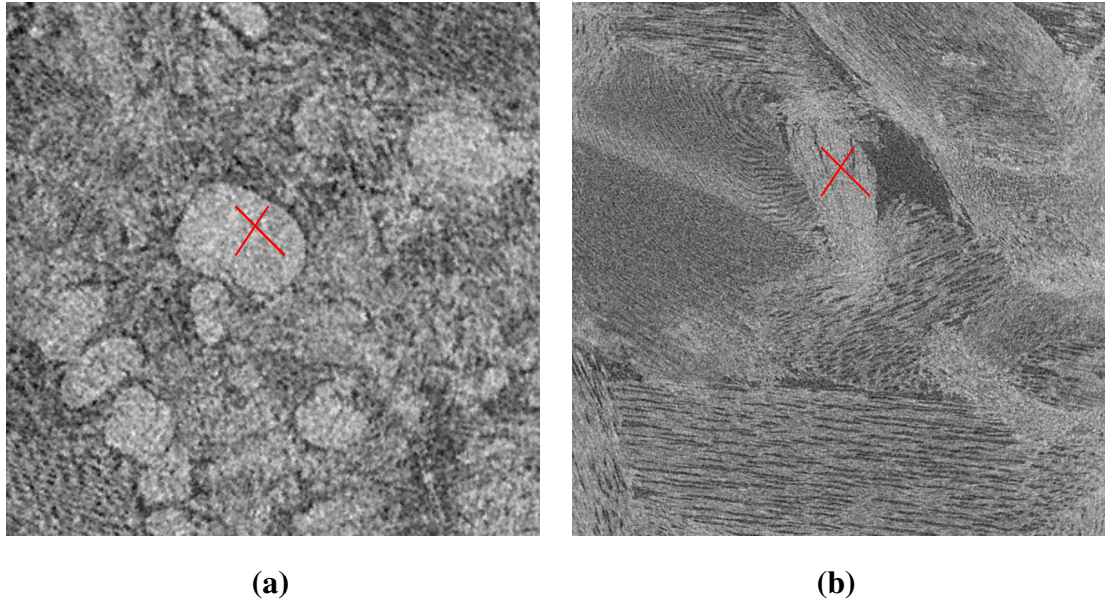


Figure 13. (a) PB image used for the survey where the red 'X' indicates what the person should outline. (b) OSB image used for the survey where the red 'X' indicates what the person should outline.

The centroid and area of people's annotations were calculated using a simple matlab program. These were the results of the survey's:

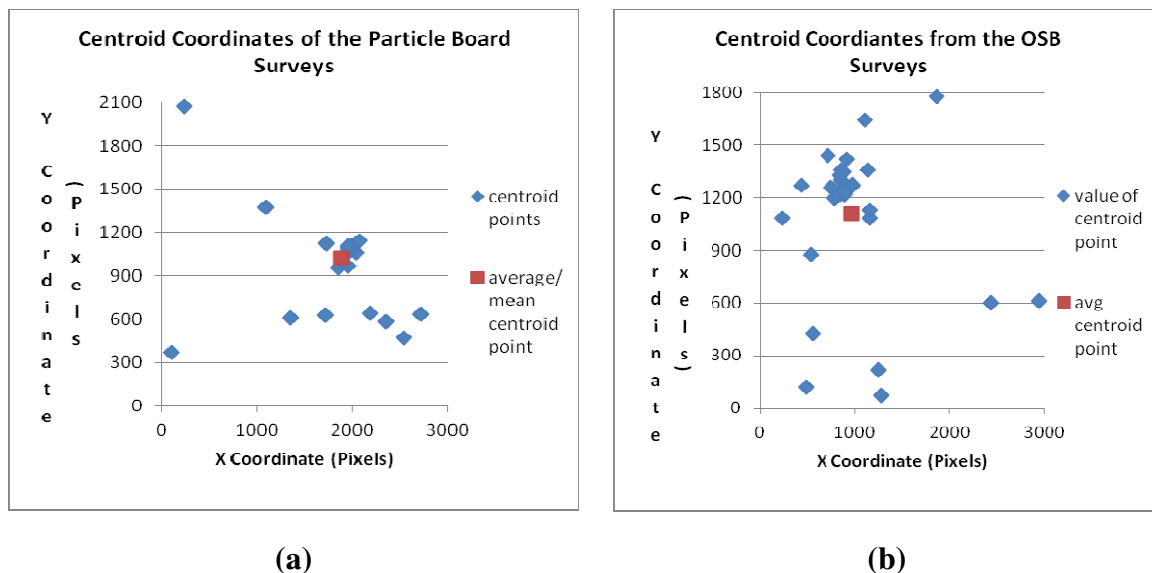
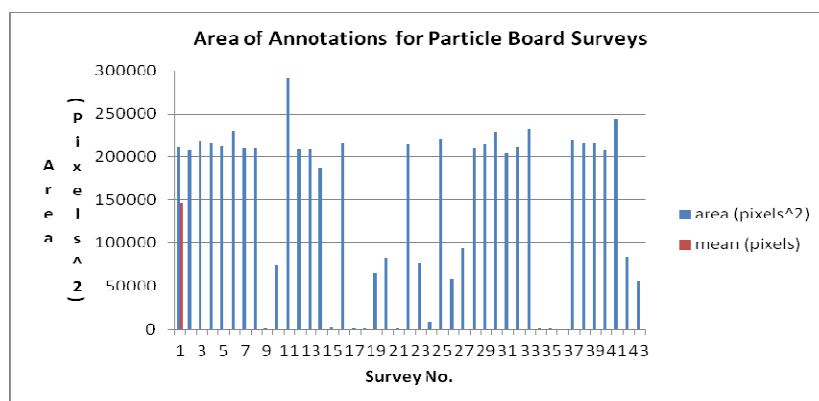
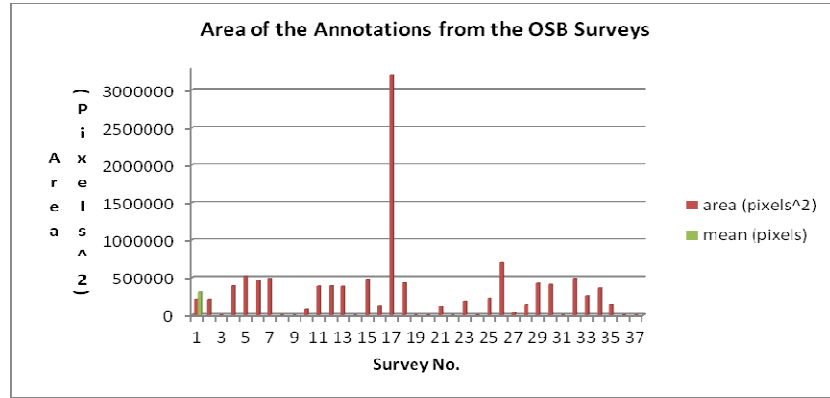


Figure 14. (a) Graph of the centroid points (in pixels) of the annotations from the PB surveys, and the average centroid point calculated based on all the PB surveys. (b) Graph of the centroid points (in pixels) of the annotations from the OSB surveys, and the average centroid point calculated based on all the OSB surveys.



(a)



(b)

Figure 15. (a) Graph of the area (pixels²) of the annotations from the PB surveys, and the average area calculated based on all the PB surveys. (b) Graph of the area (pixels²) of the annotations from the OSB surveys, and the average area calculated based on all the OSB surveys.

From Figure 14 we can see that the average centroid was closer to the majority in the PB surveys than the OSB surveys. But we must also consider the area of the annotation rather than just the centroid. The area of the annotations was more consistent in the PB surveys than in the OSB surveys, as can be observed in Figure 15. Though, the average area of the annotations for both the PB and OSB were not quite close to the majority. The reason for the average area in the PB surveys being not consistent with majority is due to the outliers, which was caused by people not fully annotating their regions. This can be seen in the surveys below.

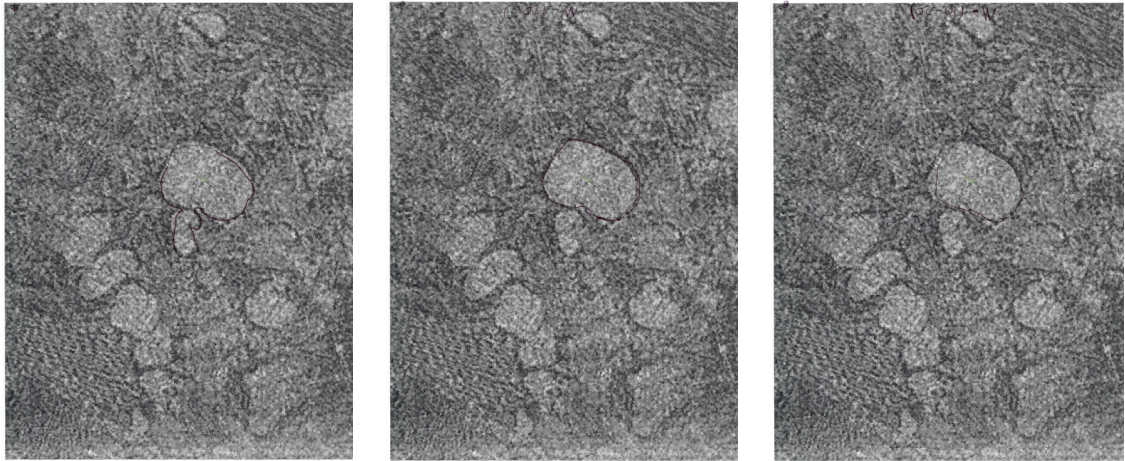


Figure 16. Three different example PB surveys annotated by random people.

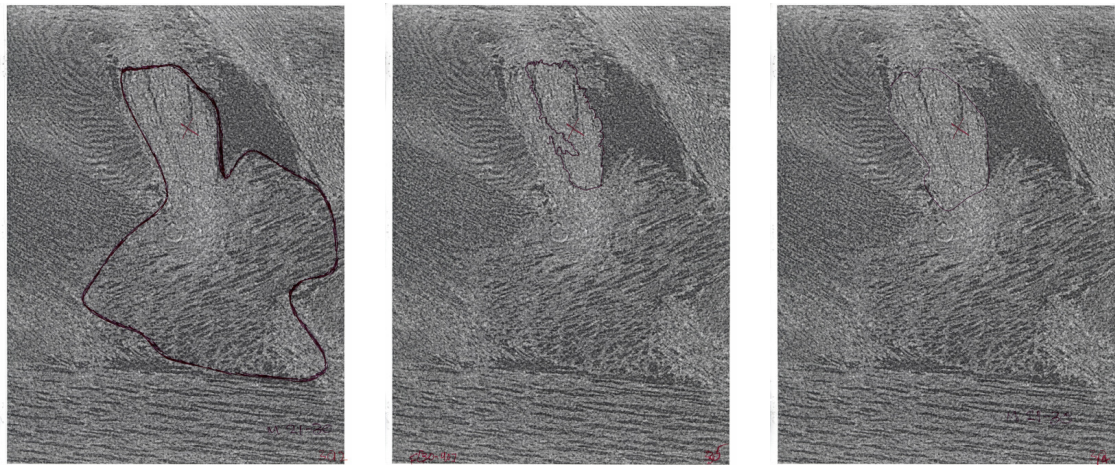


Figure 17. Three different example OSB surveys annotated by random people.

From the PB surveys we can observe that people were fairly consistent in determining what the ‘blob’ was, but the difference lied in how they annotated that ‘blob’. Observing the surveys from left to right, we can see that the first person decided that the ‘blob’ consisted of the large circular region as well as the region below it. Whereas, the other two, to the right of this survey, decided that the ‘blob’ was just the large circular region. The last survey, farthest right, shows an incomplete annotation, which makes it hard for

the algorithm to determine accurately where the centroid lies as one would have to assume where the boundary would be. And making such assumptions added to the error as well. Looking at the OSB surveys we can see a larger variation in the way people perceived what should be annotated. The first survey, farthest left, consists of an annotation that practically consumes the whole image. Whereas, the surveys on the right of it consisted of smaller annotations that appeared more similar to each other than compared to the first. Therefore the last two surveys would have centroids and areas that are approximately more similar to each other than the centroid and area in the first survey.

To get a better understanding of which annotations are more accurate, these should be compared against annotations done by experts in the Wood Science department. This would allow us to compare the amount of variation in annotations done by experts versus a random group of untrained annotators.

It was due to these results that the game was designed to be played among many players that had to agree on the same region of interest as well as annotate this region in such a way that their centroid and area of their annotation was close to the average. In addition, it is required that the same image should be annotated multiple times in order to place a higher confidence on the annotation of these 'objects' within the image. As to how many annotations are necessary in order to be confident enough that the region is correctly annotated is not yet known. From looking at the graph, it appears that 50 or less people may be sufficient for the PB image set, but more than 50 people is necessary for the OSB image set. Being that the OSB images appears to be the most difficult to decipher, as to what regions contained the same strands going in the same direction, if the players were capable of moving through the images back and forth they might have done better as this would allow them to be able to see the movements of the strands as one moves through the material. This technique was used by the high school students multiple times during the annotation of these images.

Based on these results, in order to handle for the outliers and variation in data, the LabelMe technique of annotating images should be incorporated. In addition, images should be repeatedly annotated by a large amount of people in order to place a large confidence on which annotations are correct. The annotations with a higher degree of agreement should be kept in one database while the others should be kept in another for further checks so as not to pollute the dataset.

Chapter 5 – Results

From what we studied about games and end-users we have designed a basic game outline to meet the requirements of a game in order to try to make it fun. We will then explain the basic game and its details that will explain how we can obtain high quality results and other implementations to help the players with this task.

5.1 Game Outline

Considering what was mentioned previously, such as what a game is and how we can turn this task into a successful GWAP while considering for the different behavior of end-users to expect, an outline for a game was created. The goal of the game is for players to annotate regions of interest ('blobs' in the PB images, wood particles in the WPC images, and regions or strands oriented in the same direction in the OSB images). To add to the struggle of the game, players will compete against others to annotate these regions before the others. A loose structure will be applied to the game allowing players to annotate any regions within the image. In addition, the endogenous meaning was not found to be applicable to this game for the time being.

In order to make the game pleasurable we include music and animation to add to the aspect of sensation. To create a sense of fantasy, 'sparkles' are included during the animation. The 'sparkles' can sometimes appear magical; hence providing the pleasure of fantasy. The narrative will consist of someone explaining how to annotate the images, so that players will know how to play before actually playing. The challenge of the game will be that players must try to annotate the same regions as the others and be the first four to annotate this same region. In addition, the players must try to be as accurate as possible when annotating. We have included other players in this game to add to the fellowship aspect part of the game as we have determined would add to the 'fun' of annotating rather than by doing this task individually. There are other reasons for

including other people in the game, such as helping us to ensure high quality results and get rid of outliers and cheaters, which is explained later in the details of the game. As for the discovery aspect, players will constantly be exposed to a new image which in a sense is like discovering a new area in a map of a game, which helps to keep the game interesting. Expression is one of the strong pleasures in this game. A player is able to express themselves with a username as well as the color of their choice to annotate with. Another option could be a profile picture that would pop up when a player wins. Masochism is the basic structure included in the game such that players must try to annotate regions of interest similar to the other players in order to succeed.

5.1 Basic Game Play

First the general game is introduced and then we will go into the details of it such as how to break down this task into smaller tasks so that players might find it more interesting, how players can achieve points, what data will be collected, how to handle for different number of players available, and how to handle cheating so as to obtain good results.

The game proposed for this study is ESP 2.0, which is an interactive game between seven random players that will be matched up to play together. All players will be assigned the same task of annotating (segmenting out the object/particle to be detected in) the same image. The goal of the game is for all players to annotate the same ‘objects’ in the image whether it is the adhesives in the PB images, the wood particles in the WPC images, or the same regions of strands in the OSB images. Once four players have annotated the same region, before the other players, points are given out to these four players. An alternative would be to allow the first six players to have a chance to gain points, but only the four closest to the average would gain points while the last two do not. If the image is too difficult, the players can choose to “pass” in which another image will be shown to them and no points will be awarded to any of the players. The players will continue to annotate the same image until at least four of them have agreed that the image is completely annotated. When this occurs, they can move on to the next image. In

this case, if a player has decided that their image is fully annotated, they can no longer annotate this image and will continue on to play with another team or a computer, in single mode, which is later explained, so that they do not have to wait for the other team members to finish. Other players are not notified of this occurrence (if a player decides the image is completely annotated), so that they are not influenced by other players' decisions. We incorporate this tool of skipping and deciding if an image is fully annotated so that players do not start to hallucinate 'objects' in the image if one really does not exist such as can be seen in Figure 18, which is fairly difficult to determine if a 'blob' exist in the PB image or not.

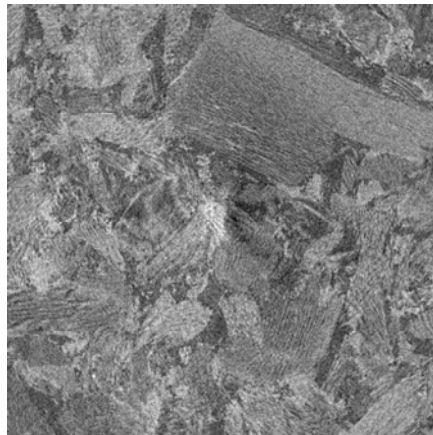


Figure 18. PB image, which is an example of how complex these images can get in terms of identifying the objects or in this case the 'blob'.

We now address more of the details of the game as mentioned earlier.

5.2 Break Down Task

The hardest challenge of creating a game is turning this mundane task into something fun, which can become quickly boring as many of the images require many annotations just in a single image, resulting in players to have to constantly look at the same image

for a period of time that might exceed their attention span. A solution to this would be to break down the image into square regions so that it appears to the players that they are moving on to a new image. In this case, the image would be cut up into 8 parts of overlapping regions to help merge the data together as some of the regions of interest may lie on the edge. Refer to Figure 19 for an example.

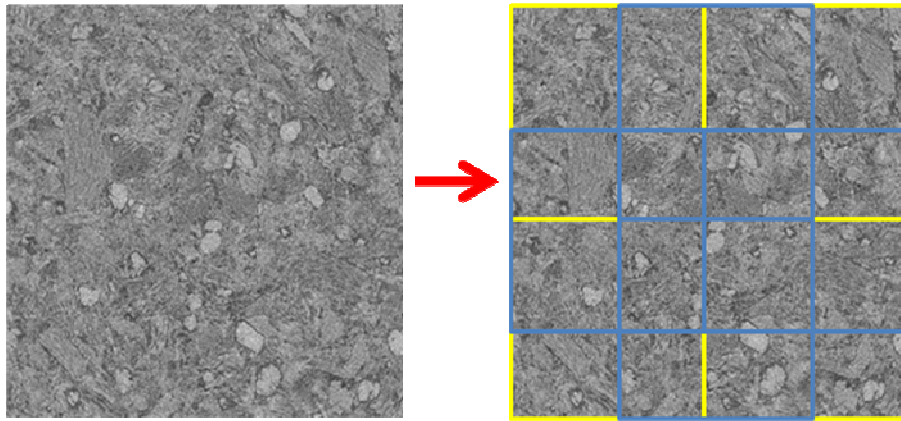


Figure 19. Example PB image, where it would be cut up into 8 slices. The 4 blue squares indicate the overlapping regions to handle for the edges made by the 4 yellow squares.

5.3 Game Points

How points can be achieved by players will now be explained. The exact number of points associated with each action is unimportant, but the general proportions will be mentioned. Furthermore, the different point strategies used by ESP 2.0 to keep the players engaged are mentioned.

To start off players with points (new users), will enter a learning stage in which they will have the chance to get familiar with the tools as well as learn the rules to help them understand what needs to be done in order to be successful. This will help to lessen their frustration, which in turn will help to increase their ‘fun’. During this stage, they can

earn points by replicating the same annotations done in previously annotated images in our database. If the user has decided that they fully understand what needs to be done, they can move on to the next round and play with others. Only a limited number of images will be provided in which they can gain points for from this round, so that they cannot continually play in this round to get higher scores. Any user can enter this stage even if they are not a new user, but cannot gain more points from it once the max number of points in this stage has been received. Refer to Figure 20 for the example flow chart of this.

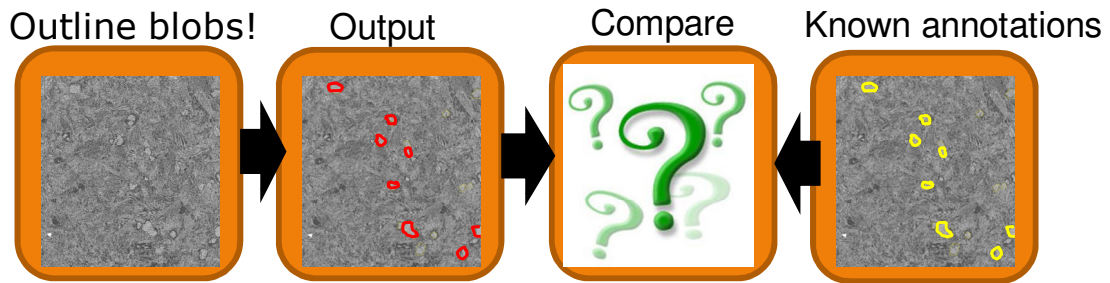


Figure 20. Flow chart of the learning stage in the game. The players will be presented with an image that they will annotate from which their outcome will be compared against the known database containing the annotated version of the image.

Points are given equally when all four players annotate the same region before the other three. Having players compete for points will help to motivate them to play as it involves some challenge for them to overcome in order to get ahead of others. For each same region that they agree upon in the image, they are awarded the same number of ‘bonus’ points. In addition to these ‘bonus’ points, we give all players points for annotating the same region, but these points are based off of how close everyone is to the average centroid and area from that region that was annotated. This will give people the incentive to be accurate. Because, though, some people may beat others in annotating the same region, they will feel inclined to do a poor job about it, and we do not want to reward this behavior. For example, someone may decide to annotate the whole image in

which case they would always beat everyone in the annotations because that region overlaps all other regions that are annotated, but this would be wrong (refer to Figure 21 for an example of this).

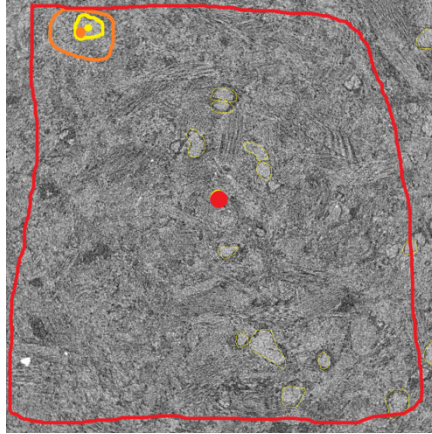


Figure 21. Example image demonstrating how players might cheat the system by annotating the whole image in which case they would earn points, by beating other players at annotating the same region, because their region would overlap all regions of interest.

So in actuality, even if you annotated the region before the others, but in a poor manner, they can still get fewer points than a person that was more accurate at annotating this region, but failed to do it quicker than the first four. This is because the bonus points will be some set constant like 1, whereas the amount of points you can receive based on how accurate you are can exceed this amount. If for some reason, four people cannot come to an agreement and they have considered the image completely annotated, then no points are awarded to any player and a new image will appear. This image, though, is not discarded, but kept in a separate database to be checked among other images in the database (images will be repeatedly played so as to get a better accuracy on the annotations). If these rejected images are played several times and are never considered to be a match with any other, it will then be discarded. When at least three players agree to “pass” an image, no points are awarded to either player and a new image will appear.

We would like players to stay consistent with the annotations they make as each region of interest is actually a 3D object within the material from which the volume will be calculated. To help them stay consistent, players that have annotated the same region will be shown this same image with the annotation again in the following image. This will allow those players to see if that region of interest still exists within the next image and if so they have the advantage to annotate it before the other players and gain more points. This is only shown to the players that have previously annotated it, so as not to influence others that have decided it may not exist.

Due to the different complexity of the annotations required by different composite materials, different points are awarded to each, and each composite material represents a different ‘level’ within the game. The WPC image set will be used in the first level, then the PB image set, and last of all the OSB image set (refer to Figure 22).

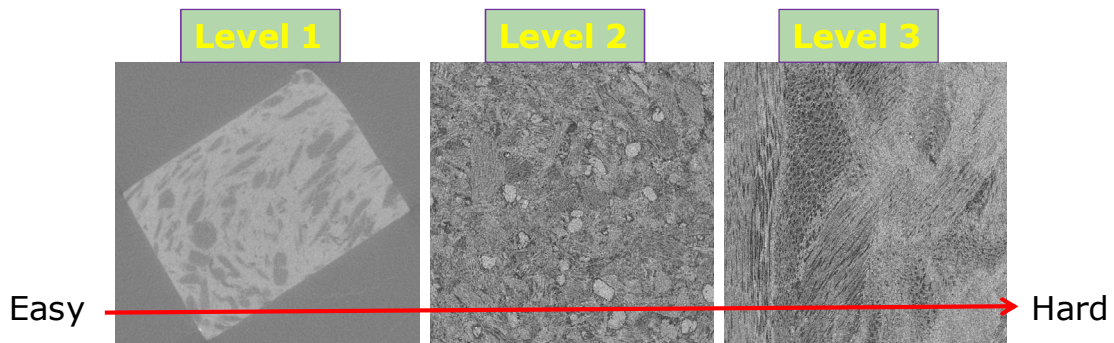


Figure 22. Levels in the game. The first level will consist of WPC images for the players to annotate, followed by the PB images in the next level, and the OSB images in the last level of the game.

After about so many points are awarded to each player within the first level of the game, they will then advance to the next and so forth.

The reason for creating levels is to instill challenge within the game while reinforcing incremental success, which helps to encourage them to continue playing. (Of course some people may never reach all levels and thus another option must be created allowing players to advance to the next level if they choose to. Further details need to be worked out as to when to allow this to happen.)

5.4 The Single Player Game

ESP 2.0 is a seven player game, but there may be times when there are not enough players to group up everyone thus a single player mode is created. In this case the single player is matched with a computer.

Previously annotated images will be used as input into the game in which the results of the single player will be compared against. This will strengthen the confidence of the annotations done in that image as well as the regions that remain in question (regions that have been annotated by only a few players or none at all). In addition, times can be saved for each annotation done in the image, so that the player can play against the range of times done for that particular region. For example, a random set of 6 (6 is chosen to obtain the results of 6 other players) images where the confidence levels are low compared to the other images in the dataset are chosen to be checked. For each annotation within the image, the times are saved by the players that annotated each region. For a particular region of interest, the fastest person to annotate that region may be 30 seconds, and the slowest person, among the four in the set, may be 1 minute. If the single player can annotate that region within that time range, they can gain ‘bonus’ points.

If in the case that previously annotated images are not in the database, we do not deny the player to proceed to play as any help should be accepted. So in this case, we will allow the player to annotate and give them points for each annotation and save these

annotations in a different database to be checked by other players. To ensure that the player is producing valid results, they must be tested. This can be done by randomly presenting the player with an image in which the correct annotations are already known or for which an image has been annotated many times by other players and therefore the confidence of it being correct based on previous annotations of the same regions are high. If a players output does not match with these results they should be noted as suspicious and their results should not be considered. Therefore, if 50% of the images presented to them are previously annotated, and this player produces results that always agree, then we can be 50% confident that this players results are accurate and this probability will increase with repetition.

5.5 The Two Player Game

When there are less than 8 people we must pair off the players. In this case, both players are given the same image as before. When both players annotate the same region of interest, points are awarded to each player, and then a new image will appear for them to annotate. The more objects the players annotate and agree upon, the more points are awarded to each.

5.6 Cheating

We would like the output of the game to produce accurate results, which can be difficult if players start to cheat the system in such a way that they are rewarded with points in the game even if they produce bad results. ESP 2.0 is a collaborative game in which multiple people work together to maximize their score. When each person works independently without the influence of others we can gain correct information. In the case that people know each other, the data will become corrupt. For instance if people know each other or are able to communicate then they can tell each other which regions to annotate, which

may end up not being any region of interest at all. Therefore, the following ways to avoid these situations are mentioned below.

There are a multiple anti-cheating mechanisms, some of which have been inherited by Peekaboom. These mechanisms have been set up to detect and deal with cheating. Cheating was tested to be rare in Peekaboom, but precautions must be taken.

- To prevent people being paired off with friends because they start the game at the same time, Peekaboom proposed the idea of waiting n seconds before pairing people, where n is the number of seconds until the next matching interval. As more people play, we can ensure partners are random and that no one can be paired up together just because they clicked the start button at the same time[7].
- IP addresses will be checked, which prevents people within the same geographical proximity or people sitting close to each other to be paired[7].
- We limit their ability to cheat by not allowing them to communicate with each other. This prevents them from chatting and telling the other which regions they are annotating.
- We limit their ability to annotate any region by requiring that three other players must agree on the same region. Due to this, people will be unable to annotate regions that are not of interest just to gain points. And requiring that four people to agree on the same region helps to prevent this as it would be less common for people to agree upon the same regions if these regions are chosen without it being of interest. Points are based on the area and centroids of the annotated regions that are agreed upon by multiple players. This is because some people may decide to annotate the whole image or annotate a region in a sloppy manner such that they gain points because the regions they annotated are the same.

- In addition, the data collected is aggregated from other players to eliminate outliers[7]. To do this, the images are put into play multiple times so that regions that are missed are given multiple chances to be detected and it gives those regions that are detected a higher confidence that they are correct regions of interest. Also, many of the images have multiple regions of interest, more than perhaps people are willing to put in the time to do, and this will help to reduce regions that are of interest from being skipped.

5.7 Implementation

We will now discuss the implementation of the game, the tools available to the players for annotating, other resources the players can use to help decipher objects as images become complex, and other add-ons to keep the players interested.

The game is to be implemented under the client-server model. Like Peekaboom, the client application is a Java applet, and the server is written in Java. The server will save the matches and scores to a disk. The collection of data will then be formatted into desired formats (e.g. images saved as tiff).

This GWAP is implemented as an output-agreement game where players are matched up with each other randomly[6]. In each round, players are given the same image to annotate. The game instructions indicate that players must try to annotate as many regions of interest as they see and produce the same output as the other team members. Players are not allowed to communicate with each other in any way. In order to 'win', players must produce the same output, but are not required to do so at the same time. In our case, players compete to match up with the results of four other people before the others.

The implementation of the game contains other features to improve game-play that is mentioned below.

Annotations: How the players will annotate (outline) the regions of interest will be done using a computer mouse or if using a touch screen device a player can use their finger. A line will be created and show up behind the mouse or finger as it moves. The line will have a width of 1 pixel, and the color of the line can be chosen by the players such that it is a bright color of high contrast against the given image (previously shown in black and white). It was also suggested that we allow for people to use joysticks or the keys on the keyboard, and allow players to have a ‘car’ that they can control that leaves a line behind. Alternatively, players can use the technique used by LabelMe[1].

Eraser: Mistakes happen often. Sometimes players mark an area by accident or need to re-annotate their region if they made it too large or too small. For this case, players will be given an eraser so that they may erase any mistakes they make while annotating as we would like the annotations to be as clean as possible.

Zoom: Some of the regions of interest are very small and can be hard to decipher, therefore, a zoom tool is necessary for the player to be able to zoom into the image when necessary.

Move back through images: We need players to stay consistent with the regions that are annotated. If they are not consistent, this will affect the calculation of the volume of that ‘object’ and that ‘object’ will be split up into parts. Thus for regions that are annotated by at least the first four players, this image will be shown again to those players that have annotated this region when presented with the next image. Often times it is hard to stay consistent or be able to decipher if a ‘blob’ is a piece of adhesive or not without being able to view the other images. Having the capability to view the other images allows a person to better observe the movement of these ‘objects’ and how they differ from the surroundings in the image. Thus by allowing players to move back to the

previous image or be shown the previous image will help players to be consistent. In addition, we allow players to move forward through the composite material to the future images, but also give them a button to move back to the current image they are annotating, so as not to get lost. Refer to Figure 23 to see an example of consecutive images with the same ‘blob’ annotated in each of them in the same location.

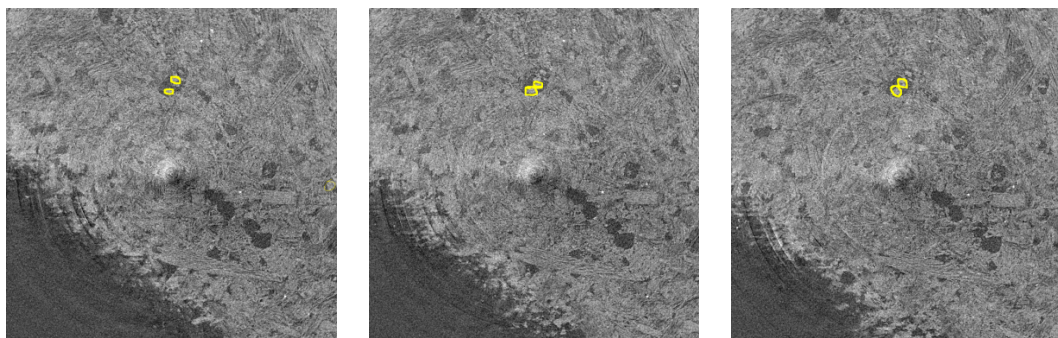


Figure 23. Three consecutive PB images showing the same ‘blob’ being traced throughout each image. This is an example of the necessity for consistency when annotating each image as the ‘blobs’ need to be traced from beginning to end because all the images stacked together creates a 3D structure of the composite material. In other words, each image is just a slice in the material.

3D structure: With the help of the 3D structure, which is basically just a sequence of images stacked together, the players can use this to help identify better as to what a ‘blob’ in the image is. This was found to be useful for the high school students especially when annotating the PB images. An example of a set of the PB images stacked together is shown in Figure 24.

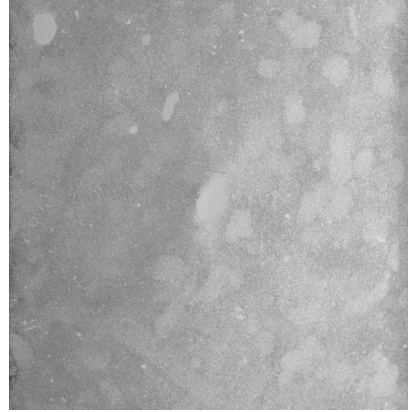


Figure 24. 150 consecutive images taken from the PB dataset and stacked together to create a portion of the 3D structure of the PB composite material using ImageJ. This is a screenshot showing one face of the structure, which is rotated about the y-axis at a very small angle so as to observe the 3D version of the ‘blobs’ within the material.

This is just a screen capture of the 3D structure that is made available using ImageJ that allows one to rotate the object around any axis by 360 degrees in order to get a better visualization of it.

Animation: In addition, to keep the game pleasing to the eye and interesting, winning players will have their lines animated with sparkles to signify the gaining of points, with players’ lines closer to the average sparkling even more. Text might also appear, for instance: “Casey Wins!” in the color of the winning line. Points can be shown flying to each players box that shows their total points. Animations are there to be distractions from the relative mundane task.

Top score list and ranks: We display the top scores of the day as well as the top scores from the overall games played. And the top 10 players with the highest scores will be displayed. Everyone will be ranked so that they know where they stand among the other players.

Chapter 6 – Discussion/Evaluation

With the design of the game, we will now discuss the outcomes of this method of approach such as the data that we will be able to collect from the game, what other applications we can apply this game to, a few ethical issues that we must consider when working with end-users, and the overall outcome of this thesis.

6.1 Collecting Image Data

The goal is to create a database for the testing of the automatic image segmentation program. Here it is discussed as to what information is collected by the game.

On input, an image will be given to the players. All players will proceed to annotate the image. The similar regions, in which four players or more have all annotated, will be saved and any other annotations that are not agreed upon will be saved in another database to be checked against future players. If future annotations do not match up with these regions, they will be discarded. The area of those regions that are considered legitimate will be saved along with their locations in the image. For annotations that are consistent throughout each image, the volume of those regions will be calculated, which will provide further information as to how large, for example, the adhesive is in the composite material. It should be noted that the images are played repeatedly to get a better confidence about the accuracy of the area and volume; therefore, these calculations are capable of being changed over time.

6.2 Generalizing Our Approach

This section presents how this approach , of solving a problem by having people play games online, can be generalized to many other problems in computer vision and Artificial Intelligence. For example instead of finding ‘blobs’ in wood composite materials, we can use this approach to understand the internal structure of other materials that are being studied in the Material Science department, so they can better understand those materials’ strengths and weaknesses to help them produce stronger materials. In addition, we can use this approach to help improve segmentation algorithms, and help to determine how many clusters are necessary for different scenes in an image. This technique can also be applied to other fields in medicine for researching MRI images of the brain, which can help doctors be more accurate when looking for different anomalies in the images.

We now discuss the ethical issues that must be addressed when creating a game for end-users.

6.3 Ethical Considerations

This section presents the ethical issues we must consider to ensure players have all the necessary information made available to them when accepting to participate in the game. At the beginning of the game, players are informed of the purpose of the study, and can choose to not play. If they proceed it is by means of their own consent to volunteer knowing the conditions. Players are informed that the purpose of the game is to help create a database to be compared against the output of an automatic image segmentation algorithm to calculate its accuracy in detecting these ‘objects’.

In the next section, we will sum up the outcomes of this study based on all the above.

6.4 Outcomes of this thesis

Based on our work, we now have a formative study on how to engage end-users to annotate ground truth in CT scanning images, a dataset of approximately 5300 annotated images consisting of the three different types of composite materials, and a function for calculating error between the ground truth and the algorithm. The function was derived because calculating error of the algorithm for these types of images is not easy. For meaningful objects like cars, we can say with 100% confidence if the algorithm did not work or not because we know what a car is by simple observation. If a car was not detected, when we see a car in the image, we know that the algorithm failed. Or if the algorithm does not match up with the annotations of the annotators, but we see that it is correct and the annotation by the annotators is wrong, we can say that they simply missed the object by accident. Though, in the case of these ‘blobs’, for example in the PB images, we cannot say for sure if there exists a ‘blob’ or not and there are no features yet that define what these are, so we must base the error of the algorithm on ground truth annotations done by other annotators. And because people’s perceptions of these ‘objects’ are all different, we need many people to get involved to help determine where these ‘objects’ are. The more people that agree on a particular region, the more we can say with some confidence if these ‘objects’ exists or not. So if the algorithm were to miss this region, we can say that it has an error based on the percentage of people that have annotated this region (agreeing that there exists an ‘object’ there) rather than outputting that the algorithm failed completely because that is not known for sure.

Chapter 7 – Conclusion

The design of ESP 2.0 is created to be an online game for creating a database of annotated images. Segmenting images is very time consuming for a single person, and they might not always be correct or consistent, therefore, a game is created for this purpose from which multiple people can participate in and have their work checked against others. With constant repetition, checking's, and more people to participate, we can gain a large and more accurate dataset of annotated images faster without any cost. This tool can be used throughout the Wood Science department for which they need to identify the internal structure of other composite material that are in question, and help further their studies in renewable resources for a more green and sustainable economy. In addition, ESP 2.0 has various other applications in computer vision for segmentation of other images for which it can help a computer to learn and segment automatically. This game can be applied not only to create a database of images for testing, but also to help create a training set of images used by Artificial Intelligence.

Based on the different outputs observed from the surveys, the algorithm used to read in these results must account for human annotation errors and variations of types. If not, the accuracy of the segmentation algorithm will suffer from false errors hindering the advancement of the study of segmentation in the computer vision field.

In addition, we now have a database consisting of approximately 5300 images consisting of the three different composite materials, which will be made available throughout the computer vision field to help further the study of automatic image segmentation in complex textures.

Future work in this study would be to implement this game and test for how effective it is in obtaining accurate results, and how 'fun' people consider this game to be. For this, pilot studies would need to be created. We would also like to benchmark

other algorithms being used in computer vision for segmentation of images, other than the NCuts and the Berkeley Segmentor, to determine how well they work. Last of all, we would like to implement our own algorithm to segment these images and test its output.

Bibliography

- [1] B. C. Russell, A. Torralba, K. P. Murphy, W. T. Freeman, “LabelMe: a database and web-based tool for image annotation,” *International Journal of Computer Vision*, pp. 157-173, vol. 77, no. 1-3, 2008.

- [2] Carl Vondrick, Deva Ramanan, Donald Patterson. "Efficiently Scaling Up Video Annotation with Crowdsourced Marketplaces," *European Conference on Computer Vision (ECCV) Crete, Greece, 2010*.

- [3] Chow, A., and Schafer, C. “Gamification of Annotating Images,” *SAE report*, 2011.

- [4] Costikyan, G. “Game Design Theory,” *Game Developers Conference*, 2002.

- [5] Todorovic, Sinisa. and Muszyński, Lech. “Advancing Bio-Based Composites by Automated Image Analysis,” *OSU CS Research*.

- [6] von Ahn, L., and Dabbish, L. “Designing games with a purpose,” *Communications of the ACM* 51, 8 (2008).

- [7] von Ahn, L., Liu, R., and Blum, M. “Peekaboom: A Game for locating objects in images,” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montreal, Apr. 22–28). ACM Press, New York, 2006.

- [8] Wah, C. “Crowdsourcing and Its Applications in Computer Vision,” *UCSD CSE Research Exam*, 2011.

- [9] S. Savarese, J. Winn, and A. Criminisi, “Discriminative object class models of appearance and shape by correlations,” in CVPR, vol. 2, pp. 2033–2040, 2006.
- [10] R. Fergus, P. Perona, and A. Zisserman, “A sparse object category model for efficient learning and exhaustive recognition,” in CVPR, vol. 1, pp. 380–387, 2005.
- [11] C. R. Brice and C. L. Fennema, “Scene analysis using regions,” *Artificial Intelligence*, vol. 1, pp. 205–226, 1970.
- [12] A. M. Darwish and A. K. Jain, “A rule based approach for visual pattern inspection,” *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 10, no. 1, pp. 56–68, 1988.
- [13] R. Basri and D. Jacobs, “Recognition using region correspondences,” *IJCV*, vol. 25, no. 2, pp. 145–166, 1997.
- [14] J. Peng and B. Bhanu, “Closed-loop object recognition using reinforcement learning,” *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 20, no. 2, pp. 139–154, 1998.
- [15] P. L. Worthington and E. R. Hancock, “Object recognition using shape-from-shading,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 5, pp. 535–542, 2001.
- [16] A. R. Ahmadyfard and J. V. Kittler, “Using relaxation technique for region-based object recognition,” *Image and Vision Computing*, vol. 20, no. 11, pp. 769–781, 2002.

- [17] R. Zhang and Z. Zhang, “Hidden semantic concept discovery in region based image retrieval,” in CVPR, vol. 2, 2004, pp. 996–1001.
- [18] I. Weiss and M. Ray, “Recognizing articulated objects using a region-based invariant transform,” IEEE Trans. Pattern Analysis Machine Intelligence, vol. 27, no. 10, pp. 1660–1665, 2005.
- [19] S. Geman and D. Geman, “Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images,” IEEE TPAMI, vol. 6, no. 6, pp. 721–741, 1984.
- [20] D. Mumford and J. Shah, “Boundary detection by minimizing functionals,” in CVPR, 1985, pp. 22–26.
- [21] T. Lindeberg, Scale-Space Theory in Computer Vision. Norwell, MA, USA: Kluwer Academic Publishers, 1994.
- [22] J. Shi and J. Malik, “Normalized cuts and image segmentation,” IEEE TPAMI, vol. 22, no. 8, pp. 888–905, 2000.
- [23] D. Comaniciu and P. Meer, “Mean shift: A robust approach toward feature space analysis,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 603–619, 2002.
- [24] C. A. Bouman and M. Shapiro, “A multiscale random field model for Bayesian image segmentation,” IEEE Trans. Image Processing, vol. 3, no. 2, pp. 162–177, 1994.
- [25] P. Arbelaez, M. Maire, C. Fowlkes and J. Malik. “Contour Detection and Hierarchical Image Segmentation,” IEEE TPAMI, vol. 33, no. 5, pp. 989-916. 2011.

[26] von Ahn, Luis, Dabbish, L. "Labeling Images with a Computer Game," Proceedings of the SIGCHI conference on Human factors in computing systems, Vienna, Austria, pp. 319-326, 2004.