Visual Servoing for Micro Manipulation

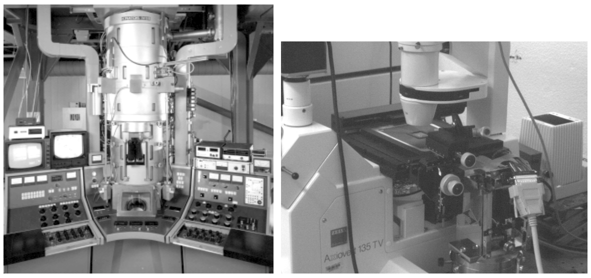
B. Parvin, D. E. Callahan, W. Johnston and M. Maestre
Imaging and Distributed Computing Group
Information and Computing Sciences Division
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
Email: parvin@george.lbl.gov

Abstract

A framework for automated micromanipulation of microscopic objects using visual information is developed and applied to diverse problems in biology and material science. These applications include microdissection of stretched and immobilized DNA molecules using an optical microscope and dynamic studies of crystal structure formation observed with a high power transmission electron microscope. Manipulation in this context refers not only to mechanical positioning or dissection of a specimen, but also to the variation of environmental parameters that elicit a response in the specimen. The common thread between these applications is a collection of computational techniques for image compression, autofocusing, object detection, tracking, and servo loop control. A second feature of our system is a distributed software architecture that allows the proposed dynamic experiments to be conducted remotely over the wide area network. In this context, the visual information processing is used to compensate for dynamic changes and the latencies in the wide area network. This system is now available over the global Internet.

1 Introduction

In this paper, we present a novel framework for manipulation of microscopic objects. We demonstrate the utility of this framework with two diverse applications that include microdissection of DNA molecules and shape equilibrium examination of in-situ crystal formation. The specifics of each application are discussed later. However, both of these scientific applications are labor intensive and can benefit from automation. The common thread between these applications is a set of computational techniques that include compression, autofocusing, detection, tracking, and servo-loop control. Another novelty of our system is in the remote operation of the proposed experiments over the wide area network. This is particularly significant for in-situ studies since rapid dynamic changes on the specimen and the unpredictable latency over the WAN have hitherto made this class of experiments unsuitable for remote operation. However, we have demonstrated that through intelligent visual interpretation and the use of this information for control, we can compensate for the latency in the WAN. The testbeds for the proposed experiments are an inverted epi-fluorescence optical microscope and a 1.5 MeV transmission electron microscope, of which the latter is now available for on-line use over the global Internet. Figure 1 shows the the high voltage electron microscope (HVEM)[1] and the robotic arm assembly mounted next to an inverted optical microscope.



(a) (b)

Figure 1: Microscopes: (a) high voltage electron microscope (HVEM); (b) robotic arm mounted near inverted optical microscope.

In the rest of this section, we discuss the significance of the two scientific applications. Then in sections 2 and 3, we will summarize the computational environment together with the collection of the proposed algorithms, respectively. In section 4, we will provide an example of software architecture for the dynamic studies, followed by concluding remarks in section 5.

1.1 Micro-dissection of DNA molecules

Precise automated micro-dissection of stretched and immobilized single DNA molecules has several important applications. If multiple (20-200) samples can be recovered from the same region of identical molecules that have been stretched to the same extent, then it is possible to amplify this material using the polymerase chain reaction (PCR). The microdissected and amplified DNA may then be used as a probe to identify members of large genomic libraries that contain DNA identical to that of the dissected region.

Human genomic libraries are large arrays of clones (yeast or bacterial cells) in which each clone contains one specific fragment of the human genome. The human genome consists of approximately 4 billion nucleic

*This work is supported by the Director, Office of Energy Research, Office of Computation and Technology Research, Mathematical, Information, and Computational Sciences Division, and Office of Basic Energy Sciences of the U. S. Department of Energy under contract No. DE-AC03-76SF00098 with the University of California. The LBNL publication number is 38809.

[1]This is a 1.5 MeV transmission electron microscope that is Operated by the National Center for Electron Microscopy

acid bases from 24 different chromosomes, and each human chromosome contains a single strand of DNA that is 50-250 megabases (Mb) in length. This genomic DNA is broken into fragments, and each clone of a genomic library contains a single fragment. Yeast clones containing yeast artificial chromosomes (YACs) may contain intact human DNA fragments as large as 1 Mb, whereas other types of more stable bacterial clones (BACs, PACs, or P1s) contain increasingly smaller intact fragments (75-300 Kb). In all cases, the clones in genomic libraries contain random and different but sometimes overlapping DNA fragments. The members of the library must be identified in a way that reconstructs the original order of the fragments in the human chromosomes. This process is first performed using genomic libraries, containing clones, with very large fragments (YACs). Next, members of the ordered YAC library serve as a templates on which to order clones from other libraries where the members contain smaller DNA fragments. One goal of the Human Genome Project is to obtain ordered sets of overlapping clones containing DNA fragments short enough (500-2000 base pairs) to submit to automated DNA sequencing machines. These ordered sets of clones can then be propagated and shared between laboratories for purposes of gene discovery and medical research.

The proposed application is designed to target regions on YACs or BACs where there are gaps in previously ordered sets of P1 clones. Closing gaps is currently the most time-consuming process in obtaining ordered sets of clones. Small, specific regions (5-10Kb) of the gap are recovered, amplified and then used as probes to identify clones that fill in the missing regions. The target region is specified in terms of fractional length. Since an extended YAC DNA can extend beyond one field of view, it would be difficult for a human operator to identify and dissect the correct region along a featureless DNA molecule. Any attempt to chemically process the DNA (e.g., fluorescence in situ hybridization, or FISH) so that a human operator can visualize the target region is likely to damage the DNA and make it impossible to amplify. In addition, for successful PCR amplification the microdissections should be performed quickly (to avoid excessive illumination), precisely, and repetitively (20-200 times on separate molecules). Thus, automation of the process is essential.

From an algorithmic perspective, the system needs to detect the position of a micro-capillary over the coverslip, select a DNA molecule, designate a position on the molecule, scrape a piece of that molecule, and verify the scraped area along the molecule.

1.2 Dynamic in-situ microscopy

Dynamic in-situ microscopy refers to a class of scientific experiments in which a specimen is perturbed by an external stimulus. The response of the specimen can be monitored and optimized during the course of the experiment. The stimulus could be in the form of temperature variation, electromagnetic field variation, or the variation of chemical or biological composition of the environment. The interaction of the external stimuli and the specimen could result in sample drift, shape deformation, changes in focus, and other responses, e.g., the physiological responses of an ensemble of living cells. The primary objective of the proposed experiment is to study crystal shape changes as a function of thermal cycles. The application monitors the shape of a crystal structure as it is cycled between liquid and solid states. This is referred to the as shape equilibrium problem, where the thermal path is optimized for identical shape features at solid and liquid states. During the in-situ experiments, the operator makes constant adjustments to the instrument maintaining focus and compensating for various drifts. It is a labor intensive task –requiring a high bandwidth video link– that is nearly impossible for teleoperation due to network latency. Therefore, teleoperation can be achieved only through the use of automated video analysis tools for manipulation and compensation of the necessary experimental parameters on a local area network. From a computational perspective, the system needs to detect precipitates[2], track their shape, alter the temperature, and compensate for drift and focus. In contrast to the previous application, in-situ studies require near-real-time response to changes in the scene. This requirement is met with parallel implementation of several key algorithms that will be discussed later.

2 Computational environment

The computational environment that implements the automated control in the local environment must be able to acquire images, process them at the required bandwidth, and manipulate a number of functions for operating the HVEM or the optical microscope. Our strategy for partitioning the required operations is based partly on design philosophy, i.e., scalability, modularity, and cost, and partly on the availability of data acquisition components for various hardware platforms. For these reasons, a Sun Microsystems workstation is used for image capture, a Digital Equipment Corporation (DEC) symmetric multiprocessor is used for CPU intensive operations, and a PC is used for data acquisition. Physically, the Sun and PC are operating near the microscope, while the DEC is located in another building of the Laboratory, connected via a LAN. The Sun and DEC are on a FDDI ring (100 Mb/s) for high speed image transfer as shown in figure 2. In this configuration, the local Sun workstation is mainly used for testing and on-line quantitative analysis by local users.

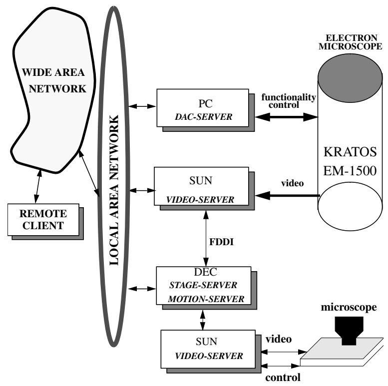


Figure 2: Computational environment for remote in-situ microscopy. The italic names specify the servers that run on each platform.

3 Computational techniques

From an algorithmic perspective, a common set of requirements for the proposed applications includes 1) image compression and autofocusing, 2) self calibration, 3) object detection, 4) tracking, and 5) servo-loop control. These are well known techniques in the field of computer vision and we have built on existing knowledge [1, 7, 8, 6] to construct our system. In the context of remote microscopy, the system must provide the look and feel available to the local

[2] a crystal structure, e.g. lead, germanium, etc.

operator, and hide the inherent latency in the wide area network. The look and feel is achieved through an appropriate user interface. The hiding of the network latency is achieved through *visual servoing*. The computational components of the system are outlined below.

3.1 Image compression and autofocusing

Both image compression and autofocusing use the wavelet transform as their underlying principle. We use Daubechies kernels [4] that are simple, orthogonal, highly localized, and separable for two dimensional processing. The main advantage of the wavelet transform is that it can represent local feature activities at multiple scales through spatial decimation. During image compression, the low order wavelet coefficients are ignored and the remaining ones are encoded in blocks of 16-by-16 pixels. The remote user has full control over what percentage of wavelet coefficients are used for compression. Autofocusing has two modes of operation: initialization and run-time.[3] The difference between the initialization and run-time mode is based on the scope of the search for the best focal position. The goodness of the focus is measured by the sum of the wavelet coefficients. At run-time –as the specimen is heated– small adjustments are made in the focal position to compensate for 3D changes of the precipitate position. Similarly, in the case of DNA microdissection the coverslip is not perfectly flat and small adjustment to the focal position is made as the stage, holding the coverslip, is moved from one location to the next.

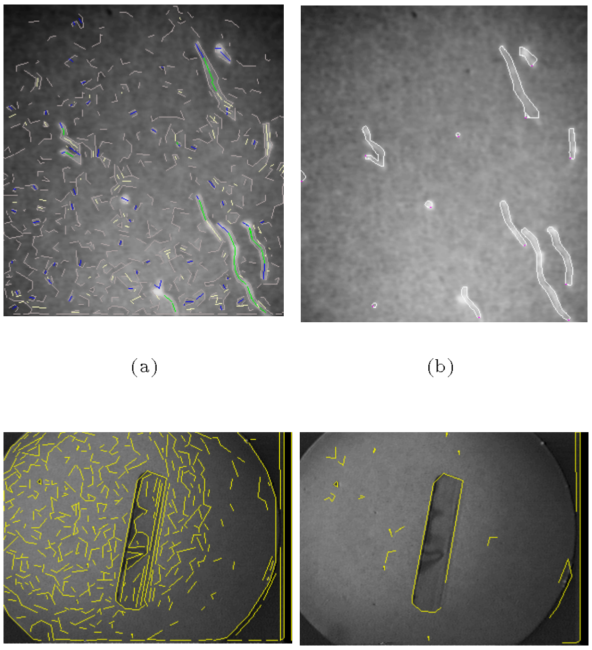
3.2 Self calibration

The motion of the XY stage or the robotic arm is precisely mapped into the actual pixel size. This mapping is an affine transform that is estimated by either low or high level features. Most microscopy applications have a narrow depth of field. As a result, the motion is constrained to a 2D flat world that simplifies the calibration stage. In the case of micro-dissection of DNA molecules, the transform is obtained by projecting a deformable contour model on a capillary. The arm is then moved into two known positions and re-acquired with the deformable model. Calibration of the XY stage –either optical or electron microscope– is performed with the optical flow field [1]. The image motion, as perceived by the XY translation stage, is affine (rotated and slightly off the scale in one axis), and a pyramid implementation of the OFF is used to obtain the motion parameters rapidly.

3.3 Detection

The system supports three modes of object detection. These are, respectively, for tubular, convex, and circular objects. DNA molecules, observed under the fluorescence microscope, are modeled as a tubular objects. Precipitates, observed under a transmission electron microscope, are modeled as convex objects. And living cells, observed under the optical microscope, are modeled as circular objects. Object detection is based on perceptual grouping principles [5, 9], and it is initiated by grouping line segments obtained from Canny's edge operator [3]. The detection system provides a coarse description of the objects of interest in the form of bounding polygons. This description is then refined, projected, and tracked in subsequent frames using the tracking component. In this context, object detection occurs only in the first frame for the purpose of initialization. Detection of DNA molecules is achieved by grouping U-shapes and anti-parallel segments. In a sense, the grouping is performed along the axis of symmetry. Detection of precipitates is achieved by accumulating line segments as long as the convexity constraint is satisfied. The actual search process is based on dynamic programming and the approach is beyond the scope of this paper [7, 8]. Examples of detection of DNA molecules and precipitates are shown in figure 3.

[3] The focus is adjusted either by a 16-bit D/A converter controlling the lens current for HVEM or a stepper motor controlling the Z-axis for the optical microscope.



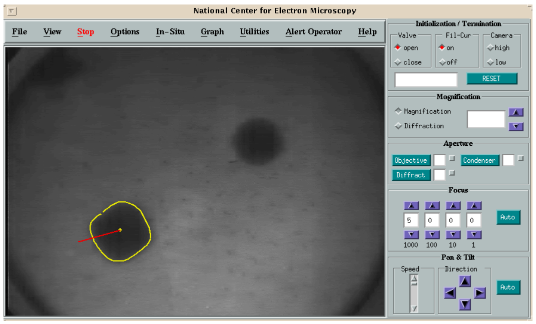
(a) (b)

(c) (d)

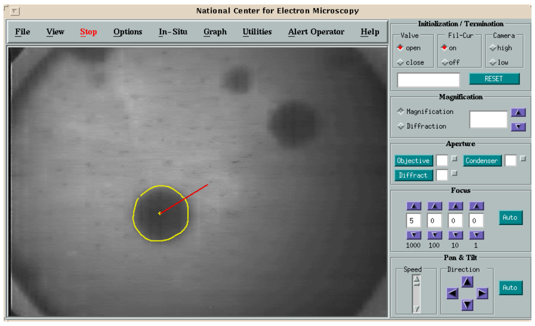
Figure 3: Detection of DNA molecules and precipitate: (a) original image with Canny's edges; (b) detected tubular objects; (c) original image with Canny's edges; (c) detected convex sets.

3.4 Tracking subsystem

Our system supports two modes of tracking based on high and low level feature activities corresponding to deformable contours and optical flow field. During dynamic in-situ experiments on the electron microscope, it is desirable to to compensate for thermal drift as the specimen is heated. We have built a closed-loop system using optical flow field computation to stabilize the stage where the specimen is located. Current throughput is at 4 Hz using a multi-grid implementation of the optical flow technique that relies on minimizing the difference of consecutive frames by estimating the affine transform between those two frames. This is the same technique that is used for self calibration. The actual software implementation uses threads to realize the necessary parallelism under the OSF operating system. The second mode of tracking is based on high level shape changes, where a variant of the deformable model, based again on dynamic programming [7], is used to quantify and track objects of interest under the microscope. We use a multigrid implementation of this algorithm for maximum speed-up and higher tolerance for large motion. The algorithm has performed well in the presence of shading, noise, nonuniform illumination, and reduced contrast. The system automatically tracks the shape, controls the drift, and hides the network latencies from the remote user. The drift control is based on tracking and compensating for the centroid of the contour. This is shown in figure 4, where the centroid is shown with a cross-hair on the reconstructed image from the wavelet coefficients, and the direction of the motion is shown with an arrow. In addition to the topological changes in the shape, during the heating and cooling experiment, thermal drift reverses its direction as well, which is also reflected in the figure.



(a)

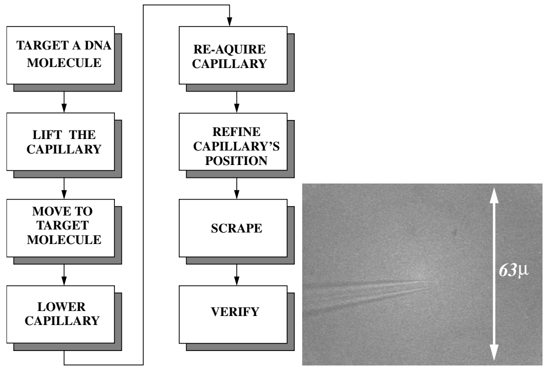


(b)

Figure 4: Tracking and compensating for drift during heating and cooling cycles. Note that direction of drift is reversed as the specimen is cooled: (a) Precipitate is initially faceted as it is heated; (b) Precipitate becomes round at high temperature.

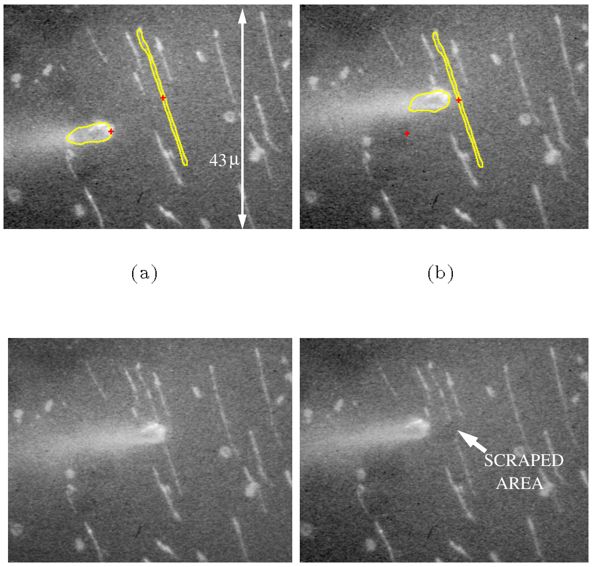
3.5 Servo loop control

There are two aspects of servo loop control: positioning of the robotic arm for scraping of DNA molecules and thermal drift correction during in-situ experiments. Both of these applications use deformable contours to acquire, refine, and reposition the mechanical assembly. However, the placement of the robotic arm is imprecise, and needs to be improved for increased accuracy. This is accomplished by moving the microcapillary near the scraping site, reacquiring the microcapillary with the deformable contour, repositioning it to where it should be, scraping the site, and then verifying the scraped site by simple thresholding. This protocol is graphically shown in figure 5 and an example of microdissections is shown in figure 6. In con-



(a) (b)

Figure 5: Protocol for scraping of a DNA molecule shown with a capillary viewed under transmitted light source.



(a) (b)

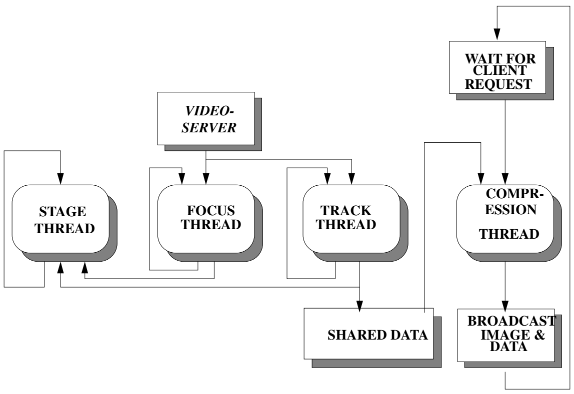
(c) (d)

Figure 6: Microdissection of a single DNA molecule (lamda DNA at 48.5 Kb) under fluorescence light: (a) original image with capillary and a number of DNA molecules; (b) capillary is reacquired near the scraping site; (c) molecule is scraped; and (d) the scraped area is verified.

trast, the thermal drift correction is continuous, smooth and linear, and a Kalman filter model is used for predicting the motion. Kalman filtering has been used extensively for smoothing, filtering, and prediction as reported in the literature [2]. In general, the model provides smooth compensation for drift and shape tracking coupled with high tolerance for larger speed. Our implementation uses position and velocity to represent the internal state of the precipitate. In this context, the model is used to predict the trajectory of the motion. As a result, instead of making incremental correction to the XY stage platform, we place the stepper controller at a constant speed in the direction opposite to the thermal drift. The speed is then refined at the tracker sampling interval. The detail of the Kalman filtering model is given in Appendix A.

4 Software architecture for dynamic in-situ experiment

The software architecture follows a distributed client-server model for scalability, performance, and modularity. There are four servers that can interact with each other in the architecture shown in figure 2. These are the i) video-server, ii) motion-server, iii) stage-server, and iv) DAC-server. The video-server –running on the Sun– captures images and transfer them in their entirely or partially to the motion-server. The motion-server –running on the DEC– manages all the image analysis and servoing. These modules are executed asynchronously and use a threads programming paradigm for parallel decomposition. The stage-server –running on the DEC– handles all the manual interaction between the remote user and the electron microscope, i.e., changing magnification, shifting the beam, etc. The DAC-server –running on the PC– reads and writes into the data acquisition components for a desired function. The DAC-server uses remote procedure calls for communication and the remaining servers use data streams through sockets for minimum delay. A critical component of the system is in the design of the motion-server. This server has four threads that run asynchronously as shown in figure 7. The stage-thread handles all the interaction with the stage-server, and it has been isolated for modularity and higher throughput. Average time for most interaction with the PC is about 7 ms. The tracking thread operates at 5-8 Hz depending to the size of the precipitate, and runs with a concurrency of two. The compression-thread runs at 1.4 Hz over the shared data, and the focus-thread runs on a single thread over the target region when the tracking thread is inactive.



(a)

Figure 7: State diagram for motion-server indicates four distinct threads that run asynchronously.

5 Conclusion

A framework for automated micromanipulation of microscopic objects is applied to two diverse applications in scientific imaging. The common thread between these applications is a collection of computation tools that allow these experiments to be conducted with little human intervention. A second feature of our system is in the software architecture that allows the proposed experiments to be conducted remotely over the wide area network.

Acknowledgment: The authors thank the support of J. Taylor, B. Crowley, D. Owen, M. Okeefe, and U. Dahmen for this project.

A Kalman model

The state space representation of the system is as follows:

$$\begin{aligned} X(k+1) &= AX(k) + n_x(k) \\ y(k) &= CX(k) + n_y(k) \end{aligned} \tag{1}$$

Where X and y correspond to the state and observation vector, n_x and n_y are noise vectors, and A and C are constant matrices. The state vector consist of position and velocity components. In this context, the above set of equations can be rewritten as:

$$\begin{aligned} x(k+1) &= x(k) + v(k) + n_x(k) \\ v(k+1) &= v(k) + n_v(k) \\ y(k) &= x(k) + n_y(k) \end{aligned} \tag{2}$$

Then, the filter equations will be:

$$Gain: \quad K(k+1) = P(k+1|k)C^T[CP(k+1|k)C^T + R]^{-1} \tag{3}$$

$$Where \quad P(k+1|k) = AP(k|k)A^T + Q^* \tag{4}$$

$$Update: \quad \hat{X}(k+1|k+1) = \hat{X}(k+1|k) - K(k+1)[C\hat{X}(k+1|k) - y(k+1)]$$

$$P(k+1|k+1) = (I - K(k+1)C)P(k+1|k) \tag{5}$$

References

[1] J. Bergen, R. Hingorani, and S. Peleg. A 3-frame alorithm for estimating 2-component image motion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:886–896, 1992.

[2] T. J. Broida and R. Chellappa. Estimation of object motion parameters from noisy images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:90–99, 1986.

[3] J. F. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, November 1986.

[4] I. Daubechies. *Ten Lectures on Wavelet*. SIAM, Philadelphia, PA, 1992.

[5] A. Huertas, W. Cole, and R. Nevatia. Detecting runways in complex airport scenes. *Computer Vision, Graphics, and Image Processing*, 51(2):107–145, August 1990.

[6] B. Parvin and et al. Telepresence for in-situ microscopy. In *IEEE Int. Conference on Multimedia Systems and Computers*, 1996.

[7] B. Parvin, C. Peng, W. Johnston, and M. Maestre. Tracking of tubular molecules for scientific applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:800–805, 1995.

[8] B. Parvin, S. Viswanathan, and U. Dahmen. Tracking of convex objects. In *Int. Symp. on Computer Vision*, 1995.

[9] A. Sha'ashua and S. Ullman. Structural saliency: the detection of globally salient structures using a locally connected network. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 321–327, Tampa, FL, 1988.