

LEGINON: A SYSTEM FOR FULLY AUTOMATED ACQUISITION OF 1000 ELECTRON MICROGRAPHS A DAY

C.S. Potter, H. Chu^{*}, B. Frey, C. Green, N. Kisseberth, T.J. Madden, K. L. Miller, K. Nahrstedt^{*}, J. Pulokas, A. Reilein, D. Tcheng^{**}, D. Weber and B. Carragher

Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign
405 N. Mathews, Urbana, IL 61801

^{*} Dept. of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801

^{**} National Center for Supercomputing Applications, Urbana, IL 61801

Abstract

We have developed a system to automatically acquire large numbers of acceptable quality images from specimens of negatively stained catalase, a biological protein which forms crystals. In this paper we will describe the details of the system architecture and analyze the performance of the system as compared to a human operator. The ultimate goal of the system is to automate the process of acquiring cryo-electron micrographs.

PACS: 07.78; 07.05.P; 61.16.B; 87.64.Dz

Keywords: Microscopic methods, instrument control, automated data acquisition, three-dimensional electron microscopy

1. INTRODUCTION

Molecular microscopy is, and will continue to be, one of the most important structural approaches in cell biological investigations. Currently, the technique requires the acquisition of very large numbers of high quality images from an electron microscope controlled by an experienced microscopist. This is a labor-intensive and slow methodology and it is clear that this situation must change if important biological problems are to be addressed in an expeditious manner. There is increasing interest in the field for fully automating the entire process of acquiring high quality transmission electron micrographs.

Typically, a microscopist identifies potential features of interest by visual inspection of a low magnification field of view. High magnification images of these identified features are then acquired using techniques which minimize the exposure of the specimen to electron beam damage. As a result, the high magnification image is never visually examined prior to acquisition. The quality of the high magnification image is assessed only after acquisition when the image can be analyzed and a decision made as to whether it warrants further processing. An experienced microscopist assimilates this quality assessment information and uses it to refine the choice of potentially relevant low magnification features. A simple brute force method in which the entire low magnification field of view is systematically examined is impractical because the field of view

is very large and the scale change between the low and high magnification images is typically two orders of magnitude.

It should also be noted that the term automation is quite often used where "computer assisted" or "semi-automated" might be more appropriate (see for example [2]). While there are a number of excellent systems [1-3] which relieve the operator of many of the tedious tasks at the console, the operator needs to identify the regions of interest and provide input to the system during the data acquisition process. Our goal is to develop a system which can acquire several thousand images in a day with no input required from a human operator.

We have developed a system, called Legion [4-5] to automatically acquire large numbers of acceptable quality images from specimens of negatively stained catalase, a biological protein which forms crystals. Acquiring good quality images of this specimen is often used as a test for students taking a course in electron microscopy and thus provides an excellent driver for the research methods that must be developed to solve the general problems of automated image acquisition. Furthermore, as catalase is an ordered crystalline structure, assessment of this order provides us with an objective measure of the quality of the automatically acquired images. In this paper we will describe the details of the system architecture which we have developed and analyze the performance of the system for the automated acquisition of a thousand images a day.

2. METHODS

A general outline of the steps required for the acquisition of high magnification images is as follows:

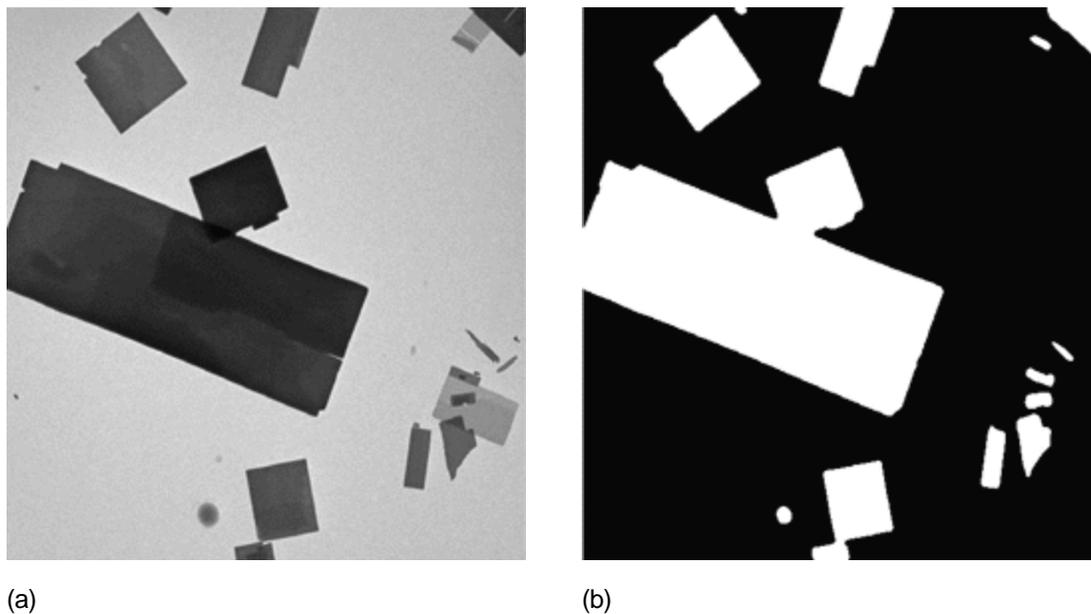
- Step 1.) Acquire low magnification (660x) image.
- Step 2.) Identify potential features of interest.
- Step 3.) Locate feature in center of field of view.
- Step 4.) Autofocus at 38,000x.
- Step 5.) Acquire high magnification (38,000x) image.
- Step 6.) Assess quality of high magnification image.

High magnification images are acquired for every feature of interest identified in the low magnification image and the entire procedure is repeated for every square on the grid. Each of these steps will be described in more detail below.

2.1 Specimen Preparation

Catalase crystals were prepared based on the method of Sumners and Dounce [6]. A low magnification image [660x] of the catalase preparation is shown as [fig. 1a](#). The preparation results in large rectangular areas of crystalline protein. However, the number of layers of protein is quite variable and the individual crystals often overlap. Various stain and contamination artifacts are also common.

Figure 1.



2.2 Microscopy and Image Acquisition

Images were recorded using a Philips CM200 equipped with a Gatan MSC CCD. The cold trap liquid nitrogen dewar was modified to increase the cryogen capacity to enable the system to run overnight without refilling.

2.3 Low Magnification Image Acquisition

A low magnification image of each grid square is automatically acquired by systematically moving the goniometer to the center of each grid square in a spiral sequence. The size and orientation of the grid squares is determined from an initial calibration which takes only a few minutes to complete. This procedure can also be used as a stand alone application to automatically acquire low magnification preview images of an entire grid [7].

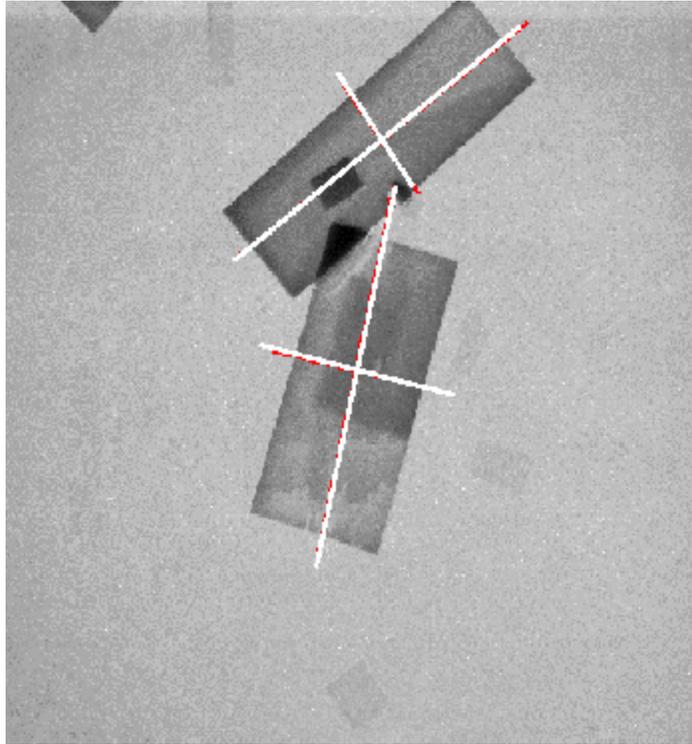
2.4 Low Magnification Feature Identification

Each low magnification image is processed to identify large contiguous areas of density by a method which uses cross correlation of the image with a template (fig 1b). Image feature metrics (size, mean, variance, centroid) are calculated and stored for each of the identified contiguous regions. These image features are later used in deciding whether a high magnification image of the region will be acquired; for example, regions that are too small are rejected. Within each acceptable contiguous region, a variety of methods can be used to select a target at which to acquire a high magnification image. The simplest method is to use the center of the region as a targeting selection and this works well for non-overlapping crystals.

Quite frequently, however, several crystals overlap and an additional analysis is used to model these clusters in order to identify the individual crystals. A mixture of orientation-adaptive, scale-adaptive masks are fit to the subimage using the expectation-maximization algorithm [8]. After testing out a number of masks, the algorithm selects the one that gives the best fit. Each mask identifies a specimen or a significant portion of a specimen and the centroid of each mask is taken as a possible target location for obtaining a high-magnification image. Fig. 2 illustrates how this pattern recognition algorithm can identify two catalase crystals, despite the fact that they

are overlapping. The orientation and size of each mask is then used to extract a normalized rotation-invariant, scale-invariant image of the corresponding specimen. These image features are used as a targeting selection criteria for subsequent high magnification image acquisition. For example, in [fig. 2](#) a high magnification image is acquired for each of the two target points indicated.

Figure 2.

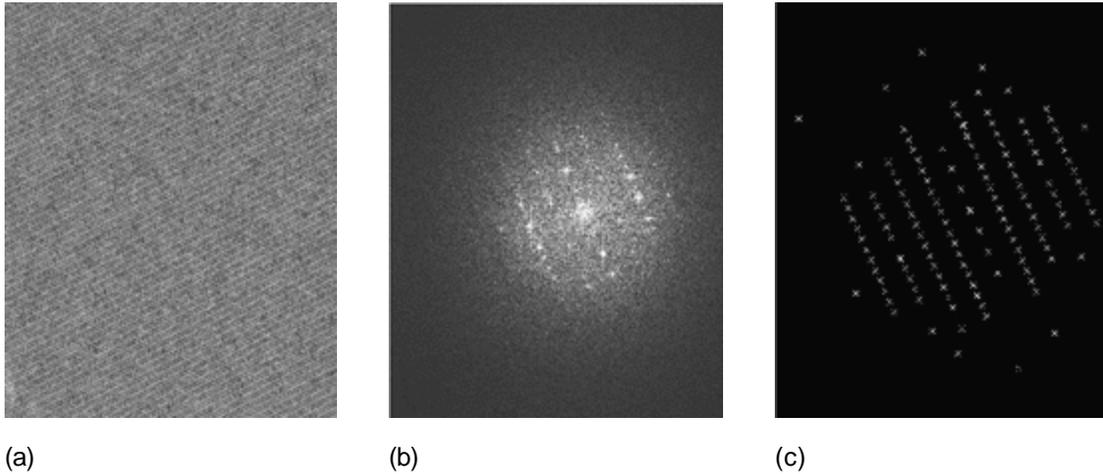


One final method used for target selection was a computer assisted interface where the operator manually selected targets from the low magnification image. This was developed to compare the performance of the automated system with that of a human operator. These results are described below.

2.5 High Magnification Image Acquisition

A high magnification image is acquired at each of the targeting locations identified in the low magnification analysis. The first step is to center the selected feature within the field of view and this is achieved by repositioning of the goniometer. The goniometer has been calibrated and programmed to correct for backlash errors but its precision is still such that feature location can be inaccurate by up to 500nm (approximately within the specifications of the microscope). To compensate for these inaccuracies we have implemented a second refinement step in which correct positioning is verified by cross correlation with the targeted sub-region selected from the low magnification image. The magnification is then increased to 38,000x, the image is focused automatically using a beam tilt algorithm [9], and a high magnification image is acquired ([fig. 3a](#)).

Figure 3.



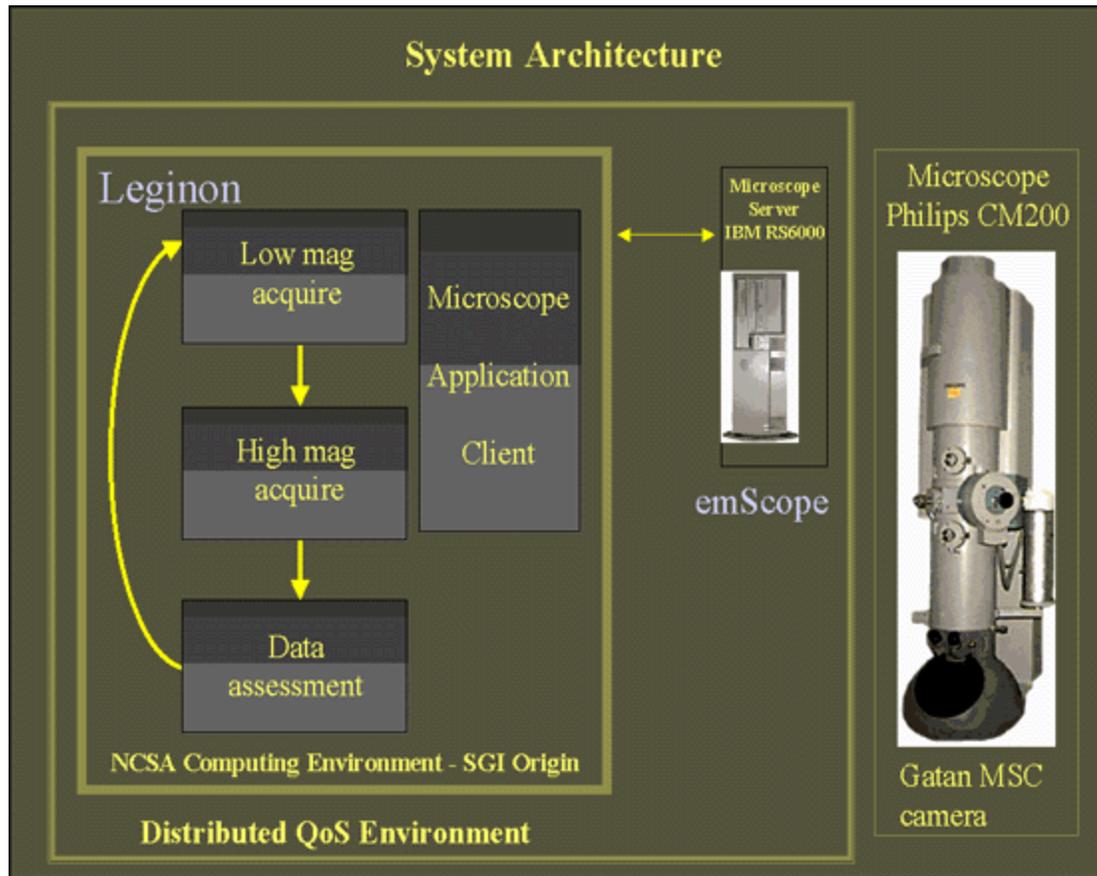
2.6 High Magnification Image Quality Assessment

The image quality of each high magnification image is automatically assessed by calculating the power spectrum (fig. 3b), identifying diffraction spots (fig. 3c), and measuring the signal to noise ratio of each diffraction spot [10]. The machine quality rating for the image is equal to the number of diffraction spots with a signal to noise greater than 3.5 (grade 1 and 2 spots).

2.7 Legion System Architecture

The Legion system uses a real-time distributed computing architecture as shown in fig. 4 and it consists of the following layers:

Figure 4.



EmScope Control Server: A multi-layered software library for portable and extensible control and acquisition in a distributed software environment [7]. A server has been developed using this library that provides distributed control of all functions on the microscope and camera.

Application Layer: A client application controls the entire data acquisition and analysis process. The application layer has been developed using Tcl/Tk [11] and integrates readily available image processing applications [12-13] as well as custom developed software. It integrates microscope and camera control, image analysis and processing, data management and the application logic. The application can be readily modified at the tcl scripting level.

User Interface: The application is monitored using a standard web browser that can be accessed by multiple users. The interface can display any of the steps in the acquisition process and is useful for assessing the progress of the experiment. Since data acquisition can run continuously for several days a web browser interface allows a means of assessing and monitoring the experiment from any convenient location.

Quality of Service Guarantees: A real-time distributed computing environment has been developed to (i) provide processor quality of service (QoS) guarantees and (ii) allow this application to share the computational resources with other non-real-time applications. The shared resource, CPU, is controlled by an adaptive real-time server (ART Server) [14]. The application accesses the ART server through the Tcl/Tk application layer as described above. We have implemented the ART server on an SGI O200 system that has two CPU processors running IRIX v 6.4.

3. RESULTS

The performance of the Leginson system was analyzed in terms of efficiency and accuracy as compared to the performance of a human operator.

3.1 Efficiency of Image Acquisition

Currently, the automated system can acquire on average one high magnification image every 90 seconds; which translates to approximately 1000 images in a 24 hour period. This compares with an average time of 110s per image for a human operator using the computer assisted interface as described above. The principle bottlenecks are image acquisition time, network bandwidth for transferring images, and processing performance (identification of features of interest in the low magnification images and assessment of the quality of high magnification images). With planned improvements in the network and the efficiency of the software we are confident that the automated acquisition time can be reduced to less than 45s an image.

3.2 Accuracy and Precision of Microscope Controls

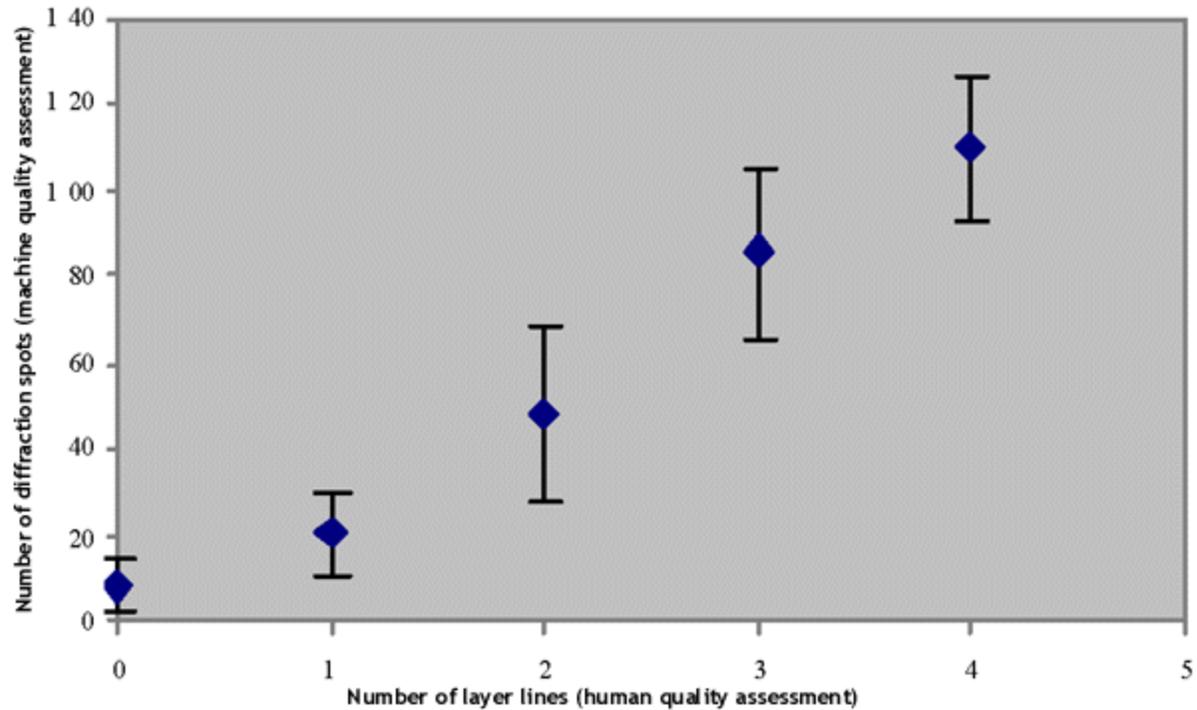
For automated acquisition it is necessary to (i) accurately control the goniometer in order to locate the features of interest in the center of the field of view of the low magnification images; (ii) maintain the feature of interest in the center of the field of view when the magnification is increased by several orders of magnitude; and (iii) accurately focus the image at high magnification. These steps require precise characterization of the response of the microscope goniometer, image shift and beam tilt coils. The accuracy with which the goniometer can be moved to a precise location is approximately 500nm and this does not allow for centering the feature of interest in a single step. Ideally, we would like to move with an accuracy of a single pixel on the low magnification image which would require accuracy to 30nm. We currently work around this problem by implementing a multi-step method in which the current location is refined by cross correlation to the target location.

Maintaining the feature of interest in the center of the field of view as the magnification changes also posed difficulties due to hysteresis effects at the very low magnification settings (660x). This occasionally causes the center of the feature of interest to shift as much as several hundred nm when returning to the low magnification setting. Again, we have implemented a refinement step to compensate for these effects.

3.3 Assessing the quality of the images acquired

The overall system performance was assessed by calculating the percentage of acceptable quality images acquired by the system during a session and comparing this performance with that achieved by a microscopist. As described above we can automatically assess the quality of our images by counting diffraction spots in the power spectrum. In order to evaluate the reliability of this method we compared the machine quality assessment to the quality assessment as made by a human observer. The human observer simply examined each power spectrum and counted the number of layer lines which could be observed (see [fig. 3b](#)). For each image, the machine quality assessment (the total number of diffraction spots with SNR >3.5) was then plotted against the number of observed layer lines ([fig 5](#)).

Figure 5.

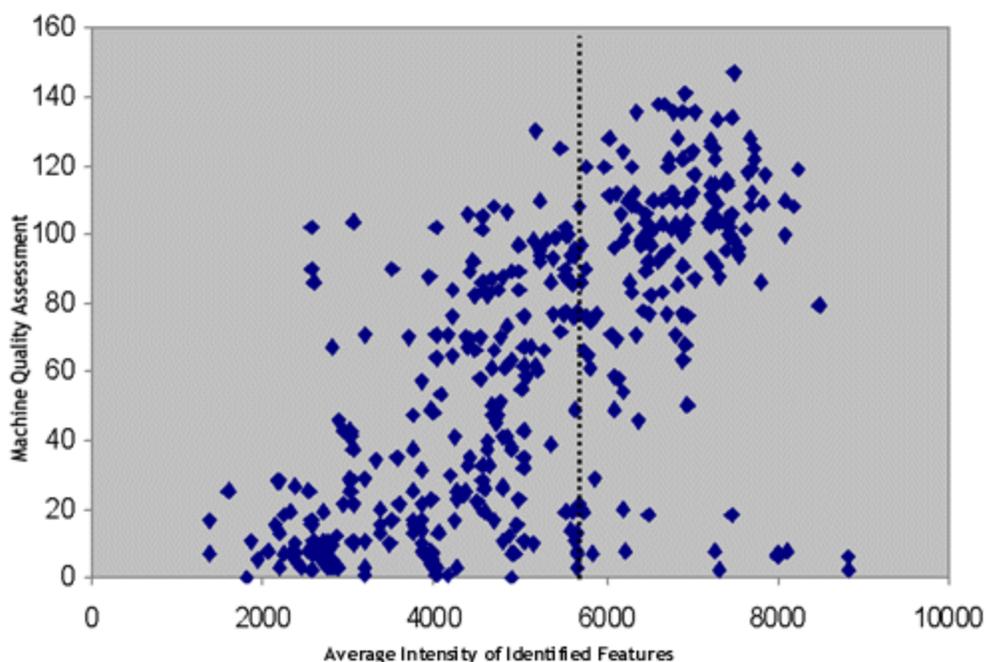


The results show that there is good correlation ($r^2=0.98$; $n=2194$) between the machine and human quality assessment. If we use the criterion of three or more layer lines as a measure of an acceptable images then this corresponds to a machine quality assessment of greater than 65. Thus automated assessment of image quality provides a reliable measure of an acceptable image and could be used to evaluate the overall performance of the system.

In order to compare the automated system to a microscopist we used the computer assisted interface, described above, to establish a baseline for the performance of a human operator. In this experiment, the microscopist identified the location of the features of interest on the low magnification image. A total of 288 high magnification images were acquired in this way and 79% of these were acceptable as defined above. In comparison, using the same grid, the fully automated image acquisition system was used to acquire 380 images of which 51% were acceptable.

Analysis of these results indicated a correlation between average feature intensity and image quality [fig 6]. This feature intensity is related to the thickness of the catalase crystal and indicates that thinner crystals result in more acceptable images. The fully automated target selection criteria was therefore further refined by incorporating an assessment of crystal thickness into the model. By acquiring high magnification images of only those features that have an average intensity greater than a preset threshold the percentage of acceptable images can be significantly improved. For example, if the threshold is set to 6000, the percentage of acceptable images improves to 86% from a baseline of 51%.

Figure 6.



4. DISCUSSION

We have developed a system for fully automated acquisition of 1000 electron micrographs a day. The Leginon system can be operated continuously over a 24 hour period and the only human intervention necessary is to fill the liquid nitrogen dewar.

There are significant variations in the percentage of acceptable images for a given grid and we have used grids of similar quality in all of the results discussed. Using very simple modeling criteria for feature selection, the number of acceptable images acquired by the Leginon system was comparable to the numbers acquired by a human operator. We have thus shown that we have developed a fully automated system that performs at least as well as a student microscopist.

There are a number of techniques that could be used to further improve the automated system's performance. First, the quality assessment could be improved by using indexing methods to more accurately identify diffraction spots and eliminate spurious spots. Other criteria (e.g. Friedel's law) could also be used to assess the quality of the crystals. Second, the feature identification and targeting methods could be much improved. For example, the variance of the intensities of the identified features is an indicator of the uniformity of the crystal thickness which should be incorporated into the targeting decision. Third additional targeting schemes should be used. For example, within large uniform crystals multiple targets should be identified. We intend to incorporate these improvements into the next application of the Leginon system which will be focused on automated acquisition of cryo-electron micrographs.

ACKNOWLEDGEMENTS

We thank Robert Josephs (University of Chicago) for his suggestion of using catalase as a test system and providing our initial specimen grids. Financial support and equipment was provided by the National Science Foundation (DBI-9730056) and the International Business Machines SUR program. Nahrstedt's research was supported by the National Science Foundation Career Grant (CCR-96-23867) and the National Computational Science Alliance. We are grateful to both the Beckman Institute and the National Center for Supercomputing Applications for institutional support.

REFERENCES

- [1] F.C. Fung, W. Liu, W. J. De Ruijter, H. Chen, C. K. Abbey, J. W. Sedat, and D. A. Agard, J. Struct. Biol. 116 (1996) 181.
- [2] G.T. Oostergetel, W. Keegstra and A. Brisson, Ultramicroscopy 74 (1998) 47.
- [3] K. Dierksen, D. Typke, R. Hegrl, A.J. Koster and W. Baumeister, Ultramicroscopy 40 (1992) 71.
- [4] C.S. Potter, B. Carragher, H. Chu, B.J. Frey, R. Josephs, C. Lin, N. Kisseberth, K.L. Miller and K. Nahrstedt, Microscopy and Microanalysis 4 Supp. 2 (1998) 259.
- [5] C.S. Potter, B. Carragher, H. Chu, B.J. Frey, R. Josephs, C. Lin, N. Kisseberth, K.L. Miller and K. Nahrstedt, Proc. 14th International Congress on Electron Microscopy, Vol 1 (1998) 8.
- [6] J.B. Sumner and A.L. Dounce, Meth. Enzymol. II (1955) 775.
- [7] N. Kisseberth, M. Whittaker, D. Weber, C.S. Potter and B. Carragher, J. Struct. Biology 120 (1997) 309.
- [8] A.P. Dempster, N.M. Laird and D.B. Rubin, Proceedings of the Royal Statistical Society, B-39 (1977) 1-38.
- [9] A.J. Koster and W. J. de Ruijter, Ultramicroscopy (1992) 40: 89.
- [10] R. Henderson, J. M. Baldwin, K.H. Downing, J. Lepault and F. Zemlin, Ultramicroscopy 19 (1986) 147.
- [11] J.K. Ousterhout, Tcl and the Tk Toolkit (Addison-Wesley, Reading, MA 1994).
- [12] Potter, C.S. and P.J. Moran, Biomedical Image Processing III and Three-Dimensional Microscopy 1660 (1992) 767.
- [13] J.P. Schroeter and J. P. Breaudiere, Struct. Biol. 116 (1996) 131.
- [14] K. Nahrstedt, H. Chu, and S. Narayan, Journal on High-Speed Networking, IOS Press, 1998. In press.