JOURNAL OF COLLOID AND INTERFACE SCIENCE 171, 474–482 (1995)

High-Precision Automated Measurement of Surface Separation for the Surface Forces Apparatus

ROGER A. QUON, JOHN M. LEVINS, AND T. KYLE VANDERLICK[1]

University of Pennsylvania, Department of Chemical Engineering, 220 South 33rd Street, Philadelphia, Pennsylvania 19104-6393

Received July 14, 1994; accepted October 27, 1994

Multiple beam interferometry is a simple, yet powerful, technique that can be used to determine changes in the distance between two surfaces; under optimal conditions, the resolution of the measurement is less than 1 Å. Multiple beam interferometry is utilized effectively in conjunction with the surface forces apparatus to measure the forces acting between two opposed surfaces as a function of their separation. The conventional method of measurement is based on visual observations of interference fringes which shift continuously in wavelength as the distance between the surfaces is varied. Data collection is a slow and demanding process, and precision is limited by both the skill and the stamina of the experimenter. We report here on an automated system designed to measure the wavelengths of interference fringes produced in a typical surface forces experiment. The measurement precision obtained with the automated system exceeds that of the conventional manual method, while the measurement time required is comparable. The distinguishing feature of our system is the development and implementation of a versatile algorithm designed to quantify the shapes of the interference fringes. Capabilities of the automated system that are not afforded with the manual method include accurate curve fits to the observed fringes and simultaneous measurements of all fringes in the field of view.

Key Words: surface forces apparatus; multiple beam interferometry; automation; measurement.

1. INTRODUCTION

The surface forces apparatus (SFA) is now a well-known device that is used to measure the forces exerted between two opposed surfaces as a function of their separation. The design of the apparatus allows the surfaces to be moved relative to one another in increments ranging from micrometers down to angstroms (1, 2). The apparatus is ideally suited for investigations of both colloidal interactions and adhesion. The former concerns the variation with separation of the force acting between two surfaces immersed in a fluid medium. The latter concerns the forces between, and associated deformations of, two surfaces in contact.

Central to the force measurement technique is the use of multiple beam interferometry (MBI) to measure the distance between the surfaces. The fundamentals of MBI are discussed in detail by Tolansky (3). For experimental convenience, the surfaces employed in the force apparatus are curved; and so the gap between them is not uniform. When studying colloidal interactions, MBI is used to determine changes in separation between the surfaces, and the relative configuration of the surfaces about their point of closest approach. For investigations of adhesion, MBI is used to measure the deformation of the surfaces about the region of contact as a function of the applied load (4).

One of the most difficult and laborious tasks associated with any surface forces experiment is the measurement of the spatial positions of observed interference fringes—data used to determine the surface separation. In the standard measurement procedure, the experimenter views directly interference fringes at the exit port of a spectrometer, and tracks their locations manually by positioning a translating graticule line. The interference fringes are extremely dim and must be measured in pitch darkness. A typical experiment requires hundreds of measurements. Eye strain, as well as mental and physical fatigue, wear on the abilities of the experimenter to take precise measurements. Thus data collection is a slow and demanding process, and precision is limited by both the skill and the stamina of the experimenter.

While fringe locations can be measured in the cumbersome manner described above, no procedures have been developed to provide an accurate quantitative characterization of fringe shapes. Not only is this information the primary focus of adhesion investigations, it is of consequence to any force versus distance measurement. Estimates of the local radius of curvature about the point of closest approach are generally used to interpret force data; this is typically done according to the Derjaguin approximation (5, 6), which relates the forces acting between curved surfaces to the energy between flat surfaces. Recently, Parker and Attard showed that, as two surfaces approach, the local radius of curvature can change significantly due to the action of surface forces (7). A method to quantify these deformations would allow for more accurate measurements of the radius of curvature and thus more accurate interpretations of force data.

[1] To whom correspondence should be addressed.

0021-9797/95 $6.00

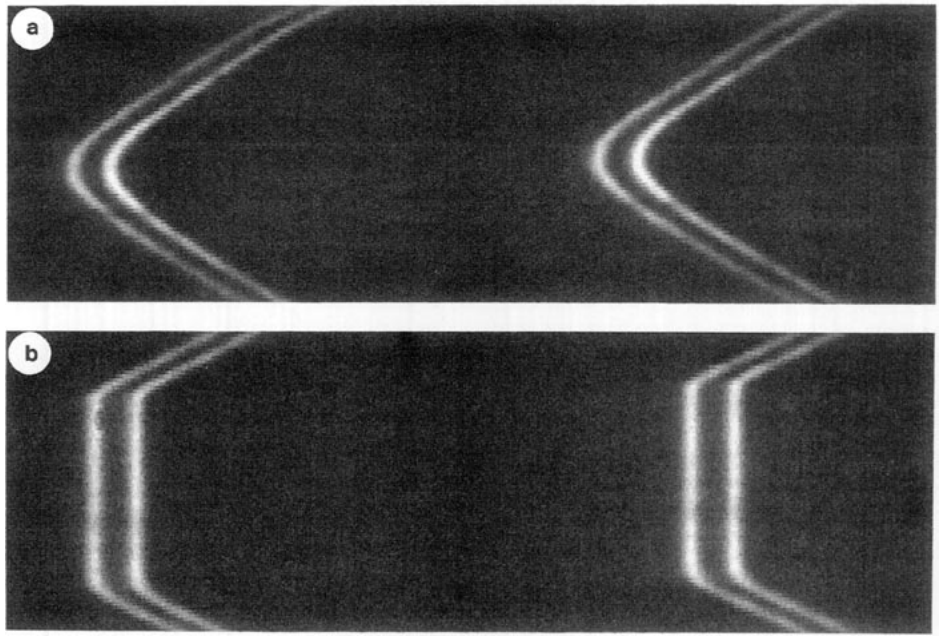

FIG. 1. (a) Image of mica–mica surfaces out of contact. (b) Image of mica–mica surfaces in contact.

The overall situation clearly necessitates an automated measurement and control system for the surface forces apparatus. Without doubt, the hub of such an automated system is the acquisition and analysis of interference fringes. Additional components beyond these—such as interfacing the motorized mechanisms used to position the surfaces—are straightforward. In 1989, Deitrick and co-workers at the University of Minnesota developed an automated system for the surface forces apparatus (8). Using this system as a platform, we have incorporated changes in both the hardware and software to produce a significantly improved fringe acquisition and analysis system. The distinguishing feature of our system is the development and implementation of a versatile algorithm designed to quantify the shapes of the interference fringes. In this paper, we describe our system, and evaluate its performance, both in terms of precision and speed, against manual measurements.

2. BACKGROUND

To better understand both the requirements and design of our fringe acquisition and analysis system, it is worthwhile to review the basics of multiple beam interferometry. The application of this optical technique requires simply the presence of two highly reflective thin films separated by one or more dielectric materials of total thickness greater than the wavelength of visible light. When white light is passed normally through such an interferometer, multiple reflections occur between the two reflective films and the light transmitted consists of a spectrum of intensity maxima known as fringes of equal chromatic order (FECO). This spectrum can be accurately predicted using classical electromagnetic theory (9–11).

In conjunction with the surface forces apparatus, multiple beam interferometry is generally implemented as follows. Two sheets of mica, each approximately 3 μm thick and 1 cm square, are prepared by depositing a highly reflective layer of silver (ca. 500 Å thick, 95% reflective) on one side only. The sheets are then glued to transparent, cylindrical supports with the silver side down. The supports are mounted in the apparatus so the bare mica surfaces are opposed to each other and the axes of curvature are at right angles. Other configurations and substrates have been employed in the apparatus (12–15) but as long as highly reflective silver films are used to promote the interference, the basic principle of fringe generation and measurement remains the same.

In an experiment, fringes are observed by passing the transmitted light through a spectrometer which disperses it, thus separating out in space the light according to wavelength. Calibration of the dispersion is accomplished by measuring the distance between spectral lines of known wavelength, such as those emitted from a mercury source. The spatial positions of the observed FECO change continuously as the separation between the surfaces is varied. Furthermore, since the separation between the surfaces is not uniform, the shape of the observed FECO is a direct representation of the relative geometry of the surfaces.

Figure 1 shows images of some typical fringes, corresponding to two different situations: surfaces which are separated, and surfaces in contact. In the former case, the fringe shape is curved, a manifestation of the cross-cylindrical ge-

ometry. In the latter case, the fringe tips are flat, reflecting the fact that the surfaces have deformed about the region of contact. The fringes in Fig. 1 appear as doublets due to the birefringence of mica. The degree of splitting of the observed fringes is controlled by the relative crystallographic orientation of the two mica sheets. Unresolved birefringence amounts to broad fringes, thus significantly downgrading the precision of the fringe location measurement. Since it is possible to guarantee resolved birefringence by being careful to orient properly the mica sheets as they are loaded into the force apparatus, we have developed our fringe analysis and acquisition package assuming this to be the case.

In the current "by eye" measurement procedure, the experimenter views the fringes through a traveling eyepiece mounted at the exit port of the spectrometer; the eyepiece contains one or more graticule lines. The location, and thereby wavelength, of any given fringe is tracked by positioning a graticule line at the perceived center of the tip of the fringe. A typical fringe—more specifically, one of the two resolved doublets—is approximately 200 μm thick. In order to resolve changes in surface separation to within 1–2 Å, shifts in fringe wavelengths must be measured to within one-tenth of an angstrom, which amounts to about 4 μm of graticule movement (this varies slightly with the thickness of mica used). The time required to make a precise measurement by this manual method is approximately 15–30 s. In addition to random error associated with this difficult procedure, the method relies on the operator's visual perception of the fringe's center and this presents a source of systematic error.

3. DESCRIPTION OF THE FRINGE MEASUREMENT SYSTEM

The automation of fringe measurement has three separate stages: (1) image acquisition, (2) fringe finding, and (3) fringe fitting. The result of the data acquisition stage is simply a digitized image of the interference fringes. The basic objective of the second stage is to identify fringes present in the captured image, and assign to each a data set consisting of the spatial coordinates of various points along the fringe. The last stage characterizes the shape of each fringe by fitting a curve through the data set. The location of the fringe tip is determined from the fringe shape profile. These three stages of the fringe measurement system are described here in detail.

3.1. Image Acquisition and Digitization

Images are acquired using a liquid-cooled slow-scan digital charged-coupled device (CCD) imaging system (Series 200, Photometrics Ltd., Tuscon, AZ). Central to the system is the scientific grade CCD detector which can be exposed for time intervals ranging from 0.1 to 1000 s. To reduce noise in the image due to thermal fluctuations, the detector is cooled to −45°C. Our detector consists of a 384 × 576 array of square pixels (25 μm per side). We note here that the system supports a variety of other CCD detectors; ones with smaller pixels can be used to increase resolution.

The camera is set at the exit port of the spectrometer, aligned so that the rows of the detector are oriented in the dispersion direction, and focused using the method described by Deitrick (8). Images are captured by opening the camera shutter and allowing light from the spectrometer to fall on the exposed CCD chip. Light incident on each pixel generates a charge proportional to the intensity and the exposure time. The signal is digitized, at 100 MHz, to a resolution of up to 16 bits of information. The CCD readout is given in arbitrary units which we denote "counts." Lastly, the array of pixel intensities is transferred to the host computer (we use a 486/25 MHz personal computer) through an IEEE-488 interface.

The exposure time is the rate-limiting step in the data acquisition process, since data transfer from the camera to the computer is essentially instantaneous. In our experience, we have found that a 3- to 10-s exposure time is adequate for most images.

We note here that an alternative to this imaging system is the use an intensified video camera with either a tube or a CCD detector. (Standard video cameras fail to detect the low intensity images.) Unfortunately, the intensifiers substantially degrade image contrast; moreover, the tube detectors exhibit appreciable geometric distortion. The signal-to-noise ratio obtained with such cameras (which operate at 30 frame/s, thus exposing the detector for no more than 0.03 s) is substantially lower than that obtained with the liquid-cooled, slow-scan system described above. As a result of image degradation and low resolution, the precision of the fringe measurement is reduced to about 8 μm, which is twice that of the manual method (8, 16). If, however, precision can be sacrificed in favor of speed of operation, these camera systems can be employed. Real-time images can then be recorded on video tape for later analysis. This approach has been used effectively to collect fringe data for dynamic force experiments (16, 17). In these experiments, however, the off-line measurement of fringe positions was carried out with a video micrometer, a procedure akin to the on-line manual method of measurement. If desired, one could digitize images stored on video tape, and then use the fringe analysis components of our system to find, and fit, the fringes.

3.2. Fringe Finding

Fringe finding involves identifying a set of points that describe the location of each fringe in the captured image. These points are distilled from the discretized image in a step-by-step procedure. Essentially, data are whittled to the point where fringes can be easily identified and detailed. To better understand the algorithm developed, it is helpful to consider first a representative image.

A typical image is shown in Fig. 2a. The shaded region is expanded below it (Fig. 2b) to illustrate its discrete nature;

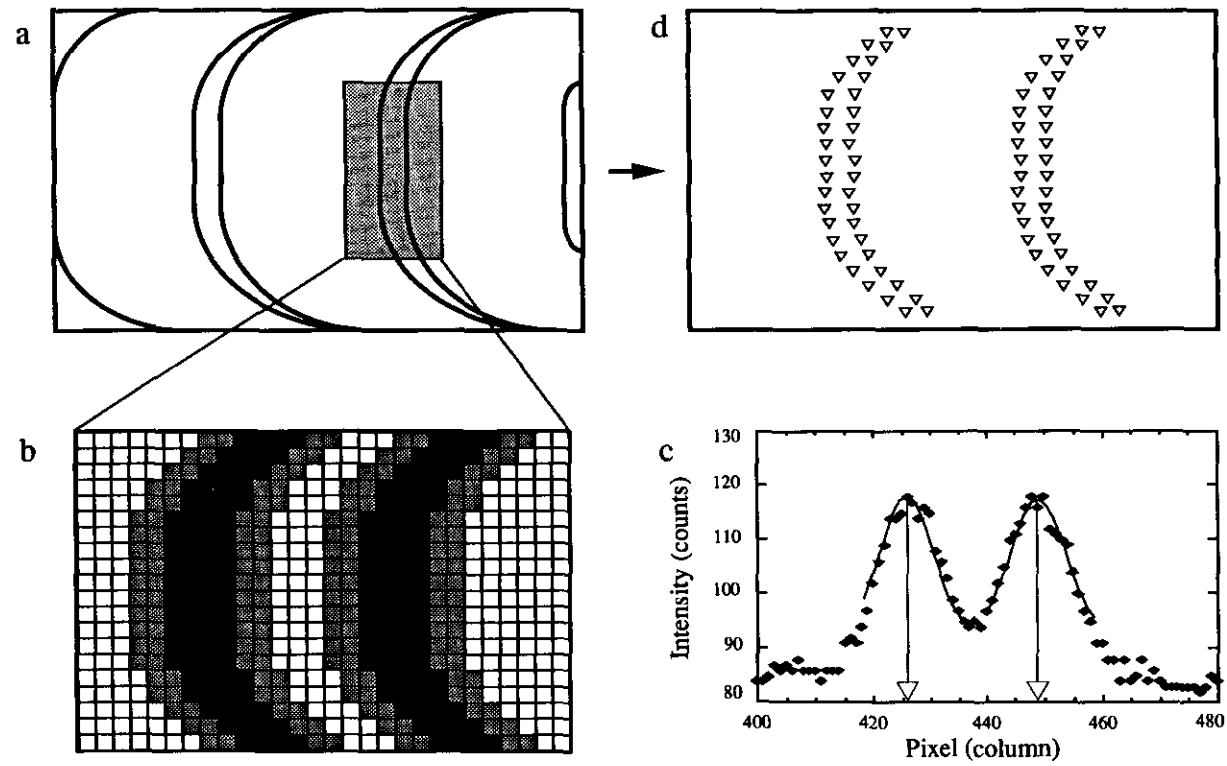

FIG. 2. (a) Representation of a typical digitized image. (b) Expanded view of the shaded region in (a). The grid represents the matrix of pixels that composes the digitized image. (c) Intensity profile along a portion of a row of an image. Filled diamonds represent the intensity at various column coordinates, the curves are a Gaussian fit of the peaked regions, and the open triangles are the means of the respective Gaussian fits. (d) Representation of the image after fringe finding. The open triangles represent points belonging to fringes seen in (a).

the underlying grid represents the pixels. Fringes (for convenience, we shall view a doublet as two separate fringes) comprise not one, but many pixels of varying intensity; darker squares have greater intensity. A realistic intensity spectrum along one row of the matrix is shown in Fig. 2c. Each peak yields one point (the open triangle) to the set that describes its associated fringe. Finally, note that the image in Fig. 2a contains nonessential, and problematic, features: tail remnants of a fringe cut by the left side of the field of view, the tip of a fringe at the right side of the field of view, and the merging of doublets in the tail region.

The overall objective of the fringe finding algorithm is to take the information shown in Fig. 2a and boil it down to the points shown in Fig. 2d. For each fringe, we refer to its collection of points simply as its data set. This data set serves as input for the fringe fitting stage.

The fringe finding process breaks into four steps—each refines information gleaned from the previous step. The first step, peak isolation, is effected through a thresholding technique. The second step removes data associated with fringes that are intersected by the left and right edges of the matrix. In the third step, peaks are assigned to fringes. Lastly, the data set for each fringe is constructed by extracting a point from each peak. These four steps are now described in more detail.

The first step, peak isolation, is based on a thresholding procedure whereby matrix values below a given count are set at zero. Thresholds are determined on a row-by-row basis. The applied threshold must serve to retain as much of the original peaks as possible while separating the twin peaks (i.e., the two fringes comprising a doublet) by a finite region of zero intensity. Thus, each resulting pixel span of positive intensity is associated with a distinct fringe.

The applied threshold is determined by an iterative scheme. The process begins with the threshold set to the highest intensity value in the row. On each pass, the number of spans of nonzero intensity (representing the number of isolated peaks) is counted. The threshold is successively lowered until the number of spans decreases; at this point, the threshold is not great enough to distinguish a twin peak. The next to last value defines the final threshold.

The second step comprises two parts. The purpose of the first part is to remove tail remnants on the left part of the matrix. This is accomplished by identifying the left-most complete fringe, i.e., one with a tip, then assigning to zero all pixels to the left of that fringe. A simple tracking procedure is employed to identify the pixels associated with the left-most complete fringe. This starts by scanning along the center row for the first nonzero pixel; this is the left-most pixel of the span associated with this fringe (it is assumed that the

fringes are reasonably centered in the image plane). From here, hop to an adjacent row, and search for the left-most pixel of the closest span; continue in this fashion until all rows have been examined.

The purpose of the second part is to remove emerging fringe tips on the right edge of the matrix. These tips are eliminated from analysis if they fail to encompass a user-determined number of rows. The range of the right-most fringe is easily found by starting at the center row of the right-most column of the matrix, and scanning above and below until nonzero elements are found; these upper and lower rows mark the bounds of the fringe. If the extent of the fringe does not meet the criterion, then the fringe is removed. We do this by setting to zero, for each row between the two bounds, the first positive span lying to the left of the right edge of the matrix. This step is repeated until no fringes cross the right edge, or fringes which do so meet the criterion.

By this point, the intensity matrix should have the following form. The only nonzero intensities correspond to fringes that have both a full tip and tail region. All peaks, i.e., contiguous spans of positive intensity, are separated by one or more pixels of zero intensity. Included in this matrix, however, may be broad peaks corresponding to unresolved birefringence in the tail extremities.

The third step of the algorithm determines the number of fringes and the size of their data sets. First, the approximate tip region is located; it is easily found by scanning down columns, beginning from the left edge, until a nonzero element is found. Next, the number of contiguous spans of positive intensity in that row are counted—this number is the same as the number of fringes in view. Using this row as a starting point, the number of spans are counted in the rows above and below, until the number of spans decreases; the origin of this decrease is unresolved birefringence in the tail extremities. For expedience, the data sets are constructed only from rows in which there is no unresolved birefringence. Within these rows, spans are easily assigned to fringes by order of their appearance.

The last step constructs the data set for each fringe. Each span generates one point. The vertical coordinate of a point is simply the height of the row of its span. One simple, but inaccurate, way of determining the horizontal coordinate of a point is to use the column coordinate of the pixel within the span having the maximum intensity. However, the method collapses if the peak is not pointed, or is significantly skewed. A more accurate method of determining the horizontal location is to fit the intensity data of the span to a Gaussian distribution. The mean of the Gaussian provides a subpixel estimation for the horizontal coordinate of the point. The solid curves in Fig. 2c are the Gaussian fits; the triangles indicate the location of the means of the fits.

The result of this procedure is a data set corresponding to points along a given fringe: $\{(x_i, y_i);\ i = 1, \ldots, r\}$, with the units of these coordinates in numbers of pixels. Here, y_i is a real number (derived from the mean of our Gaussian fit), and x_i is an integer (equal to a row number). Note that for ease of fitting, the assignment of the x and y axes is based on rotating the detector 90°.

3.3. *Fringe Fitting*

The objective of this stage is to generate a quantitative representation of the fringe shape profile by fitting a curve—henceforth referred to as a fringe curve—to each data set provided by the fringe finding algorithm. Due to the variety of fringe shapes that can be observed in practice, an analytical expression of the fringe curve does not suffice except in special cases. The most severe limitation of the automated system developed by Deitrick and co-workers (8) was a reliance on quadratic functions to approximate all fringe curves. The main advantage of our fringe characterization package is a unique fringe fitting algorithm developed to overcome this problem.

In our algorithm, we exploit the power of finite element analysis to predict accurately functions of arbitrary shape. In finite element analysis, the fringe curve is approximated by a finite number of specified basis functions, $\{\phi_i;\ i = 1, \ldots, n\}$,

$$y(x) = \sum_{i=1}^{n} a_i \phi_i(x), \qquad [1]$$

where the coefficients a_i are the unknowns to be determined. Basis functions used in the finite element method are zero except over small localized regions within the global domain. For this application, it is convenient to use Hermitian basis functions, which are constructed from third-order polynomials (18); curves approximated with Hermitian basis functions are smooth, having continuous zero- and first-order derivatives.

The key to our curve fitting algorithm is the choice of criterion used to determine goodness of fit. Because of both camera noise and the inherent pixel-wise discretization of the intensity information, the data set determined from the fringe finding analysis represents a collection of points which are scattered about the desired fringe curve. (Surface roughness can also contribute to the observed scatter in the data (2).) We henceforth refer to the inherent scatter in the data simply as "noise." The important point is that the predicted fringe curve need not, and should not, pass exactly through all the data points. Clearly, a curve which passes through every data point would do so at the expense of an increase in arc length. Furthermore, such a curve would have portions with large local radii of curvature, induced by forcing the curve to pass through data which is not inherently smooth.

To reduce the influence of noise on the predicted fringe curve, we employ a variational approach that accounts for the consequences of curve excursions caused by scatter in

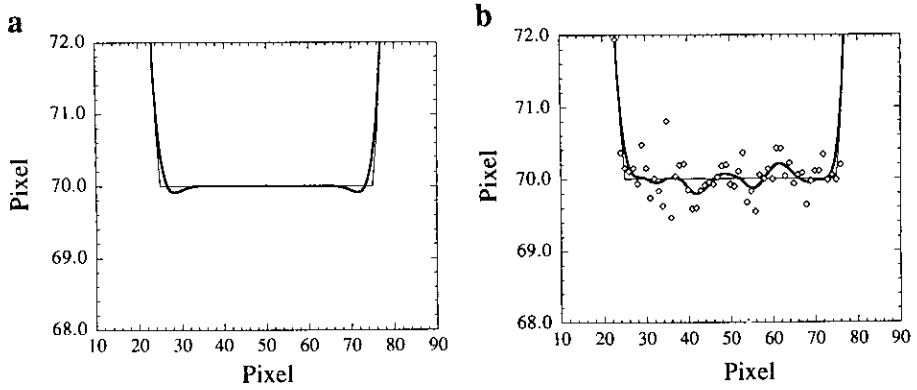

FIG. 3. (a) Comparison of Hertzian profile (light line) to the curve fit (heavy line). (b) Comparison of Hertzian profile (light line), Hertzian profile with noise (diamonds), and curve fit (heavy line).

the data set. We do this by defining the fringe curve to be that function, $y(x)$, which minimizes an objective functional, Ω, of the following form (where primes refer to derivatives with respect to x):

$$\Omega \equiv \sum_{m=1}^{r} (y(x_m) - y_m)^2 + \alpha \int_{x_1}^{x_r} \sqrt{1 + y'^2}\, dx + \beta \int_{x_1}^{x_r} \frac{y''^2}{(1 + y''^2)^{2.5}}\, dx. \quad [2]$$

This objective functional is composed of three pieces: (1) a least-squares component based on the lateral displacement of the curve from the governing data set; (2) a component proportional to the total length of the resultant curve; and (3) a component proportional to the integral of curvature squared (curvature alone is unsuitable since it can be of either sign, and its integral is constant for closed circuits). The last two pieces of this objective functional serve as filters that act to reduce the influence of scatter on the resultant curve fit. Their importance, relative to one another and to the least-squares contribution, are controlled by adjustable parameters α and β. The limits of all integrals are simply the lowest and highest rows numbers considered, thus ensuring that the predicted curve spans the vertical space defined by the governing data set.

The predicted fringe curve is then found by substituting its finite element approximation, as given by Eq. [1], into Eq. [2], and then determining the set of coefficients which minimize Ω. The unknown coefficients solve the following system of nonlinear equations:

$$0 = 2 \sum_{m=1}^{r} (y(x_m) - y_m)\phi_i(x_m) + \alpha \int_{x_1}^{x_r} \frac{\phi_i}{\sqrt{1 + y'^2}}\, dx + \beta \int_{x_1}^{x_r} \left\{ \frac{2\phi_i''}{(1 + y'^2)^{2.5}} - \frac{5\phi_i' y''^2}{(1 + y'^2)^{3.5}} \right\} dx. \quad [3]$$

We use Newton's method to solve Eq. [3]; Gaussian quadrature is employed to evaluate all integrals. Details are provided in Appendix A.

From the predicted fringe curve, the fringe tip location is given by the minimum value of $y(x)$. We estimate this value by generating a series of values derived from this function, and searching among these for the minimum.

The computational cost of the curve fitting procedure is controlled by the number of basis functions, n, used to approximate the fringe curve. This number is independent of the number of points in the data set upon which the fit is based, r. The number of basis functions, however, should be large enough so that further increase in the discretization does not change the predicted curve to within a given tolerance, usually set by numerical round-off error.

4. APPRAISAL OF SYSTEM PERFORMANCE

The appraisal of system performance is done in two parts, presented in Sections 4.1 and 4.2, respectively. First, we consider the ability of the system to predict fringe shape profiles. Our aim is to demonstrate the effectiveness of the curve fitting algorithm described in Section 3.3; we also discuss the optimal settings of the filter parameters used in the fitting procedure. In the second section, we appraise the performance of the automated system in measuring fringe tip locations, and compare its precision, accuracy, and speed of operation against the current manual method of measurement.

4.1. Fringe Shapes

We first demonstrate the effectiveness of our curve fitting algorithm in terms of its ability to predict a desired fringe profile from two simulated data sets—one differing from the other solely by the addition of random noise. In particular, we consider the idealized fringe profile associated with two smooth elastomeric spheres in contact. Using the relationship derived by Hertz to describe this relative geometry (4, 19),

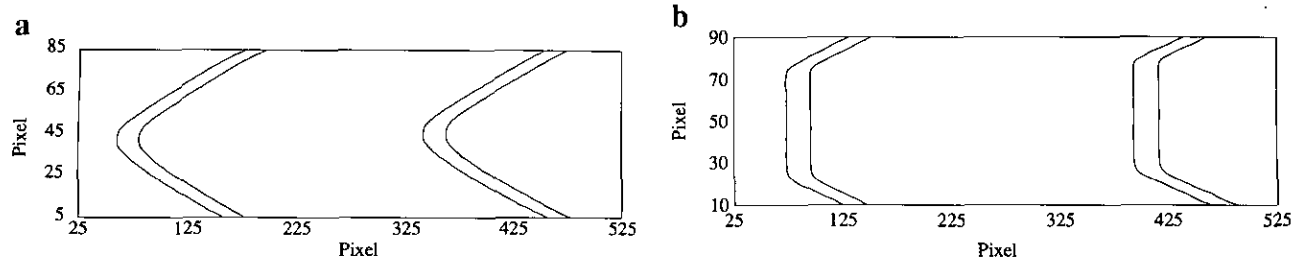

FIG. 4. (a) Curve fit of image shown Fig. 1a. (b) Curve fit of image shown in Fig. 1b.

we generated a discrete set of data points, shown in Fig. 3a, which lie along the fringe profile. The second set of data, shown in Fig. 3b, was generated by adding error to the first data set. More specifically, each data point from the first set was taken to be the mean of a Gaussian distribution having a standard deviation of .1 pixel. The magnitude of the standard deviation is on the order of the error associated with the manual measurement method (as will be discussed later). These two data sets were then used as inputs to our curve fitting algorithm.

One motivation for this study is to show that our algorithm is not only successful at predicting the underlying profile used to generate the data, but also that the curve predicted through the noisy data resembles closely that predicted through the noise-free data. In other words, a properly designed algorithm should be capable of "seeing" the desired fringe curve from within a set of noisy data. Since the influence of noise on the curve-fitting procedure is controlled by the two filters built into the objective functional, this study serves as a good means for choosing the magnitudes of the filter parameters, α and β. After exploring a large range of values, we found that $\alpha = 1.0$ and $\beta = 0.1$ were effective choices, yielding the fringe curves shown in Figs. 3a and 3b. Higher values of either weighting factor cause the predicted curve to pull away from the sharp breakaway regions where the flat contact zone ends; lower values allow the presence of noise to significantly affect the shape of the predicted curve.

The noise filters used in our curve fitting algorithm cannot distinguish sharp spatial gradients in the data induced from noise from those that may be truly associated with the underlying fringe profile. Furthermore, our algorithm is capable only of predicting curves with continuous slope, since our finite element approximation employs Hermitian basis functions. As a consequence of these two features, our algorithm is not able to predict exactly the sharp breakaway of the ideal Hertzian profile; close inspection of Fig. 3a shows that the predicted curve displays a small swoop about this region. We note here, however, that our basic algorithm could be modified to capture this type of behavior, if desired. For example, the use of linear basis functions would permit curves with discontinuous slope to be predicted. Other alternatives include using a weighted least-squares scheme to force the curve closer to data points at or near regions with sharp natural gradients, such as the breakaway zones.

Next we demonstrate the versatility of the complete automation package by employing it to find, and then fit, the fringes which appear in the experimentally observed images displayed in Fig. 1. These images spanned 80 pixel rows on the CCD chip and yielded as many data points for each fringe. The predicting fringe profiles are shown in Fig. 3. As is readily seen, the package is successful at fitting curves through fringes of distinctly different shapes—a primary goal which underpinned the development of our system. Moreover, convergence to each predicted profile is achieved with only thirty-two basis functions (the same number of basis functions was also used to predict the curves shown in Fig. 4).

4.2. *Fringe Tip Locations*

We now quantify the performance of our automation package in measuring fringe tip positions—one of its most important tasks—and compare results against the manual method of measurement. For the latter, we considered both an expert experimenter with five years of experience in the technique and a novice with only six months of experience. All findings are reported in terms of wavelengths, which are determined by comparing the position of the fringe to that of mercury spectral lines also in the field of view (for reference, in our system 1 pixel $\approx$ 0.76 Å).

First the reproducibility, or precision, of the automated measurement procedure is established. To evaluate this, we considered first the highest quality images realized in practice: these are generated when two back-silvered mica sheets are placed in adhesive contact (a representative image is shown in Fig. 1b). The tips of such fringes are flat; moreover, the fringes are immobile. This is not always the case with curved fringes which are generated when the surfaces are separated; vibrations of one surface relative to others, as well as thermal drift, can cause fringe motion and hence add significant uncertainty to the measurement. (In addition, human contact with the spectrometer, especially the turning of the translating micrometer, can promote further vibrations.) We therefore chose the measurement of contact fringes to establish the

upper boundary of the measurement precision of the automated and manual methods.

Measurements were made of three different contact fringes in the field of view; since each fringe appeared as a doublet (due to mica birefringence), six distinct tip positions could be measured. Each tip position was measured six times by the expert experimenter, and eight times by the automated system, thus allowing a standard deviation to be determined. The results are given in Table 1. The precision of the automated system is very consistent and is as good as that attained with the manual method, as manifested by the size of the standard deviations. A separate, less extensive, set of measurements were taken by the novice experimenter; his precision was approximately half that of the expert.

The results reported in Table 1 are an upper boundary to measurement precision. Under less ideal conditions (e.g., less sharp, or curved fringes) the typical precision of the expert operator is about 0.15 Å. Over the past year, we pitted the automated system against the expert experimenter under a variety of operating conditions, and we consistently found the system's precision to be as good as, or better than, that possible with the manual method.

It is important to appreciate that generally there is more than one fringe present in the field of view. The ultimate precision of the measurement of surface separation can be improved by measuring the wavelengths of as many fringes as possible. The time required using the manual method scales proportionately with the number of fringes. The automated method scales nonlinearly, requiring less time per fringe as the number of fringes increases. This is primarily because the time invested to capture the image need not be respent in analyzing additional fringes. Furthermore, the predicted shape of the first fringe curve can be used as an initial guess when using Newton's method to predict the other fringe curves. If there is unwanted fringe motion, the benefit of additional fringe analysis with the manual method is completely lost; while the automated system is even more valuable in such situations.

It is critical that the measurement procedure be both precise and accurate. Having established the former, we now confirm the accuracy of the fringe tip measurement by comparing absolute tip locations obtained by both the automated and manual methods. For this, we do not restrict the measurements to a single set of fringes within one image. Ten separate experiments were considered, yielding forty-six different contact fringes for measurement. The fringe quality was quite varied. The results are reported in terms of the absolute differences between the tip wavelength determined by the automated method and that determined by the manual method: the smallest difference was 0.00 Å, the largest difference 0.66 Å, and the average of all differences 0.23 Å. Clearly, the average difference is very small and on the order of the best precision obtainable by manual measurement. What this test also confirms is that the approach used within our system to determine the fringe's center is in good agreement with the operator's perception of the same.

TABLE 1
Comparison of Automated and Manual Fringe Wavelength Measurements

	Automated method	Manual method
Largest standard deviation of λ	0.04 Å	0.15 Å
Smallest standard deviation of λ	0.02 Å	0.07 Å
Average standard deviation of λ	0.03 Å	0.09 Å

The results reported above were all based on using the recommended values of the filter parameters, α and β. Although variation of these filter parameters affects the global shape of the fringe curve, their variation does not significantly affect either the accuracy or the precision of the fringe tip measurement.

Having established the accuracy and precision of the automated method, the last important performance index is the speed of operation. To find, and fit, the four curves shown in Fig. 4b, our personal computer (486/25 MHz processor) required just over a minute. The vast majority of the computation is associated with the fitting section of the algorithm. Given that it takes a seasoned experimenter about 20 s to measure one fringe tip accurately, the speed of the automated system is on par with that of the manual method. It is important to note, and appreciate, that a number of steps can be taken to speed up the automated process: employ faster processors, fit a limited region about the tip, and reduce the number of basis functions used. An alternative strategy is to save the data sets on-line and fit them later, off-line.

CONCLUSIONS

Our long-term goal is to integrate this fringe measurement system with hardware and software designed to control the various motors that drive the SFA experiments, leading eventually to a complete automation and control system. This control package will interact with the already existing measurement routines, and will be able to make decisions about motor movements based on fringe measurements being made concurrently. The success of the fringe measurement system seems to indicate that such a package is quite feasible, and well within reach in the near future.

The advantages of the fringe acquisition and analysis package reported here extend well beyond the original intentions. Recently, we introduced a new technique (11), extended spectral analysis of multiple beam interferometry (ESA-MBI), which significantly expands the variety of surfaces that can be investigated in the surface force apparatus. Any interference-producing system can be used, not just those which are based on highly reflective silver layers. The

transmitted fringes from virtually all other interference systems are significantly broader and dimmer, rendering the manual method of measurement essentially useless as a means to measure changes in surface separation. The underlying principle of the ESA-MBI is to capture the intensity versus wavelength spectrum of light transmitted through the surfaces; the distance between them is then determined by matching the measured spectra with that predicted from the theory of wave propagation through a stratified medium. The slow-scan camera system employed in the automated system described here is, in fact, the same detection device used to capture transmitted light in ESA-MBI. The same fringe acquisition and analysis package could be readily used to find and fit fringes produced in these experiments.

APPENDIX A

Newton's method is employed to solve the residual equations given by Eq. [3]. These residuals can be expressed as

$$R_i = R_i^{\mathrm{LS}} + R_i^{\alpha} + R_i^{\beta} = 0 \quad i = 1, \ldots, n,$$

where

$$R_i^{\mathrm{LS}} = 2 \sum_{m=1}^{r} (y(x_m) - y_m)\phi_i(x_m)$$

$$R_i^{\alpha} = \alpha \int_{x_1}^{x_r} \frac{\phi_i}{\sqrt{1 + y'^2}}\, dx$$

$$R_i^{\beta} = \beta \int_{x_1}^{x_r} \left\{ \frac{2\phi_i''}{(1 + y'^2)^{2.5}} - \frac{5\phi_i' y''^2}{(1 + y'^2)^{3.5}} \right\} dx.$$

If the coefficient, $\mathbf{a}$, is defined as $[a_1, \ldots, a_n]$, then Newton's method (18) is

$$\mathbf{a}^{k+1} = \mathbf{a}^k - \mathbf{J}^{-1}(\mathbf{a}^k) \cdot \mathbf{R}(\mathbf{a}^k).$$

Here, $\mathbf{J}^{-1}(\mathbf{a})$ is the inverse Jacobian of the system evaluated (20) at the kth iteration. Our Jacobian is defined as

$$J_{ij} = J_{ij}^{\mathrm{LS}} + J_{ij}^{\alpha} + J_{ij}^{\beta},$$

where

$$J_{ij}^{\mathrm{LS}} = 2 \sum_{m=1}^{r} \phi_i(x_m)\phi_j(x_m)$$

$$J_{ij}^{\alpha} = \alpha \int_{x_1}^{x_r} \frac{y'\phi_i}{\sqrt{1 + y'^2}}\, dx$$

$$J_{ij}^{\beta} = \beta \int_{x_1}^{x_r} \left\{ \left(\frac{5\phi_i'\phi_j' y'^2}{(1 + y'^2)^{3.5}} \right) \left(\frac{7y'^2}{1 + y'^2} - 1 \right) - \left(\frac{10y'y''(\phi_i'\phi_j'' + \phi_j'\phi_i'')}{(1 + y'^2)^{3.5}} \right) + \left(\frac{2\phi_i''\phi_j''}{(1 + y'^2)^{2.5}} \right) \right\} dx.$$

In this work, integration was performed on a local element basis using three-point Gaussian quadrature (20). The convergence criteria were

$$\|\mathbf{R}(\mathbf{a}^{(k)})\| < 10^{-6}, \quad \|\mathbf{a}^{(k+1)} - \mathbf{a}^{(k)}\| < 10^{-7}.$$

The norm, $\|\cdot\|$, is simply the root mean square. The first criterion sets the tolerance to which $\mathbf{R} \approx 0$ is satisfied by Newton's method, and the second criterion tests the sensitivity of the approximate solution to the next iteration.

ACKNOWLEDGMENTS

We gratefully acknowledge the support for this work provided by the David and Lucile Packard Foundation, the National Science Foundation through Engineering Research Equipment Grant CTS-9112964, and the MRL Program under Grant DMR91-20668. We are indebted to H. T. Davis and L. E. Scriven of the University of Minnesota for allowing us to borrow, and work with, their surface forces automated system. We owe special thanks to G. L. Deitrick for his invaluable and boundless assistance. We acknowledge the participation of Brian Motherway in the initial stages of this work. Finally, we are grateful to Jeremy Blumenfeld for lending his support and computer expertise.

REFERENCES

1. Israelachvili, J. N., and Adams, G. E., *J. Chem. Soc. Faraday Trans. 1* **74,** 975 (1978).
2. Levins, J. M., and Vanderlick, T. K., *J. Phys. Chem.* **96,** 10405 (1992).
3. Tolansky, S., "Multiple-Beam Interferometry of Surfaces and Films." Oxford Univ. Press, London, 1948.
4. Horn, R. G., Israelachvili, J. N., and Pribac, F., *J. Colloid Interface Sci.* **115,** 480 (1987).
5. Derjaguin, B. V., *Kolloid Z.* **69,** 155 (1934).
6. White, L. R., *J. Colloid Interface Sci.* **95,** 286 (1983).
7. Parker, J. L., and Attard, P., *J. Phys. Chem.* **96,** 10398 (1992).
8. Deitrick, G. L., Ph.D. thesis, University of Minnesota, 1990.
9. Clarkson, M. T., *J. Phys. D* **22,** 475 (1989).
10. Born, M., and Wolf, E., "Principles of Optics," 6th ed. Pergamon, Oxford, 1980.
11. Levins, J. M., and Vanderlick, T. K., *Langmuir* **10,** 2389 (1994).
12. Horn, R. G., Clarke, D. R., and Clarkson, M. T., *J. Mater. Res.* **3,** 413 (1988).
13. Horn, R. G., Smith, D. T., and Haller, W., *Chem. Phys. Lett.* **162,** 404 (1989).
14. Grabbe, A., and Horn, R. G., *J. Colloid Interface Sci.* **157,** 375 (1993).
15. Merrill, W. W., Pocius, A. V., Thakker, B. V., and Tirrell, M., *Langmuir* **7,** 1975 (1991).
16. Chan, D. Y. C., and Horn, R. G., *J. Chem. Phys.* **83,** 5311 (1985).
17. Israelachvili, J. N., *J. Colloid Interface Sci.* **110,** 263 (1986).
18. Strang, G., and Fix, G. J., "An Analysis of the Finite Element Method." Prentice Hall, Englewood Cliffs, NJ, 1973.
19. Hertz, H. J., *Reine Angew. Math.* **92,** 156 (1881).
20. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., "Numerical Recipes in C." Cambridge Univ. Press, Cambridge, 1988.