Semi-Automated Segmentation of Microbes in Color Images

Chandan K. Reddy[*a], Feng-I Liu[a] and Frank B. Dazzo[b]

[a]Department of Computer Science and Engineering,
[b]Department of Microbiology and Molecular Genetics,
Michigan State University, East Lansing, Michigan 48824 USA

ABSTRACT

The goal of this work is to develop a system that can semi-automate the detection of multicolored foreground objects in digitized color images that contain complex and very noisy backgrounds. Although color image segmentation is considered a general problem, our application is microbiology where various colored stains are used to reveal information about the microbes without cultivation. Instead of providing a simple threshold, the proposed system offers an interactive environment whereby the user chooses multiple sample points to define the range of color pixels comprising the foreground microbes of interest. The system then uses the color and spatial distances of these target points to segment the microbes from the confusing background of pixels whose RGB values lie outside the newly defined range and finds the boundary of the foreground microbes using region-growing and mathematical morphology. Some other image processing methods are also applied to enhance the resultant image containing the colored microbes against a noise-free background. The prototype performs with 98% accuracy on a test set compared to manually edited ground truth data. The system described here will have many applications in image processing and analysis where one needs to segment typical pixel regions of similar but non-identical colors.

Keywords: sampling, color image segmentation, color dilation, color erosion, closing, error estimation, color distance, microbes, spatial distance.

1. INTRODUCTION

Microbiologists usually color microbes with various stains to increase their contrast and visualization when examined by microscopy. These chemicals stain different components of the microbial cells or detect specific biochemical activities on the surface or inside the microbe. The color of the microbes in the digital image provides useful information on their biochemical, physiological, and phylogenetic characteristics (depending on the stain used) without the need for their laboratory cultivation. It is also possible to create color images with multiple colors, each indicating different characteristics of the microbes. Bacteria that are heavily stained will emit a stronger color. Thus, the qualitative and quantitative information content in color images of microbes can be very useful and important to the microbiologist so long as they can be accurately segmented, extracted and analyzed. The long-range goal of this work is to develop a system that can automate the detection of multicolored foreground microbes in digitized color images that also contain complex and usually noisy backgrounds. In this paper we describe our prototype system designed to semi-automate the process. The computing system described here is a proprietary product of Michigan State University and is described in Electronic Imaging 2003 held at Santa Clara, CA.

1.1. Features and phenomena of images of microbes
Like medical images, images of microbes are highly complex, causing difficulties for image processing and advanced analysis. Through observation, color images of microbes usually have the following conditions that require editing steps before they can be analyzed:

---

[*]Corresponding author: reddycha@msu.edu; Phone: 1 517 353 5964; Fax: 1 517 355 1292; http://www.msu.edu/~reddycha; Media Interface and Network Design Lab, 251 Communication Arts Building, Michigan State University, East Lansing MI 48824. After 6/03, corresponding author is Prof. F. Dazzo (dazzo@msu.edu; Phone: 1 517 355 6463 ext. 1587; Fax: 1 517 353 8649).

1.1. a. **Variable color range for the pixels that comprise the foreground microbial objects of interest.**

Each of the different stains produces a particular range of color and gradient features. The transmitted or emitted color of each stain is defined by its wavelength. However, we cannot transfer such information directly into the computer color model because the wavelength range that is characteristic of each stain may be wide, even within individual microbes of interest. This heterogeneous color range for the pixels of individual microbes in digital images may or may not be noticeable when viewed at 1 X zoom, but is very obvious when one magnifies the image and views the colors of individual pixels comprising the microbial objects. Since the combinational ratio of the RGB color pixels of microbes stained with different fluorochromes can vary, a default color range for image processing may not accurately reproduce the color of each cell for each color stain. Table 1 illustrates the RGB color ranges of various commonly used fluorochrome stains, each measured by sampling individual pixels comprising fluorescently labeled microbes in color images using the eyedropper tool in Adobe Photoshop. How to find the most precise color range is also a problem to be solved. Furthermore, eliminating the background noise without affecting the true contour (hence shape and size) of the microbe is an important task for future processing and recognition steps.

Fluorochrome Stain	Fluorescent Emitted Color	R	G	B
Fluorescein	Yellow-green	54-171	176-255	0
Rhodamine	Red	156	0-1	0-3
DAPI	Bluish-white	4-12	55-92	208-255
Green Fluorescent Protein	Green	0-21	154-172	0-28
Acridine Orange	Reddish-yellow	240-255	134-172	0-28
	Green	108-163	139-159	0

Table 1. Color range of fluorescent light emitted from bacteria stained with different fluorochromes.

1.1.b. Noise and other invalid objects in the background

In grayscale images of bacteria, the background noise appears in high frequency. If we transfer the image to the frequency domain and apply a low pass filter, some of the noise can be removed from the background. However, the backgrounds of color images of microbes are very different depending on the type of microscopical illumination used (transmitted, reflected, fluorescent) and the composition of the habitat (e.g., soil) from which the microbe was sampled. The size of invalid objects in the image background may be smaller, within, or larger than the range that includes the foreground microbial objects of interest (Fig. 1). The background may have a mixture of noise processes. All these characteristics increase the image complexity and make the segmentation procedure difficult and complex.
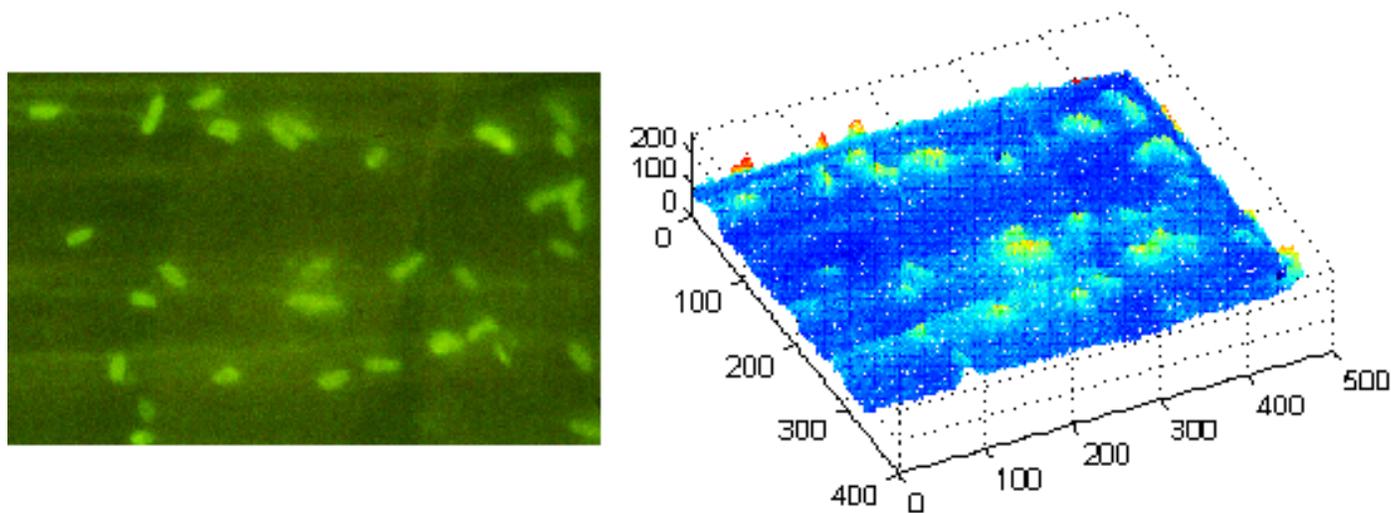

Fig. 1. Fluorescent image of bacteria colonized on the epidermal surface of a plant root. Note the foreground objects of bacteria in the complex and the heterogeneous noisy background.

1.1.c. Halo around the cell in fluorescent stained images

Because the fluorochrome stain emits light itself when excited, the more fluorescent substance the microbe absorbs, the brighter it is. Since microbes typically have curvature, they emit fluorescent light at multiple angles relative to the microscope focal plane, thus a halo of fluorescent light always surrounds the microbes. The color of the fluorescent halo is similar to the color of the microbe itself but of dimmer brightness (Fig. 2), so it is possible to mistakenly include the halo as part of the microbe. If the color segmentation cannot exclude the halo, subsequent analysis of the image would overestimate the size and may also represent a different shape of the microbe. Since sizes and shapes of microbes provide important microbiological information, this error can affect the abundance and morphotype analysis of a microbial community. We would like to eliminate the halo effect during image segmentation to extract a more accurate measure the microbe's size and shape.
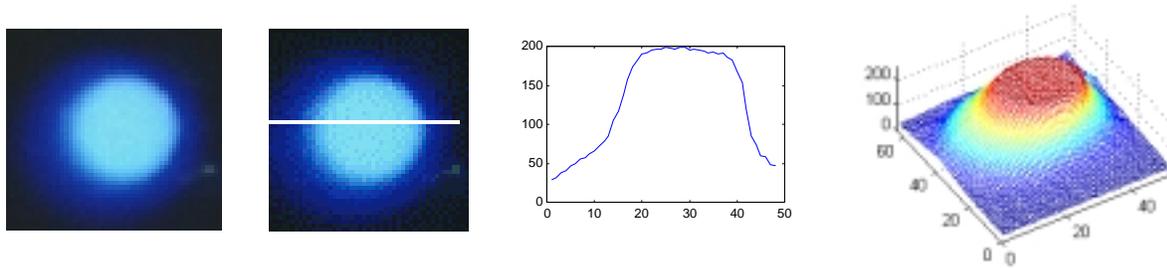

Figure. 2 Characteristics of fluorescent objects the size of bacteria in digital images obtained by fluorescence microscopy. Left to right: Example of a spherical cell emitting bluish-white light [note the fluorescent blue halo]; the same object with a line transect through its center; the corresponding 2D plot of change in blue light intensity along that line transect; and a 3D plot of the surface color intensity of the object in the image (blue channel only).

1.1.d. Darker center of the microbes

The distribution of the color stain on the microbe is determined largely from the chemical reaction between the stain and the microbe's cell components. Because different parts of the microbe have different biochemical functionality, the chemicals are not evenly distributed throughout the microbe. Some stains are freely absorbed and distribute evenly throughout the cell volume, producing a fairly constant distribution of color light intensity across the 2-dimensional projected image of the microbe. Other stains specifically interact with components present exclusively on the surface of the cell and do not enter the cell unless its cell envelope is broken. In this case, the 3-dimensional cylindrical or spherical cell will be brighter near its periphery than its center because more of the stained cell surface sampled near the curved edge of its optical median plane has been projected into a smaller area of the 2-dimensional image. When the image is segmented, the darker center might not be recognized as part of the microbe and becomes a black hole inside the object. The traditional method to process images of blood cells or bacteria is to use the opening and closing operations. If the basic element for opening and closing is chosen properly, the size of a cell will not change significantly. However, this method may change the shape of the bacteria slightly and provide incorrect perimeter information of the microbe.

Figure 3. Bacteria fluorescently labeled with an antibody reactive with a cell surface component. Note the cells' darker centers.

2. OUR METHODOLOGY

In this work, we implemented a system that can identify the foreground regions of interest based on interactive selection of a set of pixels sampled from different areas of the color image. This algorithm is very useful in detecting the microbes. We want to eliminate the image background and detect the foreground colored microbes of interest by performing color segmentation. The steps of color segmentation to find the colored microbes are shown in Fig. 4.
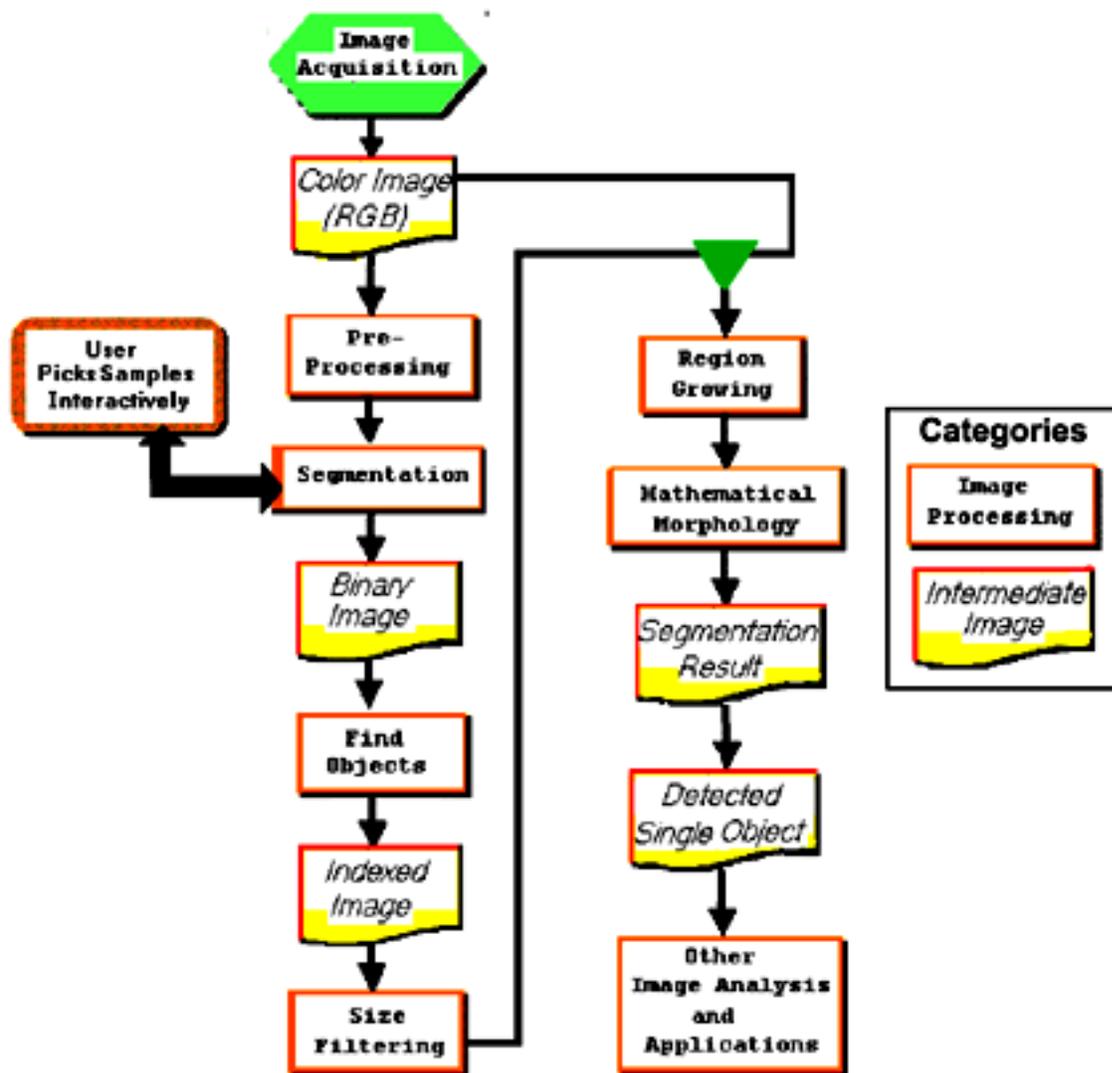

Figure 4. System flow chart to segment color images containing foreground microbial objects of interest.

2.1. Preprocessing

The image acquired by the microscope has to undergo some preprocessing before it is processed further. The quality of the color image might not be good enough because there is usually some noise scattered throughout the image. As the result, the shapes and the boundaries of the microbes may not be clear. When the user interactively samples the colored pixels of the microbe in a later step, it is possible that noise pixels in the same color range are also sampled, creating undesirable results in later steps. To prevent this problem, the user should apply a smoothing filter when the

image is very noisy. There are several methods to smooth the image. One method, commonly used in grayscale images, is to transform the image to the frequency domain and apply a low-median pass filter to remove the high frequency noise. We do not use this method here because some important information in the color image of microbes might be recognized as a high frequency noise and thus undesirably be removed. In the spatial domain, another method is to perform an average operation on the image. Since averaging is a linear operation, the speed can be very fast but image sharpness will be lost simultaneously. The Median filter has the advantage of preserving more details than these other filters that use linear operations, but its drawback is the slower speed of this non-linear operation. In the proposed system, we use a modified version of the median filter to increase its speed.

2.2. Segmentation by similarity

We want to bisect the color image into foreground and background pixels. The foreground should contain the colored pixels representing only the microbes of interest. All the microbes with other colors and all invalid objects should be classified as background. Since there is no fixed color range of a specific color stain, the microbes that should have the same reaction to the stain may show different but similar colors. Because microbes might have exactly the same color as part of the background, we need and must acquire more information than just the RGB values of a single pixel sample and must execute more than a simple threshold to produce the desired, accurately segmented result image.

One important feature of our system is it allows the user to interactively collect the information of the target foreground objects. This proposed system instructs the user to select several sample pixels in the foreground objects within the image. This is a wise design because there is no fixed color range for one stain. It is possible for a stain to produce one color range in one image and a related but still different range in another. Even within the same image, it is possible to have two similar microbes with different colors that belong to the same color group. Moreover, it is common to use different stains to detect different microbes in the same microbial sample. The color of interest can be a mixture of colors derived from different fluorescent materials. Of course we can set a RGB range for the colors of interest from mathematical analysis, but it will not fit very well because even with the same stain, the colors of the pixels that represent the foreground microbes of interest can vary significantly. We also need to consider the case when multiple stains might be applied to the same microbial community sample and be photographed simultaneously under the microscope. For these reasons we designed the user interface to acquire interactively the input color information needed to find the foreground microbes.

We compute the similarity of each pixel in the image compared to the sample points of foreground objects of interest the user picked from the image. Firstly, the system projects the image to a color space and measures the distance between each image pixel and the training sample pixels in the color space. We considered several different color spaces to project the image (e.g., RGB, HSV, CIS-XYZ, and LUV). In an experiment where each of these color spaces was examined, we found that the RGB space provided the best performance in computing the similarity of color between two pixels. In addition to analyzing the distance in color space, our proposed system also makes use of the spatial distance. Images of microbes may cover a wide range of color. The foreground pixels should have similar color when they are close to each other. We have mentioned the problem that portions of the background might have exactly the same color as some part of the foreground. If we consider the similarity only by the color information, once the user picks the colored pixels in the foreground, the system will misclassify the background pixels that have the same color as foreground objects. This is because background objects may also adsorb or absorb the stain non-specifically or they may have other endogenous components that autofluoresce within the same color range as the foreground objects in the image. Human vision can distinguish the foreground from the background easily for most cases because the human mind considers the information in the local regions more than regions at a distance. We can solve the problem of background-foreground mixing using the idea of local thresholding, which is similar to the idea used by human perception. To achieve a similar result to the local threshold, we count the similarity of each image pixel compared to the nearest sample point as a major measurement. Our assumption is that the local contrast can show the difference between foreground and background, and therefore, the foreground in one region of the image should have colors that are more similar to each other.

To compute this similarity, we can compare the current pixel to the sample point that is closest in spatial distance to it, or we can compare the current pixel to the whole set of sample points. If we compute this similarity using the nearest sample point, we may lose many objects of interest in some cases with multiple color groups. This might happened

when a sample point of one color group is closer to the current pixel but the current pixel belongs to another color group. To prevent this problem, we cannot discard the comparison result of the whole sample set. Therefore, we use a distance-weighted similarity in our proposed system. The similarity is measured by the "distance" defined by Eq. 1. Higher similarity results in a smaller measured distance. There are two parts to this equation: an overall comparison and a nearest neighbor comparison. The overall comparison computes the similarity in distance of each single pixel to the whole set of all training sample points. The distance feature is the sum of the color distances between each sample point that is weighted with its spatial distance and divided by the number of sample points. The nearest neighbor comparison compares the current pixel to the sample point that has the shortest spatial distance to it. According to the discussion in the previous paragraph, the nearest sample point should be more important than other sample points because it is in the local region of the current pixel and will provide more information for the target group in the current region.

$S = \{P_1, P_2 \cdots P_n\}$ is the set of sample points selected by user

q : The current pixel visited in the image

P_{near} : The sample point that has the smallest spatial distance to q

$$\text{Distance } (q,S) = W_{total}\left(K_1 \sum_{i=1}^{n} \frac{C_i}{(D_i + 1)^2}\right) + W_{near}\frac{K_2\ C_{near}}{(D_{near} + 1)^2}$$

Equation 1

such that $W_{total} + W_{near} = 1$

n : number of sample points in the sample set

C_i : color distance between current pixel and i - th sample point

C_{near} : color distance between current pixel and P_{near}

D_i : spatial distance to i - th sample point, D_{near} : spatial distance from q to P_{near}

W_{total} : weight for the overall similarity, W_{near} : weight for P_{near}

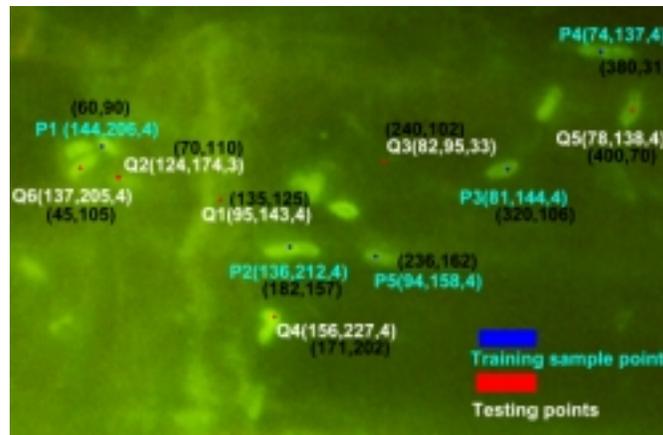K_1 , K_2 : constants for normalization

Fig. 5. Training sample points (blue) and testing points (red) for foreground and background pixels in different regions of this color image. This image is used to illustrate the system's ability to find the foreground objects of interest in a complex background containing pixels of similar color.

The nearest neighbor comparison works very well for most test images that have noisy and uneven color distributions. In the example of Figure 5, the image is partially dark and partially bright, but only the yellow-green color group of

foreground microbes is of interest. However, the nearest neighbor comparison has problems when the user wants to segment more than one color group of interest. When the image includes foreground objects representing several different color groups, the nearest neighbor comparison can belong to only one of those color groups. Thus multiple segmented images must be prepared from this primary image, one for each different color group that the user wants to segment. It is very possible that the current pixel belongs to another color group, which is not the same one as its nearest neighbor. In this case, we need to use the overall comparison in the distance formula.

Current Pixel (x, y)	Color (R, G, B)	Class	Nearest sample	Overall comparison	Nearest neighbor comparison
Q1 (135,125)	(95,143,4)	background	P2	0.0077	0.0248
Q2 (70,110)	(127,174,3)	background	P1	0.0161	0.0757
Q3 (240,102)	(82,95,33)	background	P5	0.0110	0.0195
Q4 (171,202)	(156,227,4)	foreground	P2	0.0068	0.0116
Q5 (400,70)	(78,138,4)	foreground	P4	0.0024	0.0078
Q6(45,105)	(137,205,4)	foreground	P1	0.0030	0.0113

Table 2. Distance for points in different regions in Fig. 5.

Table 2 shows that the nearest neighbor comparison method has good performance to distinguish the background and foreground pixels for a local region, but it does not work well when there is more than one color group of interest in the same image. This example clearly illustrates the high efficiency of the prototype system, even when the input image has varying colors of background pixels that are similar to the foreground. This image is a typical one where the nearest neighbor and overall accuracy measurements reveal almost equal variations in performance. The overall comparison part of the algorithm helps when there are multiple color groups of interest. Depending on how many target groups there might be in the application, we can adjust the importance of the nearest neighbor comparison and overall comparison by adjusting the weight of each term. When W_{total} is zero, we only consider the nearest sample point in the sample set. When W_{near} is zero that means we are using a uniform weight for all sample points. The values for the overall and nearest neighbor comparisons are unitless.

2.3. Find Object
After the segmentation step, a secondary image is produced in which the foreground objects of interest defined by the range of RGB values produced by interactive sampling are presented against a noise-free, monochrome white or black background, depending on the type of microscopy illumination used. However, the size and shape of the microbes might not be complete and their surface contours may not be smooth enough for image analysis directly. Therefore, the image may need further advanced processing so that each foreground microbe of interest is accurately represented. Our proposed system finds the object in the foreground using the 8-connected component method. The result of this step is a labeled image containing only the foreground objects.

2.4. Size-Limit Filter
Not all of the pixels found in the previous step will be the foreground microbes of interest. Some of these invalid objects are just the neighboring non-microbial material that absorbs some of the stain or autofluoresce in the same color range. The area occupied by those invalid objects is usually different from the size range distribution of the valid individual microbes. We can accurately estimate the pixel size range of the microbes in the image separately by image analysis[9] and then filter out those invalid objects whose size lies outside that measured range. The problem is the size range of a microbe changes somewhat with different pixel resolution (more accurate at higher sampling density of pixels in the image). Of course, some microbes are relatively large while others are much smaller, so the size range can also vary up to about 100-fold (0.5 µm to 50 µm), depending on the actual range of cell sizes within the microbial community sample being examined. Several broad-spectrum stains have microbial targets covering a large size range; other stains used as molecular probes are more specific and thus restrict the staining to only certain microbial group(s) within a much narrower size range. In the interface of our prototype system, the user can manually change the minimum and maximum limits of the pixel size of the valid objects. So objects within the same color range but outside the specified size range are eliminated in the processed image.

2.5. Region Growing

We want to find the most appropriate boundary of each foreground object of interest. The difficulty with fluorescent micrographs is that the fluorescent microbes show a bright *halo* encircling the true cell boundaries for reasons discussed earlier. We designed our program to exclude this bright background region of similar color using a region growing method. We can improve the result of region growing by applying a higher threshold to the similarity value computed in the segmentation process. Firstly, the system finds the 'seed' object of each microbe using a high threshold. At this stage, only part of the object will be recognized as an object. With addition of a small step value r, the contour of the object grows outward when its neighboring pixel has a similar color. A similar color refers to the color in the color space with the distance no more than the step value r. This region growing step will be repeated automatically for each foreground object until no neighboring pixels of similar color remain in the image.

Problems might occur in this region growing process. One is that the region commonly grows too much into the background when the object and its background differ slightly in RGB values within a small area of the image. When all regions grow using the same step value r, regions in certain areas of the image might not grow enough whereas other region(s) simultaneously grow too much. It is difficult to decide the optimal value of r for a particular object. An alternative method is to set the upper limit of iterations to a fixed value so that most regions stop growing after that set number of iterations is completed. In the proposed system, we set the system to grow no more than 50 iterations. The user can change the setting depending on the application. This number of iterations is reasonable because the microbes should have a size limit that should not become too big. Making this interactive decision is improved by prior knowledge of the true size range of the microbes under investigation.

2.6. Mathematical Morphology

Sometimes, the apparent brightness of a fluorescently stained microbe in the 2-dimensional projected image is high at the periphery and diminished internally with small darker holes (Fig. 3) that may even look transparent for reasons described earlier, and cannot be filled with the region growing procedure. It is essential to fill those small internal holes within the valid objects in the digital image, especially if measurements based on cell area analysis are to be performed. To fill these internal holes, our proposed system applies a closing operation to each object using the indexed image created after region growing. Here we use a circular disk as the element of dilation and erosion. After picking the sample training points of the foreground objects, the user can adjust the parameters for degree of similarity, the step value of region growing, and the size limit of the microbes. Our proposed system will then capture only the microbes of the color and size of interest and display the segmented result image on the screen.

3. EXPERIMENTAL RESULTS

The goal of the system is to select the regions of the color image that contain the color of interest and eliminate the rest. Examples of pixels that we want to exclude from the resultant segmented image are: invalid objects not of the color of interest, the halos of very similar color around foreground objects, and the background noise with color in the same range as the foreground objects. To evaluate the segmentation results, we want to see if the result can offer the information frequently used by microbiologists. The most common information desired are the count results, size, and shape of the microbes of interest. First, we selected 10 images from different color stains as our test set. Those digital images were acquired directly using a microscope equipped with a CCD camera, or a photographic camera producing images that were digitally scanned. For each digital image, the user picked a set of 10-20 sample training points to represent the colors of interest. The segmentation process was based on these colors of interest. Using the test set as input, the program can generate new images that contain only the foreground objects with the color of interest. The cost of time for the segmentation process depends on the size of the image and the number of sample points. The average time for processing a 600×400 image after sampling the points is only a few seconds. To evaluate the result, we generated the binarized ground truth data for every well-separated cell in each image by manual editing on the original image, as illustrated in Figure 6. We then compared the results generated by the proposed method to the ground truth data to compute the differences between the two results as the percentage error.
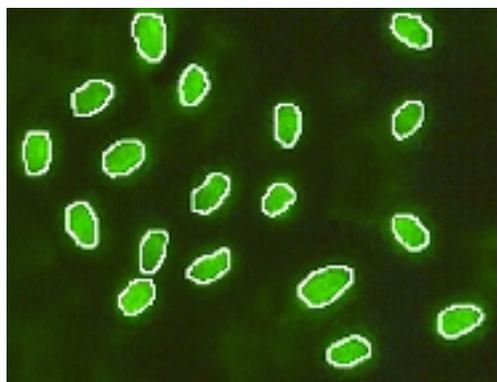
Fig 6. Manually edited ground-truth image of fluorescent bacteria.

To test if the program can segment the objects of interest with their correct size, we compared each pixel in the tested results to the ground truth data. Two different types of error are counted: false alarm (non-object pixels treated as foreground objects) and false dismissal (foreground object pixels treated as non-object background). Those errors are measured in number of pixels (area) and divided by the total pixels of the image and reported as the total percent of error in size. The area we used for this comparison is not just the total number of pixels. We use a pixel-by-pixel comparison to ensure that each pixel is classified correctly. If the segmentation result produced objects of the same size but with a wrong boundary, this measurement of area reflected an incorrect shape or boundary.

Image	Error in size and shape (% of pixels)			Analysis of the results
	False alarm	False dismissal	Total error	
Constellation-02dapi-2	0.54%	0.74%	1.28%	Although these images contain objects of many colors, the boundaries are very clear and easy to recognize.
Constellation-02dapi-6	0.67%	0.12%	0.79%	
Dapi-d	0.12%	1.23%	1.35%	
Dapi-h	0.65%	1.54%	2.19%	
Dapi-i	1.14%	0.56%	1.7%	
FitcAB0403CRA-c	0.12%	2.98%	3.1%	The boundary is not well defined.
FitcABanu843DWC-c	0.25%	1.34%	1.59%	These images have complex background and the errors are due to missing or misclassified objects.
FitcABanu843DWC-e	0.43%	0.34%	0.77%	
RitcFISH-g	0.05%	2.41%	2.46%	Some of the foreground objects are very dark and must be sampled for the system to recognize them.
RitcFISH-h	0.75%	1.93%	2.68%	
Average	0.47%	1.32%	1.79%	

Table 3. Tests of error in size of colored microbes recognized by color segmentation.

From the results in Table 3, we can see that few of the background non-object regions are recognized as a foreground objects. This type of error usually comes from the manual-drawn boundary between touching cells. The false-dismissal error is higher than the false-alarm error. This is because the boundary between the fluorescent halo and the object is drawn subjectively in the ground truth data (Fig. 6). If the boundary in the original image is not very clear, the computer might select a different boundary from what the human picked, but the shape is still similar to the shape in the ground truth data. Both kinds of error in the area representing the foreground objects are low. The sum of the false alarm and false dismissal errors equals the total error, which in this performance test is no more than 2% using our testing images, and therefore this result is very promising.

4. DISCUSSION

There are many factors that affect the results of our analysis and some are more significant than others. Those factors that significantly affect the quality of the results are described here.

4.1. Selection of sample training points

The selection of sample training points is an interactive input that can significantly affect the quality of the experimental results. Since our proposed segmentation method requires no predefined information about the targets, it can only depend on the knowledge provided by the user in the form of the sample training points. The color and the position of the sample training points are the only information describing the desired target group. If the space and color information of the sample training points do not accurately represent the features of the target group, the thresholding result will not be the best that the system can do since the given target is not defined clear enough. In our experiment, we picked 10 to 20 sample points per image. The quality of the system output is only as good as the quality of the sample points that are selected by the user. Therefore, the interactive sampling step must be performed with great care and a sampling of many pixels is recommended if much variability exists in the color composition of the target group.

4.2. Editing the ground truth data

To evaluate the performance of the system, ground truth data need to be produced as objectively as possible (Fig. 6). However, the result of color segmentation cannot be measured directly without human intervention because no absolute ground-truth data are available. In such case, we try to compare the result to manually edited ground-truth data. The procedure of manual editing can be very subjective. Even different microbiologists can define the boundary of the cells in different ways, especially when there are halos around the true boundary of the object. If we use the two different ground-truth data to measure the performance of our system on the same image, the error rates in shape and size may have very different results. The ground-truth data strongly determine the error rate reported in our experiment results. It is difficult to decide where the true boundary lies, even with experienced human eyes.

4.3. Filling holes inside the object

To solve the problem of a dark center inside the object (Fig. 3), we apply a closing operation using mathematical morphology after the region-growing step to fill the hole created by the darker center in fluorescent objects as described earlier. However, the closing operation is limited by size. Since the proposed system does not have any knowledge about the size of the target objects, this method might fail when the object size is much bigger than the mathematical morphology component. To produce the best segmentation result, users have to limit the size range of objects to a certain range, which can be set in the preference options before beginning the color recognition process.

4.4. Noisy background

The non-microbe objects might have the same color as the color stained microbe, as illustrated in Figure 1, showing fluorescently stained bacteria colonized on epidermal cells of a plant root. Color information alone cannot distinguish these background pixels from the microbes. We can apply the size-limit filter and the circularity filter to the selected objects to reduce this background noise in the result image. If this problem still exists, other image editing steps currently not included in our system described here (e.g., a morphotype classification filter) will need to be performed on the image before it can be analyzed.

4.5. Varying the parameters in the algorithm

The nearest neighbor comparison works very well for most test images with noisy and uneven color distributions. Adjusting the weights of the nearest neighbor and the overall accuracy measurements will play a major role in segmenting the microbes. Based on our experiments we found that nearest neighbor comparison gave a good performance in most cases when there is only a single color group of interest. When there is a large variation among the colors in the image, the weight of the overall similarity should be increased for better accuracy. The time taken to calculate the overall similarity mainly depends on the number of sample points chosen by the user. If there is only one group of interest, then we can assign $W_{total} = 0$ and $W_{near} = 1$. By using this approach, the time required to generate the output will be significantly decreased because it need not compute the overall accuracy in this case. Depending on the variation of the color in the image, we can adjust the importance of the nearest neighbor comparison and overall

comparison by adjusting the weight of each term. The adjustment of the weights can also be computed automatically based on the sampled points chosen by the user. However, our experiments revealed that manual adjustment usually yields better results. If these weights are not optimally set, then some foreground objects may be misidentified and omitted.

5. CONCLUSION

This work describes our computer-assisted method to detect microbes of interest in digitized color images. The computer-assisted detection can reduce time and labor costs, hence money and frustration for the whole process. With the weighted similarity measurement, our proposed system provides the flexibility to adapt to different color groups for the segmentation process. Users have to interactively pick a sufficient number of good sample points that can represent the target color group to achieve the best results. The cost of time for doing this will vary depending on the size of the image and the number of sample points picked. It is obvious that this computer-assisted system will be much faster than manual editing. In our experiment, the system performed with about 98% accuracy in object area measurements. The result of color segmentation depends on the quality of the image itself and the sample training points picked by the user. When the background is complex, the segmentation result strongly depends on the color and position of the training sample points picked by the user. It is not necessary to pick points for every region of interest, but the user has to indicate ambiguous parts by choosing the appropriate training sample points. We propose that color segmentation achieved by this new program provides an efficient first stage of image editing, and anticipate that the operating principals of this tool will open many new opportunities for quantitative image analysis of microbes differentiated by color in digital RGB images. However, as with all applications of digital image analysis, color images of the microbial community must be of high quality as a prerequisite. Our first application of this color segmentation program was to produce segmented color images for analysis of the spatial distribution of color-coded bacteria that provide a source of quorum sensing cell communication molecules *in situ* during their colonization of plant roots.

6. ACKNOWLEDGMENTS

7. REFERENCES

1. F. Meyer, "Automatic screening of cytological specimens," Computer Vision, Graphics, Image Processing, vol. 35, no. 3, pp. 356-369, Sept. 1986.
2. S. R. Sternberg, "Biomedical image processing," IEEE Computer, vol. 16, no. 1, pp. 22-34, 1983.
3. Blum, H., "An Associative Machine for Dealing with the Visual Field and some of its Biological Implications," in E.E. Bernard and M.R. Kare (eds.), Biological Prototype and Synthetic Systems, Vol 1, Plenum Press, New York, 1962.
4. C.-S. Fuh, S.-W. Cho, and K. Essig, "Hierarchical Color Image Region Segmentation for Content-Based Image Retrieval System", IEEE Transactions on Image Processing, vol. 9, pp.156-162, 2000.
5. E. Saber, A.M. Tekalp, and G. Bozdagi, Fusion of color and edge information for improved segmentation and edge linking, Image and Vision Computing, vol. 15, no. 10, pp. 769-780, 1997.
6. Fred L. Bookstein, "Shape and the information in medical images: a decade of the morphometric synthesis," Computer Vision and Image Understanding, vol.66 no.2, p.97-118, 1997.
7. G. Agam and I. Dinstein, "Geometric separation of partially overlapping nonrigid objects applied to automatic chromosome classification", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 11, pp. 1212-1222, 1997.
8. H.S. Kim and K.H. Park, Shape decomposition based on perceptual structure, in: L. Shapiro and A. Rosenfeld, eds., Computer Vision and Image Processing, Academic Press, Boston, pp. 363-383, 1992.
9. Jinhui Liu, Frank B. Dazzo, Olga Glagoleva, Bin Yu, and Anil K. Jain, "CMEIAS: A computer-aided system for image analysis of Bacterial Morphotypes in Microbial Communities", Microbial Ecology, Vol. 41, pp. 173-194, 2001.
10. John Ross, Manual of Image Processing Toolkit. 1998.
11. J. Zhou, X. Fang, B.K. Ghosh, "Multiresolution Filtering with Application to Image Segmentation", Mathematical Computer Modeling, Vol. 24, 1996.

12. Knud Thomsen, *Color Theory: Introduction to the theory of color*, Dimac. 1997.
13. Longin Jan Latecki, Rolf Lakämper, "Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution", Computer Vision and Image Understanding, Vol. 73, No. 3, pp. 441-454, 1999.
14. M.P. Dubuisson, A.K. Jain and M.K. Jain, "Segmentation and Classification of Bacterial Culture Images", Journal of Microbiological Methods, Vol. 19, pp. 279-295, 1994.
15. Miguel A. G. Ballester, Andrew Zisserman, Michael Brady, "Segmentation and measurement of brain structures in MRI including confidence bounds", Medical Image Analysis, June 1999.
16. M. H. Wilkinson and F. Schut, "*Digital Image Analysis of Microbes: Imaging,* Morphometry, Fluorometry, & Motility Techniques & Applications", John Wiley & Sons, UK 1998.
17. Michael E. Sieracki, Stephen E. Reichenbach, and Kenneth L. Webb, "Evaluation of Automated Threshold Selection Methods for Accurately Sizing Microscopic Fluorescent Cells by Image Analysis,", Applied and Environmental Microbiology, pp.2762-2772, 1989.
18. Prewitt, J., "Object Enhancement and Extraction," Picture Processing and Psychopictorics, Academic Press, pp. 75-149, New York, 1970.
19. Linda Shapiro and George Stockman, *Computer Vision*, Addison Wesley, 1998.