Extracting 3-D Structure and Focused Images Using an Optical Microscope

Ushir B. Shah and Shree K. Nayar

Department of Computer Science,
Columbia University, New York, N.Y. 10027

Abstract

In this paper, we describe a novel method for recovering three-dimensional shapes and focused images of microscopic objects. A small sequence of object images is obtained by translating the object through the focused plane of the microscope. The sum-modified-Laplacian (SML) operator is developed to compute local measures of the quality of image focus. The SML operator is applied to the image sequence, and the focus measures obtained at each image point are used to compute local depth estimates. We present two algorithms for depth estimation. The first algorithm simply looks for the focus level that maximizes the focus measure at each image point. The second algorithm uses a model to interpolate the focus measures to obtain more accurate depth estimates. The algorithms were implemented and tested on variety of biological specimen. We conclude with a brief description of a fully automated shape-from-focus system that is currently being used to inspect both medical as well as industrial samples.

1 Introduction

The advancement of three-dimensional machine vision is largely dependent on the development of efficient and reliable shape extraction methods. Several shape recovery methods have been developed for large objects with simple reflectance properties. Microscopic objects with complex and varying reflectance properties pose a challenging shape recovery problem; image intensities produced by such objects vary substantially from one sensor element (pixel) to the next. Consequently, it is difficult to obtain dense and accurate surface shape information by using existing passive or active machine vision techniques such as stereo, shape from shading, and structured light. A practical and reliable solution to this rather difficult extraction problem is therefore desirable. In this paper, we develop a shape extraction technique that uses focus analysis to recover dense depth maps of complex microscopic samples.

1.1 Background

Previously, in machine vision, focus analysis has been used to automatically focus imaging systems or to obtain *sparse* depth information from the observed scene. Horn [1] proposed focusing imaging systems by using the Fourier transform and analyzing the frequency spectrum of the image. Tenenbaum [2] developed the gradient magnitude maximization method that uses the sharpness of edges to optimize focus quality. A modification to this approach was later proposed by Jarvis [3]. He formulated the sum-modulus-difference as the sum of the first intensity differences between neighboring pixels along a scan-line and used it as a measure of focus quality. Several automatically focusing algorithms were implemented and tested by Schlag et.al. [4].

More recently, Krotkov [5] evaluated and compared the performance of different focus criterion functions. Krotkov also proposed a method to estimate the depth of an image area. Pentland [6] suggested estimating the depth of image points by evaluating image blur due to defocusing. A similar approach was applied to edge points by Grossmann [7]. Darrell and Wohn [8] have developed a depth from focus method that obtains an image sequence of a scene by varying the focus level, and uses Laplacian and Gaussian pyramids to calculate depth. Subbarao [9] suggests the change of intrinsic camera parameters to recover the depth map of a scene.

1.2 Shape from Focus

In this paper, we develop a shape-from-focus method. In contrast to previous work in this area, we avoid the following approaches.

- We do not attempt to estimate depth from a pair of images by evaluating local estimates of the blurring function. The accuracy of such

0-8186-2742-5/92 $3.00 © 1992 IEEE

a method is greatly dependent on the blurring model used. The models used thus far are only approximations to the actual physical-optics model and therefore do not ensure high quality results.

- We do not apply our method to general scenes. Depth estimation based on focus analysis relies on the presence of high frequency brightness variation in the scene. General scenes often have areas with little or no brightness variation. For this reason, experiments in the past have only produced sparse depth information.

Here, we restrict ourselves to surfaces that produce complex textured images with high frequency intensity variations. We review the image formation process and show that a defocused imaging system plays the role of a low-pass filter. The shape-from-focus method moves the unknown object with respect to the imaging system and obtains a sequence of images that correspond to different levels of object focus. The sum-modified-Laplacian (SML) focus operator is developed to measure the relative degree of focus between images. The operator is applied to the image sequence to obtain a set of focus measures at each image point. A model is proposed that describes focus measure variations due to defocusing. This model is used to interpolate between a finite number of focus measures to obtain accurate depth estimates. Experimental results indicate that the method is capable of extracting dense and accurate shape information from a few images of the object. The results demonstrate appreciable invariance to surface roughness and reflectance properties.

Motivated by these results we have developed a fully automated shape-from-focus system. The system is applicable to samples that are upto a hundred microns in size. An optical microscope is used to image the samples. The vertical axis of the translation stage of the microscope is motorized to enable controlled movement of the sample through the plane of focus of the imaging system. The stage is translated in increments and for each new position an image of the sample is obtained. The image sequence is processed on a workstation. Two results are produced by the recovery algorithm. The first is a focused image of the sample that is reconstructed from the sequence of partially focused images. The second is a depth map (3-D shape) of the sample surface. The automated system has been applied to a variety of biological samples. In this paper, we present results obtained in the case of a mature embryo of a fruit fly (drosophila melanogaster).

2 Focused and Defocused Images

In this section, we briefly review the image formation process and describe defocused images as processed versions of focused images. Fig.1 shows the basic image formation geometry. All light rays that are radiated by the object point P and intercepted by the lens are refracted by the lens to converge at the point Q on the image plane. The relationship between the object distance o, focal distance of the lens f, and the image distance i, is given by the Gaussian lens law:

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}. \tag{1}$$

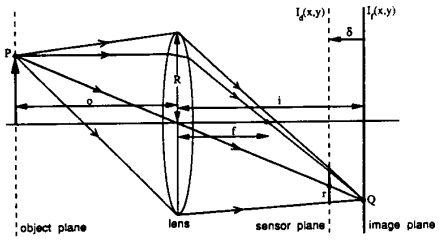


Figure 1: Formation of focused and defocused images.

Each point on the object plane is projected onto a single point on the image plane, thus causing a clear or *focused* image $I_f(x,y)$ to be formed on the image plane. If, however, the sensor plane does not coincide with the image plane and is displaced from it by a distance δ, the energy received from the object by the lens is distributed over a circular[1] patch on the sensor plane. Fig.1 may be used to establish the relationship between the radius r of the circular patch and the sensor displacement δ. From Fig.1 we find that:

$$r = \frac{\delta R}{i} \tag{2}$$

where R is the radius of the lens. It is also possible to convince oneself that the radius r of the circular patch is independent of P's location on the object plane. The distribution of light energy over the circular patch, or the *blurring function*, can be modeled using physical optics [11]. Very often, a two-dimensional Gaussian function is used to approximate the physical model [6]. Then, the blurred or *defocused* image $I_d(x,y)$ formed on the sensor plane can be described as the result of convolving the focused image $I_f(x,y)$ with the blurring function $h(x,y)$:

$$I_d(x,y) = h(x,y) * I_f(x,y) \tag{3}$$

[1]The shape of the patch also depends on the shape of the aperture of the imaging system. We are assuming the aperture to be circular.

where:

$$h(x,y) = \frac{1}{2\pi\sigma_h^2} e^{-\frac{x^2+y^2}{2\sigma_h^2}} \quad (4)$$

where σ_h, the *spread parameter*, is assumed to be proportional to the radius r [6]. The constant of proportionality is dependent on the optics, sampling, etc. We will see shortly that the value of this constant is not important in our approach. Note that defocusing is observed for both positive and negative sensor displacements.

Now let us analyze the defocusing process in the frequency domain (u,v). If $I_F(u,v)$, $H(u,v)$, and $I_D(u,v)$ are the Fourier transforms of $I_f(x,y)$, $h(x,y)$, and $I_d(x,y)$, respectively, we can express eq. 3 as:

$$I_D(u,v) = H(u,v).I_F(u,v) \quad (5)$$

where:

$$H(u,v) = e^{-\frac{u^2+v^2}{2}\sigma_h^2} \quad (6)$$

We see that $H(u,v)$ allows low frequencies to pass while it attenuates the high frequencies in the focused image. Furthermore, as the sensor displacement δ increases, the defocusing radius r increases, and the spread parameter σ_h increases. Hence defocusing is a *low-pass* filtering process where the bandwidth decreases with increase in defocusing.

From Fig.1, it is seen that a defocused image of the object can be obtained in three ways: by displacing the sensor with respect to the image plane, by moving the lens, or by moving the object with respect to the object plane. Moving the lens or sensor plane with respect to one another causes the following problems:

- The magnification of the system varies, thereby causing the image coordinates of focused points on the object to change.
- The area on the sensor plane over which light energy is distributed varies, thereby causing a variation in image brightness.

These effects are described in detail by Willson and Shafer [14]. In order to overcome these problems, we propose to vary the degree of focus by moving the object[2] with respect to a fixed configuration of the optical system and sensor (Fig. 2). This approach ensures that as the object passes through the plane S, surface points that lie on S are perfectly focused onto the image plane with the same magnification. In other words, as the object moves, the magnification of imaging system can be assumed to be constant for the image areas that are perfectly focused.

[2]Object movement is easily realized in the case of optical microscopes.

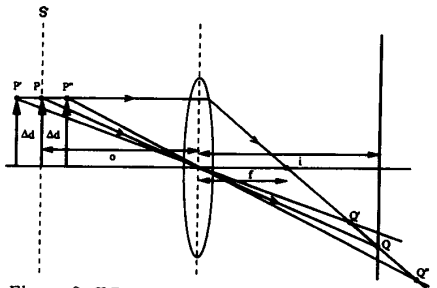


Figure 2: Effect of object displacement on magnification.

However, from Fig. 2 we see that points that lie outside of the plane S will be projected onto the image plane with different magnifications. In fact, the magnification of the defocused object points will depend on their distance from the plane S. Note that for small displacements Δd, magnification may be assumed to be constant. We will use this assumption while developing our depth estimation algorithm.

3 Shape from Focus: An Overview

Fig.3 shows an unknown object placed on a translational stage. The reference plane shown corresponds to the initial position of the stage. The configuration of the optics and sensor defines a single plane, the "focused plane[3]," that is perfectly focused onto the sensor plane. The distance d_f between the focused and reference planes, and the displacement d of the stage with respect to the reference plane, are always known by measurement. Consider the surface element, s, that lies on the unknown surface, S. If the stage is moved towards the focused plane, the image of s will gradually increase in its degree of focus (high frequency content) and will be perfectly focused when s lies on the focused plane. Further movement of the element s will again increase the defocusing of its image. If we observe the image area corresponding to s and record the stage displacement $d = \overline{d}$ at the instant of maximum focus, we can compute the height d_s of s with respect to the stage as $d_s = d_f - \overline{d}$. In fact, we can use $\overline{d}$ to determine the distance of s from the focused plane, sensor plane, or any other coordinate system defined with respect to the imaging system. This procedure may be applied independently to all surface elements to obtain the shape of the entire surface S.

To automatically detect the instant of "best" focus, we will develop an image focus measure. In the

[3]The focused plane is the same as the object plane defined in the previous section. A different term is introduced here as the object does not necessarily lie on the focused plane.

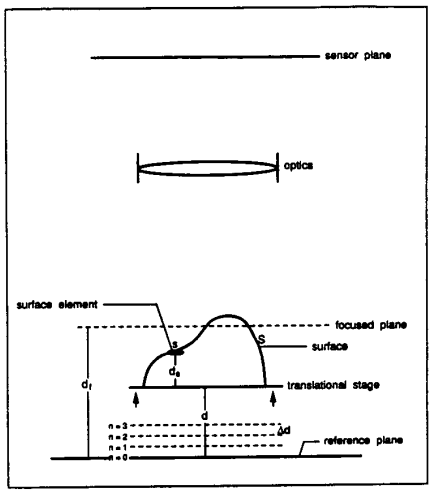


Figure 3: Shape from focus.

above discussion, the stage motion and image acquisition were assumed to be continuous processes. In practice, however, it is not feasible to acquire and process such a large number of images in a reasonable amount of time. Therefore, we obtain only a small number of images; the stage is moved in increments of Δd, and an image is obtained at each stage position ($d = n.\Delta d$). By studying the behavior of the focus measure, we develop an interpolation method that uses a small number of focus measures to compute accurate depth estimates. An important feature of the method is its local nature; the depth estimate at an image point is computed only from focus measures recorded at that point. Consequently, the method can adapt well to variations in texture type and content over the object surface.

4 A Focus Measure Operator

To measure the quality of focus in a small image area, we develop a focus measure operator. The operator must respond to high frequency variations in image intensity, and ideally, must produce maximum response when the image area is perfectly focused.

An interesting observation can be made regarding the application of focus measure operators. Eq. 3 relates a defocused image to a focused image using the blurring function. Assume that a focus measure operator $o(x,y)$ is applied (by convolution) to the defocused image $I_d(x,y)$. The result is a new image $r(x,y)$ that may be expressed as:

$$r(x,y) = o(x,y) * (h(x,y) * I_f(x,y)) \quad (7)$$

Since convolution is a linear operation, we can rewrite the above expression as:

$$r(x,y) = h(x,y) * (o(x,y) * I_f(x,y)) \quad (8)$$

Therefore, applying a focus measure operator to a defocused image is equivalent to defocusing a new image obtained by convolving the focused image with the focus measure operator. The focus measure operator only selects the frequencies in the focused image that will be attenuated due to defocusing. Since defocusing is a low-pass filtering process, its effects on the image are more pronounced and detectable if the image has strong high frequency content. An effective focus measure operator, therefore, must high-pass filter the image.

One way to high-pass filter an image is to determine its second derivative. For two-dimensional images, the *Laplacian* may be used:

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \quad (9)$$

where $I(x,y)$ is the image intensity at the point (x,y). In frequency domain, applying the Laplacian $L(u,v)$ to the defocused image $I_D(u,v)$ (eq. 5) gives:

$$L(u,v).H(u,v).I_F(u,v) \quad (10)$$

where:

$$L(u,v).H(u,v) = -(u^2+v^2).e^{-\frac{u^2+v^2}{2}\sigma_h^2} \quad (11)$$

For any given frequency (u,v), $|\ L.H\ |$ varies as a Gaussian function of the defocusing parameter σ_h. In general, however, the result would depend on the frequency distribution of the imaged scene. Though our texture is in general random, it may be assumed to have a set of dominant frequencies. Then, loosely speaking, each frequency is attenuated by a Gaussian function in σ_h and its width is determined by the frequency. Therefore, the result of applying the Laplacian operator may be expressed as a sum of Gaussian functions in σ_h. The result is expected to be maximum when $\sigma_h = 0$, i.e. when the image is perfectly focused. Since the frequency distribution of the texture is random, the widths of the Gaussian functions are also random. Using central limit theorem, the result of applying the Laplacian operator to an image point may be assumed to be a Gaussian function of the defocusing parameter σ_h. This general behavior is expected irrespective of the focus measure operator used. The focus measure operator only selects the frequencies that will play a dominant role in this process.Our experiments (section 4) as well as Krotkov's empirical evaluation of various focus measure operators [5] support the above argument. As seen in [5], image noise and magnification variations will of course degrade the performance any focus measure operator.

In the context of textured images, the Laplacian poses a problem as a focus measure operator. Note that in the case of the Laplacian the second derivatives in the x and y directions can have opposite signs and tend to cancel each other. An example of such an instance is illustrated in Fig.4; the partial derivatives are equal in magnitude but opposing in sign, i.e. $\nabla^2 I = 0$. In the case of textured images, similar instances may occur frequently and the Laplacian may at times behave in an unstable manner. We overcome this problem by defining the *modified Laplacian* as:

$$\nabla^2_M I = \left|\frac{\partial^2 I}{\partial x^2}\right| + \left|\frac{\partial^2 I}{\partial y^2}\right| \quad (12)$$

Note that the modified Laplacian is always greater or equal in magnitude to the Laplacian.

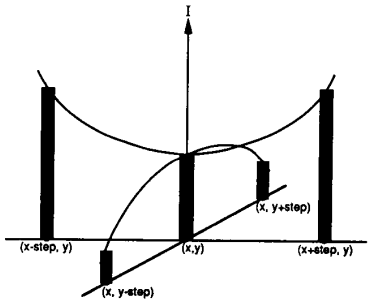


Figure 4: A texture instance with zero Laplacian value.

The discrete approximation to the Laplacian is usually a 3×3 operator. In order to accommodate for possible variations in the size of texture elements in the image, we compute the partial derivatives by using a variable spacing (*step*) between the pixels used to compute the derivatives. Hence, the discrete approximation to the modified Laplacian is computed as:

$$ML(x,y) = \quad (13)$$
$$|\ 2I(x,y) - I(x-step,y) - I(x+step,y)\ |$$
$$+\ |\ 2I(x,y) - I(x,y-step) - I(x,y+step)\ |$$

Finally, the focus measure at a point (i,j) is computed as the sum of modified Laplacian values, in a small window around (i,j), that are greater than a threshold value:

$$F(i,j) = \quad (14)$$
$$\sum_{x=i-N}^{i+N} \sum_{y=j-N}^{j+N} ML(x,y) \quad for\ ML(x,y) \geq T_1$$

where the parameter N determines the window size used to compute the focus measure. In contrast to auto-focusing methods, we typically use a small window of size 3×3 or 5×5, i.e. $N = 1$ or $N = 2$. We shall refer to the above focus measure as the *sum-modified-Laplacian* (SML). Note that as a result of definition of the modified Laplacian and the use of the threshold T_1, the SML is not a linear operator and cannot be implement as a convolution. However, it can be computed using a simple algorithm.

5 Evaluating the Focus Measure

We evaluate the SML focus measure by analyzing its behavior as a function of the distance between the observed surface and the focused plane. A detailed description of the experimental set-up is given in a later section. In the following experiments, texture samples are attached to a translational stage (Fig.3) and the distance, d_s, from each sample to the stage is known by measurement. Images of the samples are obtained using a microscope and a 512×512 pixel CCD camera. The complete imaging system has a physical resolution of approximately $0.5\mu m$ per pixel width.

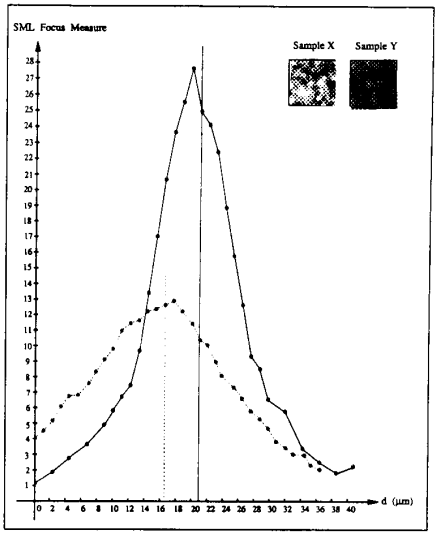


Figure 5: SML focus measure function computed for two texture samples.

In Fig.5, the focus measure functions of two samples are shown. Sample X has high texture content while sample Y has relatively weaker texture. Both samples are small sections of industrial circuit boards. For each sample, the stage is moved in increments (Δd) of $1\mu m$, an image of the sample is ob-

tained, and the SML focus measure is computed using an evaluation window size of 10×10 pixels. The vertical lines in Fig.5 indicate the known initial distances (d_f - d_s) of the samples from the focused plane. The focus measures were computed using parameter values of $step = 1$ and $T_1 = 7$. No form of temporal filtering was used to reduce the effects of image noise, as we intend to use unfiltered focus measures to estimate the depth of surface points. Though the measure values are slightly noisy, they peak very close to the expected peak positions (vertical lines in Fig.5). We see that the focus measure function peaks sharply for the stronger texture and it peaks relatively slowly and with a lower peak value for the weaker texture. However, the sharpness of the focus measure function depends not only on the texture strength but also the depth of field of the imaging system. The depth of field, in turn, depends on the magnification and aperture size of the imaging system. We will assume that the depth of field is constant for all our experiments. Note that the focus measure functions for both samples have Gaussian-like distributions near their peak values. The fringes of the focus measure functions are less symmetric as the magnification can vary substantially from one fringe to the other.

6 Sampling the Focus Measure Function

The focus measure function of an image point (x, y) may be represented as $F(x,y,d)$. Since depth estimation is a local operation, we focus our attention on a single image point, bearing in mind that the same estimation method can be applied to all other image points. The focus measure function at the image point is $F(d)$. From the previous sections we know that $F(d)$ has a Gaussian distribution near the peak, with mean value $\overline{d}$ and standard deviation σ_F (Fig.6). The mean $\overline{d}$ corresponds to the stage displacement at which $F(d)$ is maximum, i.e. $F(\overline{d}) = F_p$. As the texture content on the surface element increases, F_p increases and σ_F decreases. Each surface element, therefore, is expected to have its own F_p and σ_F values.

If we use very small stage displacements ($\Delta d \approx 0$), the number of images to be obtained and processed is too large from the perspective of practical implementation. Hence, we use large displacements to obtain a few images of different focus levels and use the Gaussian model to interpolate a small number of focus measures in the peak region to obtain depth estimates. Computing focus measures at a finite number of stage displacements is equivalent to *sampling* the function $F(d)$; at each displacement d_i the focus measure $F(d_i)$ is computed to obtain the set $\{F(d_i) \mid i = 1, 2, ...M\}$. We show in the following section that a minimum of *three* focus measures are needed perform the Gaussian interpolation. Since the Gaussian model is valid only in the peak region of $F(d)$, these three focus measures must be computed in this region. This is ensured by using the condition $\sigma_F \leq \Delta d \leq 2\sigma_F$. Note that all object points are subjected to the same displacements. Therefore, by applying the above condition to the image area that has maximum texture content, we can ensure that a few or many focus measures will be computed in the $\pm\ \sigma_F$ range at all other image points.

The value of σ_F also increases with the depth of field of the imaging system. Therefore, for objects of larger dimensions also, only a small number of images may used by increasing the depth of field.

7 Depth Estimates from Focus Measures

We now describe the estimation of depth of a surface point (x, y) from the focus measure set $\{F(d_i) \mid i = 1, 2, ...M\}$. The parameter $\overline{d}$ represents the depth of the surface point. For convenience the notation F_i is used to represent the focus measure value $F(d_i)$. We present algorithms for two different depth estimation methods. Each algorithm may be applied to all points in the image to obtain depth maps.

7.1 Coarse Resolution Depth Estimation

The first algorithm simply looks for the displacement value d_i that maximizes the focus measure and assigns that value to $\overline{d}$.

Algorithm 1

Step 1: Let $k = 1$, $F_{max} = 0$.

Step 2: If $F_k > F_{max}$, then $F_{max} = F_k$ and $\overline{d} = d_k$.

Step 3: If $k < M$, then $k = k + 1$, go to step 2. Else stop.

Step 4: If $F_{max} < T_2$, the point (x, y) belongs to the background.

This simple algorithm may be used to compute low resolution depth estimates. The performance of the algorithm is directly dependent on the selection of Δd. If Δd is small, a large number of object images are obtained and the depth maps are more accurate.

7.2 Depth Estimation by Gaussian Interpolation

The second algorithm uses the Gaussian distribution to model the peak region of the focus measure function $F(d)$ and interpolates the computed measure values to obtain more accurate depth estimates. The following algorithm uses only three focus measures, namely, F_{m-1}, F_m, and F_{m+1}, that lie on the largest mode[4] of $F(d)$, such that, $F_m \geq F_{m-1}$ and $F_m \geq F_{m+1}$ (Fig.6).

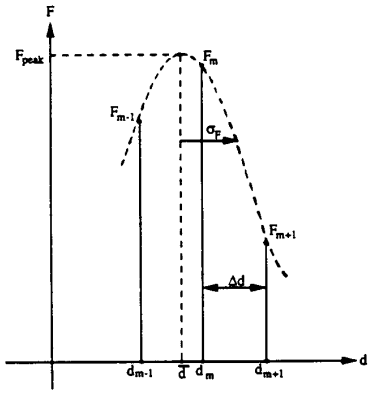


Figure 6: Gaussian interpolation of focus measures.

Using the Gaussian model, the focus measure function may be expressed as:

$$F = F_p \; exp \left\{ -\frac{1}{2}\left(\frac{d-\overline{d}}{\sigma_F}\right)^2 \right\} \qquad (15)$$

where $\overline{d}$ and σ_F are the mean and standard deviation of the Gaussian distribution (Fig.6). Using natural logarithm, we can rewrite eq. 15 as:

$$ln\, F = ln\, F_p - \frac{1}{2}\left(\frac{d-\overline{d}}{\sigma_F}\right)^2 \qquad (16)$$

By substituting each of the three measures F_{m-1}, F_m, and F_{m+1}, and its corresponding displacement value in eq. 16, we obtain three equations that can be solved for $\overline{d}$ and σ_F:

$$\overline{d} = \frac{(ln\, F_m - ln\, F_{m+1})(d_m^2 - d_{m-1}^2)}{2\,\Delta d\,\{(ln\, F_m - ln\, F_{m-1}) + (ln\, F_m - ln\, F_{m+1})\}} - \frac{(ln\, F_m - ln\, F_{m-1})(d_m^2 - d_{m+1}^2)}{2\,\Delta d\,\{(ln\, F_m - ln\, F_{m-1}) + (ln\, F_m - ln\, F_{m+1})\}} \qquad (17)$$

[4]Due to image noise and variations in magnification, the focus measure function may be multi-modal with one strong peak and one or more weak ones.

$$\sigma_F^2 = \frac{(d_m^2 - d_{m-1}^2) + (d_m^2 - d_{m+1}^2)}{2\,\{(ln\, F_m - ln\, F_{m-1}) + (ln\, F_m - ln\, F_{m+1})\}} \qquad (18)$$

Using eq.15, we can find F_p from σ_F and $\overline{d}$ as:

$$F_p = F_m \,/\, exp \left\{ -\frac{1}{2}\left(\frac{d_m-\overline{d}}{\sigma_F}\right)^2 \right\} \qquad (19)$$

If F_p is large and σ_F is small, the focus measure function has a strong peak, indicating high surface texture content in the vicinity of the image point (x, y). Hence, the values of F_p and σ_F can be used to segment the observed scene into regions of different texture content.

The following algorithm first finds the measures F_{m-1}, F_m, and F_{m+1} that correspond to the strongest peak of $F(d)$, and then uses these measures to estimate the depth $\overline{d}$ by Gaussian interpolation.

Algorithm 2

Step 1: Let $k = 3$, $F_{m-1} = 0$, $F_m = 0$, $F_{m+1} = 0$, $d_m = 0$.

Step 2: If $F_{k-1} > F_m$, $F_{k-1} > F_k$, and $F_{k-1} > F_{k-2}$, then:

$F_m = F_{k-1}$
$F_{m-1} = F_{k-2}$
$F_{m+1} = F_k$
$d_m = d_{k-1}$

Step 3: If $k < M$, $k = k + 1$, go to step 2.

Step 4: $d_{m-1} = d_m - \Delta d$ and $d_{m+1} = d_m + \Delta d$. Determine $\overline{d}$, σ_F, and F_p using Eqs. 17, 18, and 19.

Step 5: If $F_p < T_3$ or $\sigma_F > T_4$, the image point (x, y) belongs to background. Stop.

Since the values of F_p and σ_F are only useful for texture segmentation, their evaluation may be avoided to save computations.

8 An Automated System

8.1 Implementation

We have implemented a fully automated shape from focus system for the recovery of microscopic objects. A photograph of the system is shown in Fig.7. A Nikon Alphaphot-2 model microscope is used to image the objects. Objects can be magnified using objective lenses with ×10, ×40, and ×100 magnification. The object is illuminated using bright field illumination where light energy is focused on the object by the same lenses that are used to magnify

the object. A CCD camera with 512×512 pixels is mounted on the microscope to obtain digital images of the object. The z-axis of the microscope stage is driven by a stepper motor and the position of the stage can be computer controlled with a resolution and accuracy of 0.02 μm. The shape from focus algorithm is programmed and executed on a Sun SPARC 2 workstation.

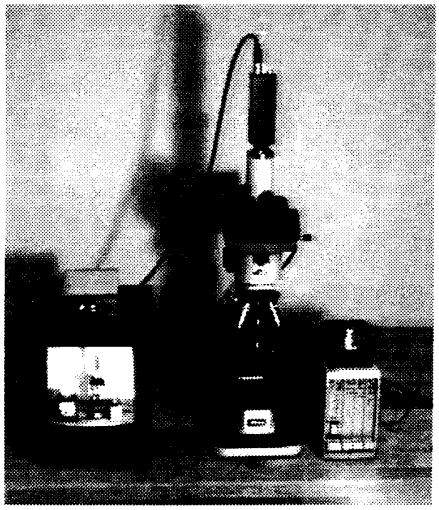


Figure 7: Shape from focus system.

The object is placed on the microscope stage and the appropriate objective lens is used to magnify the object. The focus measure parameters (T_1 and *step*) and the stage displacement (Δd) are provided to the program. The program then automatically increments the stage position, digitizes and stores an image for each new position, and uses the image sequence to compute a depth map of the object. The program also reconstructs a focused image of the object from the sequence of defocused images. The reconstruction algorithm uses the estimated depth to locate and patch together the best focused image areas in image sequence.

8.2 Results

Prior to automating the shape from focus system, experiments were conducted to determine the accuracy and feasibility of the method. The first experiment was conducted on a steel ball sample that is 1590μm in diameter. The ball has a rough surface that gives it a textured appearance. A camera image of the ball under bright field illumination is shown in Fig.8(a). Incremental displacements of $\Delta d = 100$ μm were used to obtain 12 images of the ball, and a 5×5 SML operator was applied to the image sequence to obtain focus measures. Depth maps of the ball, generated by the coarse resolution and Gaussian interpolation algorithms, are shown in Fig.8(b) and 8(c), respectively. The known size and location of the ball were used to obtain error maps by subtracting a smooth ball from the two depth maps. It is difficult to define the accuracy of the method as it depends on several factors: the surface texture, depth of field of the imaging system, and the incremental displacement Δd. The table shown in Fig.8(d) shows the error statistics computed from the error maps corresponding to the two algorithms. A total of 23235 image pixels lie within the boundary of the ball. The number of depth values computed by each algorithm depends on the values selected for the thresholds T_2, T_9, and T_4.

The automated system has been applied to a variety of biological samples. Fig. 9 shows the results obtained in the case of a mature embryo of a fruit fly (drosophila melanogaster). Since the embryo is globular in shape, no one image is perfectly focused; for any microscope stage position, only some sections of the embryo are perfectly focused while the remaining sections are defocused. A sequence of 20 images of the embryo was obtained by moving the microscope stage. Figures 9a-f show a few of these images. Using the image sequence the system produces two results: Fig.9(g) shows a reconstructed (focused) image of the fruit fly embryo and Fig. 9(h) shows two views of the depth map of the embryo. The shape-from-focus system is currently being applied to a variety of medical as well as industrial samples.

9 Summary

The shape from focus method presented in this paper uses a sequence of object images obtained by translating the object through the focused plane of an optical microscope. The sum-modified-Laplacian operator is applied to the image sequence to compute a set of focus measures at each point on the object surface. Using a model for the focus measure function, the focus measures at each point are interpolated to compute accurate depth estimates. The local nature of the depth estimation technique enables it to adapt to substantial variations in image texture.

A fully automated shape from focus system has been developed. The system has been used to recover dense and accurate depth maps and focused images of a variety of microscopic objects.

Acknowledgements

The authors would like to thank Yasuo Nakagawa of PERL, Hitachi, and Reg Willson of the VASC Group, CMU, for several discussions and valuable comments.

References

[1] B.K.P. Horn, *Focusing,* MIT Artificial Intelligence Laboratory, Memo No. 160, May, 1968.

[2] J.M. Tenenbaum, *Accommodation in Computer Vision,* Ph.D. Thesis, Stanford University, 1970.

[3] R.A. Jarvis, *Focus optimization criteria for computer image processing,* Microscope, Vol. 24, No. 2, pp. 163-180, 1976.

[4] J.F. Schlag, A.C. Sanderson, C.P. Neumann, F.C. Wimberly, *Implementation of automatic focusing algorithms for a computer vision system with camera control,* Carnegie Mellon University, CMU-RI-TR-83-14, August, 1983.

[5] E. Krotkov, *Focusing,* International Journal of Computer Vision, Vol. 1, pp. 223-237, 1987.

[6] A. Pentland, *A New Sense for Depth of Field,* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 9, No. 4, pp. 523-531, July 1987.

[7] P. Grossmann, *Depth from Focus,* Pattern Recognition Letters, Vol. 5, pp. 63-69, 1987.

[8] T. Darrell and K. Wohn, *Pyramid Based Depth from Focus,* Proc. CVPR, pp. 504-509, 1988.

[9] M. Subbarao, *Efficient Depth Recovery through Inverse Optics,* Machine Vision for Inspection and Measurement, edited by H. Freeman, Academic Press, 1989.

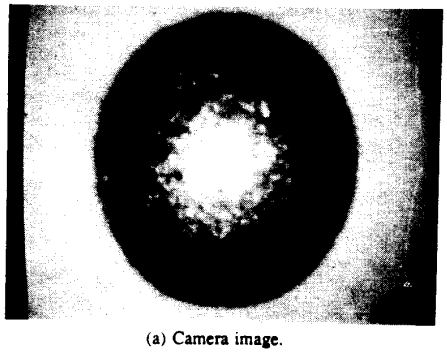
[10] S. K. Nayar, K. Ikeuchi, T. Kanade, "Surface Reflection: Physical and Geometrical Perspectives," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 611-634, July, 1991.

[11] M. Born and E. Wolf, *Principles of Optics,* London:Pergamon, 1965.

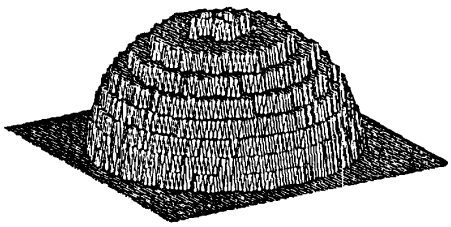
[12] B. K. P. Horn, *Robot Vision,* MIT Press, 1986.

[13] G. Westheimer, *The Eye,* Medical Physiology, Thirteenth Edition, Mosby Co., 1974.

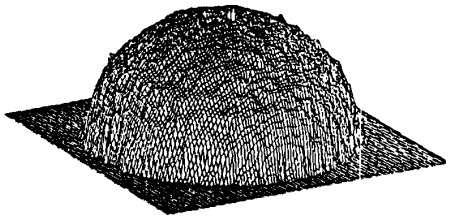
[14] R. G. Willson and S. A. Shafer, *Dynamic Lens Compensation for Active Color Imaging and Constant Magnification Focusing,* Carnegie Mellon University, CMU-RI-TR-91, October, 1991.



(a) Camera image.



(b) Depth map: coarse resolution.

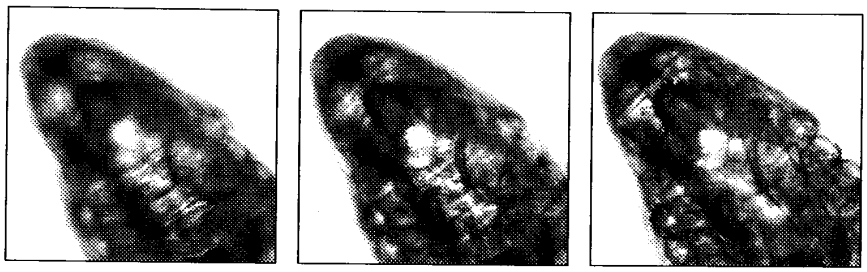


(c) Depth map: Gaussian interpolation.

Diameter of Test Sphere : 1590 μm		
	Coarse Interpolation	Gaussian Interpolation
Number of Points	22682	23257
Mean Error (μm)	7.861	3.857
Mean Absolute Error (μm)	30.32	13.815
Maximum Absolute Error (μm)	187.80	175.82

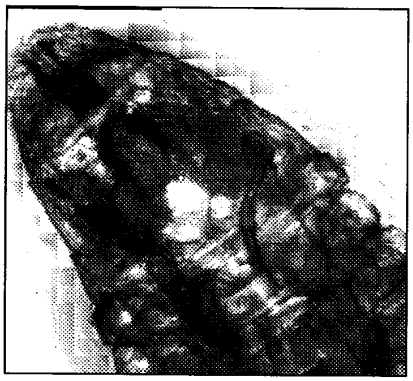
(d) Error statistics.

Figure 8: Error analysis using a spherical object.

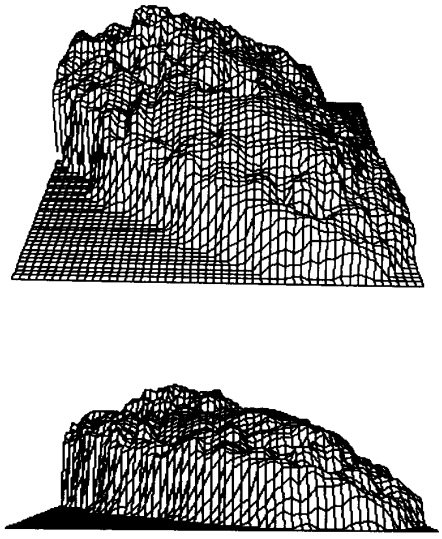


(a) i=2 (b) i=6 (c) i=10

(d) i=13 (e) i=15 (f) i=18



(g) Reconstructed Image



(h) Depth maps.

Figure 9: Results obtained for fruit fly (drosophila melanogaster) embryo.