Development and Implementation of an Efficient Automated Cell Colony and Plaque Counter

Kevin Claytor
Rice University Physics Department

Abstract

The manual counting of cell colonies, or viral plaques is a laborious, inaccurate task. The work here investigates a low-cost alternative to professional colony / plaque counters. We used an off the shelf webcam, and image processing software (MATLAB) in conjunction with well studied image processing techniques, primarily the watershed transform and the regional maxima transform. With these methods we were able to develop a hardware, software combination that was capable of counting to within 93% of manual counting, in under 10 s for a 90 mm agar plate, and with only one parameter that the user needed to modify. We suggest that this is an effective solution for labs with medium throughput colony or plaque counting.

I. Introduction

In the experiment presented here we use automated cell counting to establish the infection behavior of λLZ1 and λ2903 in *E.coli* LE392. This was done as preliminary work to determine if the modified λ phage, λLZ1, would infect the host in a manner similar to the wild type, λ2903. The phage was modified with about 400 YFP molecules [1] to make it visible under the microscope with the goal of providing single molecule statistics and behavior for lambda infection.

It was decided to develop an automated colony counter for this experiment for three major reasons. Foremost is the speed of automated colony counting, each plate only requires a few seconds to count by computer, much faster than the minutes it may require for a human to count a plate with thousands of colonies [2]. Secondly, the automated counting is more reproducible; the computer uses the same set of parameters for making the decision of classifying a colony, a human operator's judgment may vary within a sample, and moreover the classification between operators differs, resulting in counts that vary by 12% or more [3]. Finally the automated analysis can be extended to other experiments with only minor software changes, allowing the system to have a great versatility.

Colony counting is an important step in many lab procedures, and automated counting has been proposed and investigated multiple times before, likewise it is a lucrative market with several companies manufacturing them today. As early as in 1973 an off the market colony counter was tested, found to agree to 89 to 95% of a hand count, and was recommended as useful [4]. Since then numerous methods have been applied to attempt fast and accurate recognition of uniform bacterial colonies to interestingly shaped mammalian cell colonies. Usually one uses either a scanner [5, 6], or a camera with or without a backlight [3, 7-10] to capture images. A number of novel techniques have been applied ranging from using a distance transform [9], a Hough transform [8], parameter identification and fuzzy logic [6], and optimizing a Gaussian model of the image [7]. However, the counting range of most of these methods was not verified in cases

exceeding 400 colonies / plate. Furthermore, most of the algorithms, while effective, do not see wide circulation. More recently, however, software has been written and released, free of charge, that allows for multiple types of image analysis, including colony counting [11].

Professional colony counters meet many of the challenges to automated colony counting. They are able to obtain colony numbers up to 95% of a hand count [4]. They also do so with much less variation in one sample than human counters [12, 13], and finally do so in seconds [5]. Likewise, their drawback is several fold; firstly, the instrument cost is considerably higher than the comparable imaging systems (scanner, camera) mentioned above [14]. Hence this is prohibitive to their wide use, and results in a frequent queue for a device [13]. Additionally, the software is proprietary and is not able to be modified by the end user for custom experiments. Furthermore the software must be configured for the parameters of each sample before beginning counting. Many of the algorithms mentioned in the previous paragraph do not suffer from this problem. Although professional colony counters see great use in many settings, they are not idea for all, and have some major drawback that prevents them from seeing prolific use.

To complement the advantages that both the academic and professional colony counters see, and to develop a low cost, versatile system that can be put in the hands of the average researcher is the primary goal of this research. To this extent, we saw fit to design a system that was streamlined, able to count colonies without much, if any, human input. Primarily, however, the system needs to be accurate; ideally to 5% of a trained hand count, and over a wide range from 0 to 2000 or more colonies, even for a non-dyed sample. Also to be competitive against the professional systems, it would need to be fast, between 1 and 10 seconds [5]. Furthermore, the adaptability of the system would have to be maintained, the software would have to be easy to modify. With these constraints it becomes economical, in terms of money, accuracy, and time, to completely automate the laborious task of counting colonies or plaques.

II. Method

A. Colony, plaque, and florescence preparation

Bacterial plates were prepared using the following procedure. *E.coli* strain LE392 was allowed to grow overnight. This solution was diluted 100 fold and grown again to mid-log phase, about 10^8 cells / mL. This was concentrated 10 fold into LBMM solution via centrifugation. Two separate lambda phages were used, λLZ1 and λ2903 were use to infect the *E.coli*. First the phage stock was diluted and baked for 30 minutes at 30° C to evaporate chloroform from the solution. The infection was then preformed by adding phage to the *E.coli* stock and letting it sit at room temperature for 20 minutes. Following this the infection was allowed to incubate at 30° C for 45 minutes, and plated onto Kanamycin 10 μg / mL. The plates were then allowed to incubate at 30° C overnight.

B. Apparatus

To control lighting conditions a custom sample holder was constructed (see Figure 1). The configuration was designed to eliminate reflections from overhead lights and objects that would otherwise give erroneous results. Backlighting the sample also provided adequate contrast to resolve viral plaques. A 2 Megapixel Logitech QuickCam Pro [15] was used to acquire the images. It was placed 12 cm from the dish, and the built-in auto focusing routines were allowed to focus the images. The sample sat on a

plastic diffuser with patterned paper below that. This equalized the intensity from the light source, such that the sample could always be lit uniformly. For the colony and plaque counting, the light source was a 15 watt CFL light bulb.

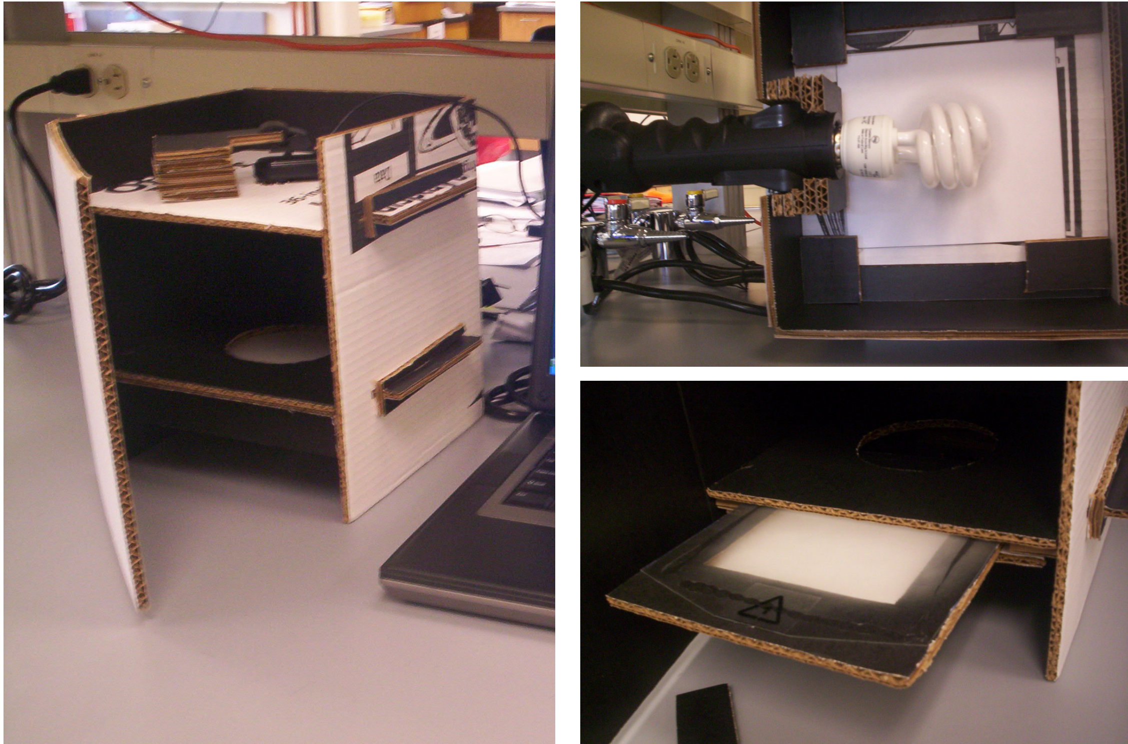

Figure 1: The experimental apparatus. Left: The box that holds the camera above the sample, also visible is the depression in which the dish sits. Upper Right: The underside of the device showing the light source for colony / plaque counting, as well as the Gaussian patterned paper diffuser. For the florescence measurements, the apparatus was placed upon the UV excitation table and blank paper was used. Lower Right: Illustrating how the diffuser could be removed for modification.

C. Analysis routines

The analysis routines were written in MATLAB using the Image Acquisition and Processing toolboxes. The code is freely available online for use and modification (see). Two main analysis method were used, the watershed transform and locating regional maxima. The code pipeline for these methods is shown in Figure 2.

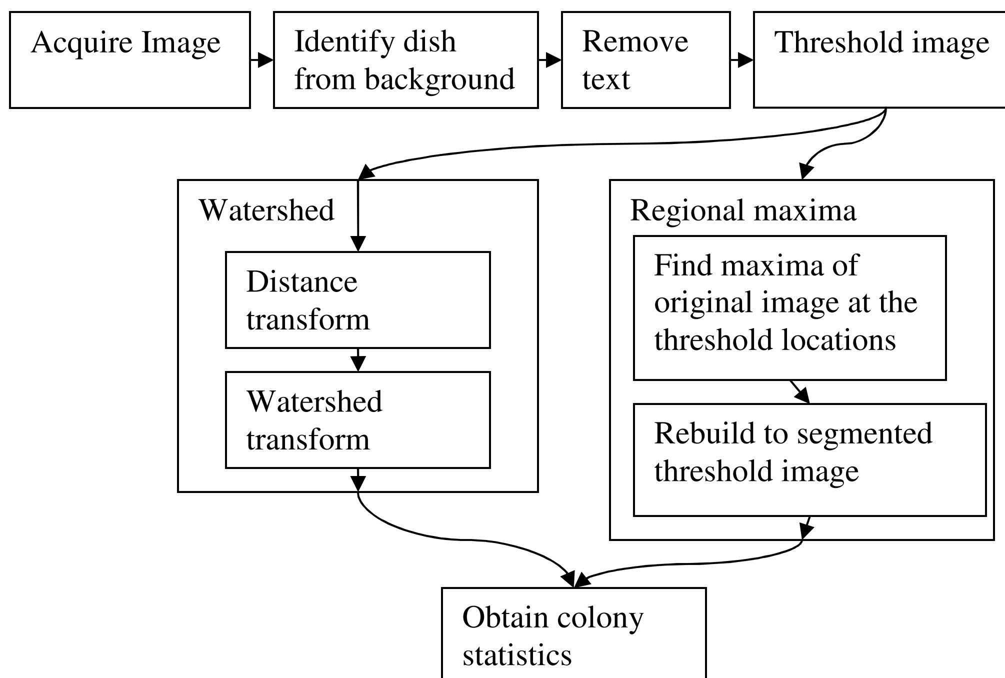

Figure 2: The analysis pipeline.

The first difficulty in detecting the colonies was developing an accurate method to threshold the images, and a metric which could alert the user to when the threshold was inadequate. To further equalize any intensity fluctuations from lighting, as well as to increase the contrast between circular colonies (plaques) and the agar, a morphological top-hat transform was used. An intensity threshold was then applied to the image, dividing it into of colonies or plaques, and background agar. A morphological close was performed with a small structuring element to remove small, non-colony artifacts. The resulting binary image, or mask, was then processed by either the watershed or regional maxima methods.

The watershed method first took the distance transform of the binary mask. The watershed transform was then done directly on this image. Alternatively, the original image value at the locations of the colony mask were taken. The regional maxima transform was then done to this subset of the data to separate merged colonies. This resulted in individual points corresponding to maxima, which were morphologically extended to fill the full binary mask. These steps, as applied to the data, are shown in Figure 3.

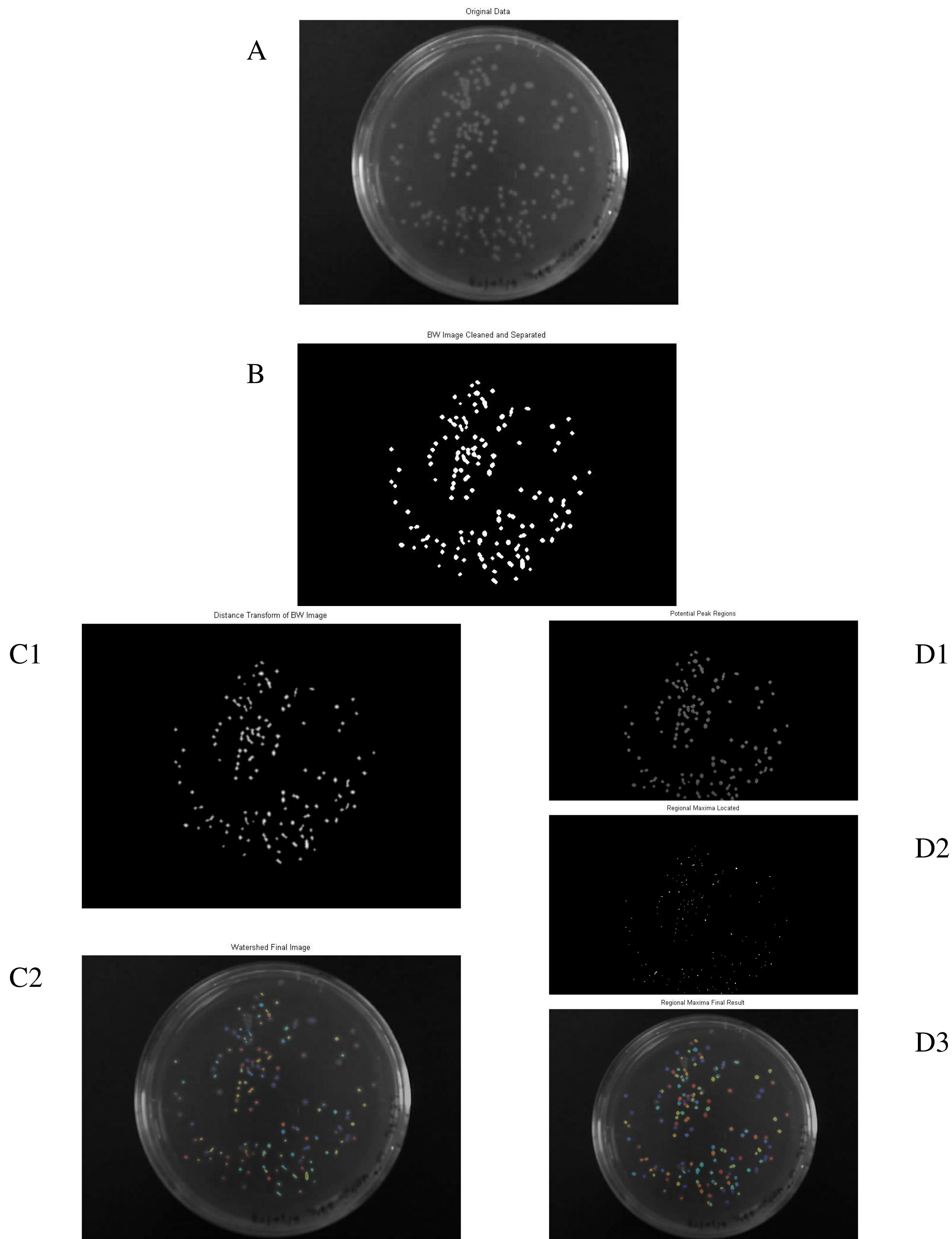

Figure 3: Captures of various stages of process in the image analysis. A) The original image. B) The binary mask after the initial threshold and text removal. C) The watershed pathway. C1) The distance transform of the binary mask. C2) The watershed transform of the binary mask in a false color overlay on the original image. D1) The values of the original image at the mask locations. D2) The result of the regional max transform (individual points). D3) The false color overlay of the recovered mask area.

The different image conditions, colony, plaque, or florescence, required minor modifications to this overall procedure. Since the sample was back lit, colonies would appear dark, thus the colony images were inverted before any further analysis was performed. No special treatment was done for the plaque images.

Occasionally text would be written on the plates. While this proved useful for focusing the camera, it contributed to erroneous counts. To remove the text in colony images, the maximum value in the central part of the dish was determined. The locations of any values in the image over this were then removed from the binary mask. Since the text was far darker than any colony, it showed much brighter than any of the colonies in the inverted colony image and was reliably removed. For plaques a similar method was used that removed the darkest sections within the outer annulus of the dish.

III. Results and Discussion

Identifying the proper threshold to use proved to be the most challenging aspect of designing software that would reliably, and without further operator input, count a sample. Since no prior knowledge of the sample was assumed, this added another level of complication. Although the built in threshold method was used, a number of additional features were implemented to make it more robust. First the threshold was taken of only the central region of the image after variations in background were removed. This decreased the sensitivity of edge effects as well as uneven background effects. Finally, the threshold was performed for the image with the text removed. Since the text was of much higher or lower intensity, this made the threshold more sensitive to the difference between the agar and the colonies / plaques. Figure 4 shows the sensitivity to the threshold. As the level was varied, there is a region for which the count is accurate. The automatic threshold selects a value that is usually in this region.

The method for determining the threshold was not perfect, however. On plates with a low density of very small colonies or plaques, the threshold was much less than the optimum. No method was found to circumvent this problem, instead a metric the threshold divided by the average intensity of the plate was used. For these low density plates correct counting would place the threshold much above the average intensity. This allowed an inaccurate threshold value to be identified and re-run with a higher threshold.

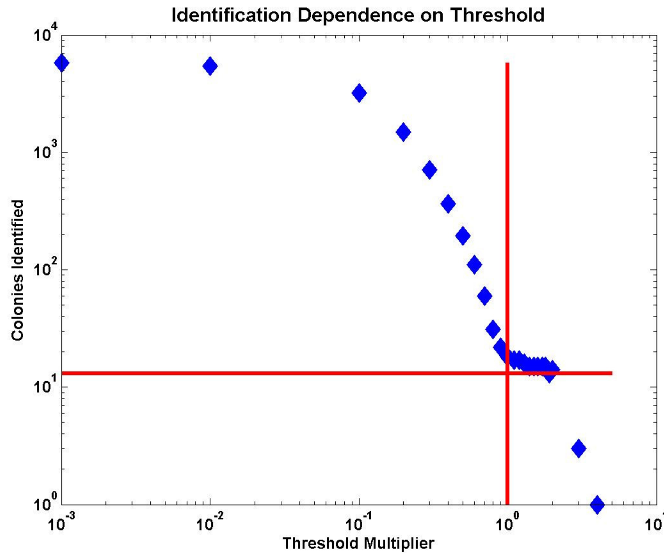

Figure 4: Results from varying the threshold around the proper value. The points are the number of colonies counted. The vertical line is the automatically identified threshold value. The horizontal line marks the number of colonies counted by hand. The automatically generated threshold value is on the verge of a region of reliably accurate counting.

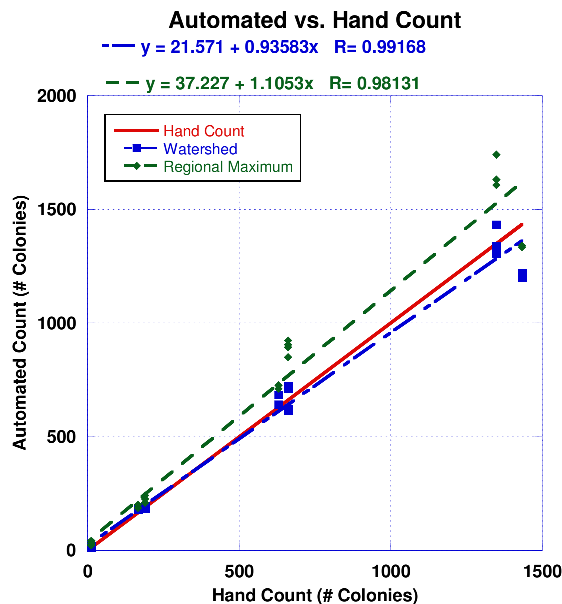

Figure 5: The values obtained for the two methods of automated counting plotted against the hand count. The solid line is a slope of unity; what would be observed if the automated count were to exactly mimic the hand count.

Data were taken for 8 plates. Each plate was imaged at four different rotations. The colonies on each plate were subsequently counted by hand. Figure 5 shows the automated count plotted against the hand count, and corresponding linear fit, for both the watershed and regional maximum methods. In this figure, a slope closer to unity

indicates an automated count that obtains a value closer to the human count. The y-intercept indicates the extrapolation to how many colonies would be counted on an empty dish. Hence an intercept closer to zero is more desirable.

Figure 6 presents an alternative presentation of the data. This plots the difference between automated count and hand count. As such over counting yields a positive value, and undercounting a negative one. The unfocused images have been removed from this analysis. The error bars represent the standard deviation of the remaining measurements. Analysis times were 8.7 ± 0.4 s for the watershed method and 21.2 ± 4.6 s for the regional maxima method.

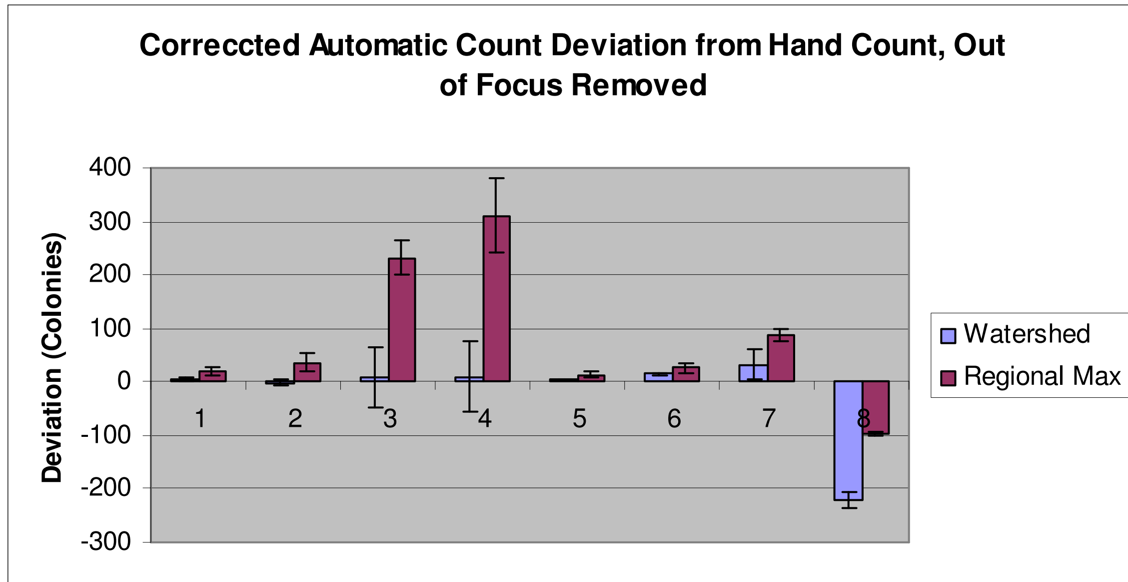

Figure 6: An alternative method of displaying the data in Figure 4. Here the difference between the hand count and the average automated count is shown. Out of focus images were removed from the analysis and the error bars indicate the standard deviation in counts of the remaining images. Plates one through four consisted of colonies roughly 1 mm in diameter, with increasing density. Plates five through eight were of colonies roughly 2 mm in diameter, with increasing density.

One question that arises is why the count for sample 8 was significantly lower than expected. These data were obtained before the printed equalizer was implemented, as such the center of the field was saturated and colonies could not be identified in the central region. While this would affect all plates, the high density plates would be affected to a greater extent as there would be more missed colonies in the saturated, uncounted region. Although a light 3 pixel Gaussian filter was used to smooth out short distance noise, it was found that the out of focus images would have wildly inaccurate counts. In all out of focus images, it was found that the text removal failed and the text would subsequently be counted as colonies. Competing with this factor was a drastic smoothing of the colony features; intensity, shape, and intensity distribution, resulting in an undercounting of colonies. If the former had a greater effect, the count would be higher than the in focus image, if the latter was more significant the count would be

considerably lower than the in focus image. Interestingly, one of these methods would nearly always dominate, rather than interact to produce a seemingly normal count.

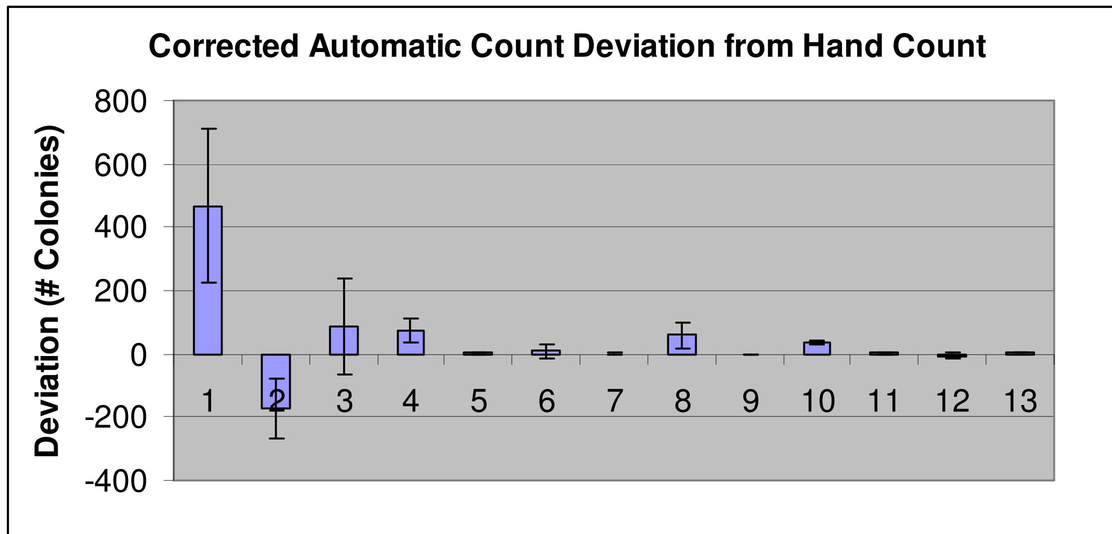

Figure 7: The data for the m.o.i. experiment shown in Figure 8. The average colony size was < 1mm in diameter. Samples 1 through 5 are the 2903 strain and samples 6 through 13 are the florescent LZ1 strain. Due to the small colony size, a higher threshold was required for most cases. Samples 5, 9, 11, 12, 13 required multiplying the threshold by 10, sample 7 required twice the automatic threshold, and sample 10 required twice the automatic threshold.

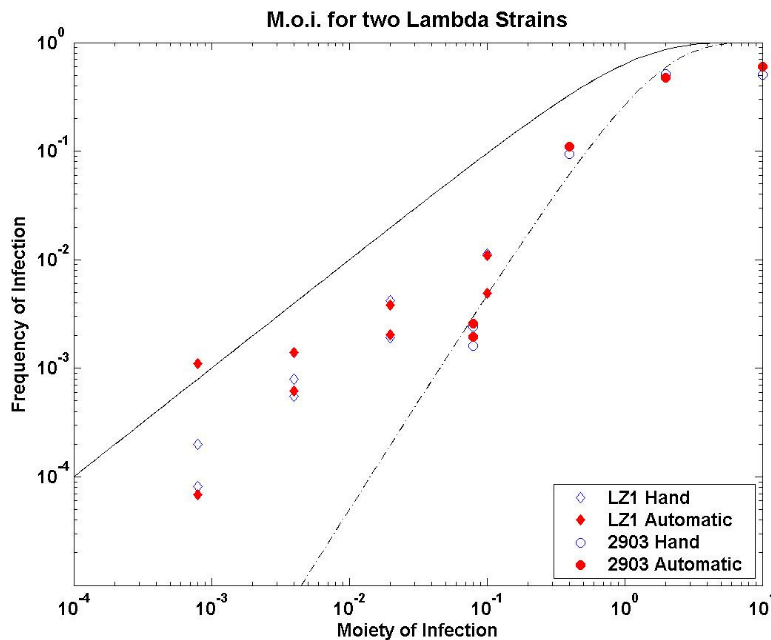

Figure 8: Frequency of lysogenization for two strains of viruses. Using both the hand and automated counting methods. LZ1 is a florescent strain modified from [1], while 2903 is

a wild type strain. The solid line is the theoretical fit for lysogenization induced by one or more viruses, and the dashed line is for lysogenization induced by two or more viruses.

The automatic counter was used to perform a moiety of infection experiment. Figure 7 shows the deviation of the automatic count from the hand count. Although the threshold needed to be modified in several cases, especially with the smaller LZ1 colonies, only three multipliers were used to obtain reasonably accurate results. Figure 8 shows the results of the m.o.i. experiment for both the hand and automatic counts. The 2903 wild type strain follows the expected curve for lysogenization by two or more viruses [16]. The florescent LZ1 strain, however, does not follow the same curve. It has the slope of the expected curve for lysogeny induced by one or more viruses, however, the data points lie below the expected curve. This is most likely due to having fewer viruses than expected, or an expected m.o.i. greater than that used. This is likely as the phage stock was several weeks old and may have deteriorated in health in that time. Of interest here is that Figure 8 indicates that the florescent LZ1 strain may be damaged in its lysogeny behavior.

IV. Conclusions

This research has shown that modern low-cost webcams paired with simple generic algorithms can perform automated colony and plaque counting as effectively as professional colony counters. The accuracy of 93% with the watershed method is very near the goal of within 95% accuracy, and functions in under 10 s. The user variable parameters were kept to a minimum, requiring only a threshold adjustment, and the program is still able to count a wide range of colonies. The effectiveness was validated with a m.o.i. experiment, where the hand and automatic count both show the same results, namely that the florescent strain used here has altered lysogeny behavior. With additional minor modifications to the code and setup, a number of different experiments could be performed, maintaining the system's flexibility. With the development and validation of the system presented in this paper, there remain no deterrents for the modern biological laboratory not to have an automated colony counter.

V. Acknowledgements

I would like to acknowledge the National Science Foundation funding this research under grant number; PHY-0243675.

VI. References

The MATLAB code used to develop this project, as well as the schematics of the camera system, are all copyright under the Creative Commons license (Attribution, Non-Commercial Share Alike (http://creativecommons.org/licenses/by-nc-sa/3.0/)), and may be found online at the following URL: http://owlnet.rice.edu/~kec4482/

1. Alvarez, L.J., et al., *Propagation of fluorescent viruses in growing plaques.* Biotechnol Bioeng, 2007. **96**(3): p. 615-21.
2. Putman, M., R. Burton, and M.H. Nahm, *Simplified method to automatically count bacterial colony forming unit.* J Immunol Methods, 2005. **302**(1-2): p. 99-102.

3. Bollea, A., et al., *Automatic Quantitation of Cell Colonies on Petri Dishes by Computerized Image-Analysis.* Cytotechnology, 1995. **17**(3): p. 185-191.
4. Goss, W.A., R.N. Michaud, and M.B. McGrath, *Evaluation of an automated colony counter.* Appl Microbiol, 1974. **27**(1): p. 264-7.
5. Dahle, J., et al., *Automated counting of mammalian cell colonies by means of a flat bed scanner and image processing.* Cytometry A, 2004. **60**(2): p. 182-8.
6. Marotz, J., C. Lubbert, and W. Eisenbeiss, *Effective object recognition for automated counting of colonies in Petri dishes (automated colony counting).* Comput Methods Programs Biomed, 2001. **66**(2-3): p. 183-98.
7. Bernard, R., M. Kanduser, and F. Pernus, *Model-based automated detection of mammalian cell colonies.* Phys Med Biol, 2001. **46**(11): p. 3061-72.
8. Barber, P.R., et al., *Automated counting of mammalian cell colonies.* Phys Med Biol, 2001. **46**(1): p. 63-76.
9. Mukherjee, D.P., et al., *Bacterial colony counting using distance transform.* Int J Biomed Comput, 1995. **38**(2): p. 131-40.
10. Yourassowsky, E. and E. Schoutens, *Automated count and size evaluation of colonies of bacteria grown in a zonal concentration gradient of antimicrobial agent.* Appl Microbiol, 1974. **28**(4): p. 525-9.
11. Lamprecht, M.R., D.M. Sabatini, and A.E. Carpenter, *CellProfiler: free, versatile software for automated biological image analysis.* Biotechniques, 2007. **42**(1): p. 71-5.
12. Hattier, T., *Variance Rotations*, K. Claytor, Editor. 2008.
13. Herman, C.J., et al., *In-use evaluation of the Omnicon automated tumor colony counter.* Cytometry, 1983. **3**(6): p. 439-42.
14. Hattier, T., K. Claytor, Editor. 2008, Cellular Technology Limited. p. 2.
15. www.logitech.com.
16. Kourilsky, P., *Lysogenization by bacteriophage lambda. I. Multiple infection and the lysogenic response.* Mol Gen Genet, 1973. **122**(2): p. 183-95.