Fast Detection Based on Semi-Transient Signals in AFM

Peng Huang and Sean B. Andersson
Department of Mechanical Engineering and Division of Systems Engineering
Boston University, MA 02215

Abstract—In this paper, a high-speed width detector is derived for use in atomic force microscopy (AFM). The algorithm rapidly determines the width by detecting the two edges of the sample during a fast scan. The algorithm is designed primarily as a first step towards detection of a single macromolecule moving on a biopolymer. This detector is an important component of a new approach in AFM, a control system that directly tracks the motion of these single macromolecules rather than deriving their motion from a sequence of images. Such an approach promises a much higher temporal resolution than is achievable in time-lapse imaging.

I. INTRODUCTION

The ability to follow the dynamics of single macromolecules allows researchers to study molecular processes, increasing our understanding of molecular biology and helping to elucidate a variety of genetic diseases [1], [2]. Several techniques have been developed, including single particle tracking in fluorescence microscopy [3] and imaging with both special optical devices [4] and with scanning probe systems such as the atomic force microscope (AFM) [5] [6]. These tools and techniques have different strengths and weaknesses. For example, optical approaches allow for the study of molecules moving inside living cells but typically require modification of the target through fluorescence labeling and offer limited spatial resolution.

Among these tools, AFM provides a unique set of capabilities, including the ability to observe systems in their physiological environment, a resolution on the order of nanometers or better, and the ability to measure mechanical properties directly. Therefore, it has been used extensively in exploring the dynamics of biological systems, such as directional transport by protein motors [7], interactions between proteins [8] [9] and other cellular behaviors [10].

The primary drawback of AFM with respect to its application to dynamic processes is its slow imaging rate, with commercial instruments typically taking seconds to minutes to acquire a single frame. As a result, there is a great deal of ongoing work on increasing this rate. Approaches include improvements of mechanical components [11], advanced controller designs for piezo actuators [12] [13] as well as changing the conventional scanning pattern [14]. As a result, systems with near video-rate speed have been developed [15].

For some systems, however, even video-rate is not fast enough. For example, the dynein motor has been reported to move at speeds of up to 1.7 μm per second. Even at 30 frames per second, the motor would take six steps between the start and end of imaging a single frame. To study such high speed systems, an alternative approach is needed.

Such an approach begins with the recognition that individual measurements in AFM are very fast, with cantilever resonant frequencies in the tens to hundreds of kHz. The imaging rate is limited by the fact that as the cantilever moves across the sample, the controller must maintain the system in steady state, rejecting the disturbance represented by the sample. When considering dynamics of single macromolecules, however, it is the motion, not the structure, that is of primary importance. Approaching the problem from the point of view of detection and tracking rather than imaging allows us to avoid the issues related to imaging.

The general method of tracking features or single particles in scanning probe microscopy was first developed and applied in Scanning Tunneling Microscopy (STM) through the use of a small scan to keep the tip of the microscope positioned over a single atom [16]. A similar technique was used in [17] to account for thermal drift in STM. The “protein sandwich” method of tracking the motion of a protein held under the tip of an AFM cantilever was explored in [18].

There have also been efforts in AFM at using high-speed detection to infer the presence of samples from high-speed scans. In particular, in [19] a scheme for tapping mode AFM was developed in which the edges of the sample were detected by processing transient signals. The approach provides a significant increase in the scan rate. However, it can be difficult to discern a rising edge (moving on to a sample) from a falling edge (moving off of a sample), making it hard to estimate the width of a sample in real time from the data.

Our target application is the direct tracking of a single macromolecule moving on a biopolymer. Examples include protein motors walking in the plane parallel to the substrate or carrying a large cargo [20] and the motion of RNA polymerase during DNA transcription [21]. Such systems involve features that prevent the use of the previous methods while at the same time providing structures that can be taken advantage of to create a new tracking scheme. In particular, the presence of such a molecule on its track results in a change in the apparent width of the track. Inspired by the high speed detector of [19], [22], the current paper develops
a high level detection scheme that utilizes the transient signal to measure the width of a sample in real time. Accordingly the presence and location of the moving macromolecule is indicated and the contour of its track is described. This detector is integrated with the local raster-scan algorithm developed by the authors [14]. This algorithm uses the measurements of the system in real time to steer the tip to move along the track defined by the biopolymer. The combined detector and local raster scan algorithm allow the tip to be moved rapidly along a biopolymer and the width value to be estimated in real time.

II. OVERVIEW OF THE DETECTION SCHEME

The detection scheme begins with an AFM cantilever operating in tapping mode and being driven repeatedly across a biopolymer according to the local raster scan algorithm [14]. The scan pattern is illustrated in Fig. 1 in which the tip trajectory is shown as a sequence of dots representing the measurement points. Measurements on the underlying sample are highlighted in yellow. Our detection scheme relies on monitoring the difference between the actual cantilever motion and that of a reference model.

![Fig. 1. Illustration of the tip and cantilever motion. The local raster-scan algorithm drives the tip in a sinusoidal pattern (dots) over the sample. Measurements on the sample itself are highlighted in yellow, those on the substrate are in black. As the tip scans, the cantilever transitions from steady state at A, to a transient state caused by the interaction with the sample at B, and back to a steady state at C.](image)

Suppose that AFM is operated with a sufficient tip speed and the width of the sample is thin enough such that the z-controller cannot settle into its steady state while the tip is crossing the sample. Such a tip speed would be detrimental to the imaging application since it would essentially be impossible to produce a credible height value. For detection, however, a fast tip speed has benefits. The duration of the transient process contains information about the width of the underlying sample. In addition, changes of the width of the sample allow the detection of the moving macromolecule as stated in Sec. I. Unlike an imaging application, our tracking scheme requires that the cantilever is held in its transient state while the tip is on top of the sample and relies on the information extracted from this transient process.

To describe the algorithm, consider the segment of the tip trajectory from points A to C in Fig. 1. At A and C, the cantilever (illustrated in green) is tapping on the substrate and is assumed to be in steady state. It is this steady state that forms the reference model for comparison. As the tip proceeds to B, the transition up onto the sample is a disturbance that drives the cantilever into its transient state. Assuming the z-controller is slow relative to the tip speed across the sample, the cantilever will remain in the transient state until some time after moving back to the substrate. As the tip continues over the substrate, the controller has time to act, returning the cantilever to its steady state upon reaching C. The detection scheme uses the difference between the reference (steady state) dynamics and the true dynamics to estimate when the tip is on the sample, backing out the width from the position of the tip during that time.

A block diagram of the scheme is given in Fig. 2. In the next section, we walk around this diagram and describe each component.

![Fig. 2. A block diagram of the detector. Amp is the amplitude output of AFM while Amp_m denotes the amplitude output of the reference model.](image)

III. THE WIDTH DETECTOR: A WALK AROUND THE LOOP

As discussed in Sec. II and illustrated in Fig. 2, the detector uses the difference between a reference model and the true cantilever dynamics to detect the edges of the sample. Those edges are then used to calculate the width. This width is used in two ways; to drive the local raster-scan algorithm and to produce a "binary" image. Below we describe each element in the block diagram.

A. Model Reference

The dynamics of the cantilever have been shown to be well modeled by a second-order system with nonlinear feedback to capture the tip sample interaction [23]. These dynamics are given by

\[
y + \frac{\omega_0}{Q} \dot{y} + \omega_0^2 y = \frac{Q}{K} F + \omega_0^2 b,
\]

where \( y \) is the tip displacement, \( K \) is the spring constant, \( \omega_0 \) is the resonant frequency, \( Q \) is the quality factor, \( b \) is a sinusoidal excitation and \( F \) is the tip-sample interaction force.

Several models for the interaction force have been developed (e.g. [24], [25]). Here we utilize the Derjaguin-Muller-Toporov (DMT) model of [24]. This model captures the interaction in tapping mode as well as the convolution effect due to radius of the tip [26]. The DMT model expresses the force as

\[
F_{Is}(z) = \begin{cases} 
-\frac{HR}{6(z_s - z)^2} & z_s - z \geq a_0, \\
-\frac{HR}{6a_0^2 + \frac{4}{3}E^* \sqrt{R}(z_s - z + a_0)^{3/2}} & z_s - z < a_0,
\end{cases}
\]

where \( z_s \) denotes the distance between the sample and tip of the undeflected cantilever, \( z \) is the tip deflection towards the sample, \( H \) is the Hamarker constant, \( R \) the tip radius,
$E^*$ is the effective contact stiffness and $a_0$ is the interatomic distance.

As illustrated in Fig. 2, the reference model and the physical cantilever are driven by the same sinusoidal drive to synchronize their dynamics in the steady state. The reference model, however, receives no input from the physical sample and represents the steady-state dynamics on the substrate. Note that while the form of the steady state solution is well-known (see (3) below), the complexity of the nonlinear tip-sample interaction prevents one from writing an analytical expression for the parameters in that solution; the reference model thus provides a system-specific solution (up to the accuracy of the model and its parameters).

**B. Edge detection**

Since the reference model scans only over the substrate, its dynamics remain in the steady state, given by

$$ y_m(t) = A_0 \sin(\omega_0 t + \phi_0), $$

where $y_m(t)$ denotes the output of the model. The amplitude $A_0$ and phase $\phi_0$ are constant under the assumption that the nominal tip-sample distance is constant [27]. In practice this condition is enforced by the $z$-direction controller.

Upon transitioning from the substrate onto the sample, the physical cantilever is driven into its transient state. These dynamics can be described by

$$ y(t) = A(F(t)) \sin(\omega_0 t + \phi(F(t))) + \eta(t), $$

where $y(t)$ is the output of the physical system, $F(t)$ is the tip-sample force and $\eta(t)$ is the sensor noise. After moving back to the substrate and after sufficient time has passed, the $z$-controller will drive the physical cantilever back to the steady state described by (3). The detection problem is then one of determining when the system is in its transient state. Define the error $\delta y(t)$ as

$$ \delta y(t) = F((y(t)) - (y_m(t))) = \Phi(t) + \gamma(t), $$

where $F(\cdot)$ is an averaging filter, $\Phi(t)$ is the dynamic part of $\delta y(t)$ and $\gamma(t)$ accounts for the measurement noise. Note that the filter is performed on the absolute value of the residue between $y(t)$ and $y_m(t)$ over a period of half of the cycle of the cantilever. As a result, the oscillation of the residue is accounted for with the filter determining the envelope of the signal. It is this envelope that is used in the analysis below.

Due to noise, it is reasonable to consider a moving window of data for detection rather than a single measurement. We therefore define the vector $\vec{Y}$ by

$$ \vec{Y} = [\delta y(k-M+1), \delta y(k-M+2), \ldots, \delta y(k)]^T, $$

where $M$ is the size of the moving window. (See Sec. IV for an example of choosing $M$). Note that this signal $\vec{Y}$ has two possible states; either all its elements are sampled from the steady state of the cantilever or some of them are not. Accordingly, the binary hypothesis testing [28] is utilized here to determine the most likely case based on the data.

The two hypotheses are

$$ H_0 : \vec{Y}(k;M) = \vec{\gamma}(k;M), $$

$$ H_1 : \vec{Y}(k;M) = \Phi(k;M) + \vec{\gamma}(k;M). $$

Here $H_0$ is the hypothesis that the cantilever is in its steady state and thus the error signal is all noise while $H_1$ is the case where some transient dynamics appear. Due to the complexity of $\Phi(k;M)$, we approximate it as a box function with parameters selected by experience.

We use the maximum likelihood approach to decide between the hypotheses based on the data, and assume Gaussian statistics on the noise. We define a threshold test on the probabilities of the two hypotheses as

$$ P(\vec{Y}|H_1) = \frac{1}{\sqrt{2\pi M |\Sigma|}} e^{-\frac{1}{2}(\vec{\gamma}(k;M) - \vec{\Phi}(k;M))^T \Sigma^{-1}(\vec{\gamma}(k;M) - \vec{\Phi}(k;M))}, $$

$$ P(\vec{Y}|H_0) = \frac{1}{\sqrt{2\pi M |\Sigma|}} e^{-\frac{1}{2}\vec{\gamma}(k;M)^T \Sigma^{-1}\vec{\gamma}(k;M)}, $$

where $\vec{\Phi}$ is a known constant vector approximating the magnitude of the first peak of the transient process in $\Phi(i)$ and $\Sigma$ is the covariance of $\vec{\gamma}(k;M)$. In general, the threshold $\eta$ should be determined from the requirements on the probability of detection and of false alarm. In practice, however, a value is typically selected based on experience and simulation.

Taking logarithms of both sides of the above equation yields the likelihood ratio test,

$$ L(\vec{Y}) = \vec{\theta}^T (\Sigma^{-1} + \Sigma^{-1}) \vec{Y} - \vec{\theta}^T \Sigma^{-1} \vec{\theta} > \eta_{H_0}, $$

From (9), an indicator signal determining the position of the probe “on” or “off” the sample is generated. This then yields the start and end time of the transient process of the cantilever. The duration of this transient process is then received by the width computation described below to estimate the width value of the sample. Note that the approximation of $\Phi(k)$ and of the peak magnitude given in $\vec{\theta}$ affects the shape of $L(\vec{Y})$ and leads to an error in the estimation of the end time of the transient process. This error then propagates into the width calculation.

**C. Width Computation**

Under our assumptions, the cantilever continues to be driven into its transient while it scans across the sample. Once leaving the sample, it begins to settle back to steady state. The settling time $T_s$ is defined by [29]

$$ T_s = -\frac{\ln \epsilon}{\zeta \omega_0}, $$

where $\epsilon$ is the tolerance fraction to be chosen by the user. The total time spent in the transient state, $T_t$, is then a sum of the time spent crossing the sample itself (denoted $t$) with the settling time. The time to cross the sample during the $i^{th}$ detection cycle is then given by

$$ t_i = T_i - T_{si} = T_i + \frac{\ln \epsilon}{\zeta \omega_0}, $$

where $\zeta$ is the damping and $\omega_0$ the resonant frequency of the cantilever.
Once the time $t_i$ is calculated, it is used in two ways, to modify the indicator signal to generate the “hit” signal discussed in Sec. III-D below and to calculate the length of the tip trajectory across the sample through the relation $l_i = V_{tip} t_i$. This length can be translated into the width using the knowledge of the scan pattern.

As shown in Fig. 3, the sinusoidal pattern of the tip trajectory under the local raster-scan algorithm is defined by a spatial frequency of $f$ and an amplitude of $A$. Then the width of the sample can be calculated from $l_i$ using the knowledge of the angle $\phi$ to be

$$w_i = l_i \cos \phi = V_{tip} t_i \cos \left[ \arctan \left( 2fA \right) \right],$$

(12)

Fig. 3. Illustration of the geometry of the sinusoidal local raster-scan pattern. The sinusoidal tip trajectory is defined by the spatial frequency $f_i$ and amplitude $A_i$.

D. “Hit” generator

The local raster-scan algorithm operates by estimating the geometric properties of the sample and predicting its spatial evolution forward to determine a tip trajectory until the next crossing. If the indicator signal in (9) were used directly, then the detection would alternate between the two sides of the sample. This alternation skews the estimates of the sample parameters due to the width of the sample, leading to poor tracking. In prior efforts, and ad hoc attempt to overcome this was made by simply ignoring every other detection.

With the width detector, however, a more direct approach is possible. Since the total transit time across the sample is estimated on each crossing (given by $t_i$) the detection can simply be delayed by $t_i$ to approximate a detection at the center of the sample, allowing every detection to be used.

E. Image generation

It is important to give visual feedback to the user in real time. There are a variety of means by which one could do this. In this work, we build a false height image by assigning each measurement a height of 0 if it is off the sample and a height proportional to the estimated width at the current measurement location if it is on the sample. The resulting data is non-raster and can be interpolated into an image using a prior algorithm developed by the authors in [30]. See Fig. 7 for an example of an image produced using this approach.

IV. SIMULATION RESULT

To illustrate and verify the performance of the proposed detector, it was incorporated into a SIMULINK-based model of an AFM. Details on the AFM simulator can be found in [14]; here we describe only the details relevant to the detection scheme.

A. Dynamics in Three Axes

The cantilever model used in this paper was a typical design for an AFM operated under tapping-mode in air. The resonant frequency was 210 KHz, the spring constant was 0.4 N/m and the quality factor was 100. (In order to capture imaging in liquid, the resonance frequency and quality factor should both be reduced to account for viscous damping.) A generic second order system was utilized to represent the dynamics of the piezo actuator in all three axes. For each axis, a semi-automatic tuned PID controller [31] was used to provide a high closed loop bandwidth. The Bode plots for the open loop dynamics, PID controller, and the closed-loop system for the $z$-axis are shown in Fig. 4. The other axes were similar and are omitted for space reasons.

Fig. 4. Bode plots for the $z$-axis of the simulated AFM. The resonant frequency of the piezo appears as the first peak in the open loop system (blue) and was approximately 800 Hz. The semi-automatic tuned PID controller (green) notched out the resonance of the piezo. The closed loop dynamics (red) show a bandwidth of approximately 10 kHz. Note that the second peak is the cantilever resonance at 210 kHz.

B. Sample Model

A simulated sample with a varying width is illustrated in Fig. 5. The sample approximates an underlying biopolymer
(in yellow) and a particle on that polymer (red). The dimensions of the sample can be selected based on the particular biopolymer-macromolecule we are trying to mimic. For the example considered here, we chose to model an RNA polymerase on a DNA strand. Therefore the strand width was selected to be 2.5 nm and the particle to be 16 nm long with a maximum width of 11.25 nm and a height of 10 nm [32]. For simplicity, the lateral profile of the particle was taken to be sinusoidal.

C. Sampling Rate

In order to analyze the state of the cantilever, a fast sampling rate is needed relative to its dynamics. There is no need, however, to sample and actuate on the piezoelectric actuators at this high rate. We therefore consider two sampling rates, $f_{sc}$ for the cantilever and $f_{sa}$ for the actuators.

The choice of $f_{sc}$ is fundamentally limited by the Nyquist rate relative to the cantilever dynamics. In practice, however, one typically chooses a rate 5-10 times faster than the resonant frequency. Here we use 1 MHz, a rate fast enough for most cantilevers but also reasonably implementable in a physical system.

The choice of $f_{sa}$ is more involved. It should be chosen based not only on the actuator and controller dynamics, but also based on the rate at which the tip is moving across the sample. If it is too low, then the system will be slow to respond to edge detections and poor tracking will result. In this example, we selected a tip speed of 20 µm/sec and thus an average transit time of approximately 125 µs. Based on this a sampling rate of 50 kHz was selected, for an average of approximately 6 samples per crossing.

It is important to note that 50 kHz is beyond the closed-loop bandwidth in the $z$-direction (see Fig. 4). We are not, however, attempting to control at this rate but only to measure and estimate. This highlights the difference between tracking and imaging since in imaging one typically operates well below the closed-loop bandwidth.

D. Simulation Results

With the parameter settings discussed above, the simulated sample was scanned with a local raster pattern shown as in Fig. 6(a). The change in the scanning width at approximately (118,194) nm was caused by the simulated particle on the DNA. The red curve represents the estimated center of the DNA strand.

To simulate noise in the measurements, samples from a Gaussian white noise process were added to the cantilever deflection measurements. The signal-to-noise ratio was set to 50 dB. Note that SNR in this setting is defined by the ratio of the peak-to-peak range of the sensor (20 V in our example) and of the peak-to-peak range of the sensor value under a zero input. The window size was set to 50. At a tip speed of 20 µm/sec, this would correspond to a window lasting approximately $\frac{1}{5}$ of the way across the string-like sample. Then we applied the likelihood ratio test algorithm in (9). Here we chose the threshold as $5 \times 10^4$ based on extensive simulations. The resulting likelihood ratio and the indicator signal are shown in Fig. 6(b). As stated before, in general one would choose the threshold $\eta$ to establish desired levels of false positives and false negatives.

The width value estimated from (12) and its true value are shown in Fig. 6(d). The error between these two is due to several sources, including the noise, the choice of threshold, and the approximation of the dynamics of the error signal through the use of the box function (see Sec. III-B), scan rate, cantilever properties as well as controller gains. The estimator almost uniformly over-estimates the width, indicating a positive bias. Such a bias is not necessarily detrimental, however, since the goal is to determine the location of the particle on the biopolymer and thus it is the change in width that is most important.

Using the method discussion in Sec. III-E, a reference image as the visual feedback was generated; this is shown in Fig. 7.

With the particular choices of cantilever dynamics, tip velocity, and local raster-scan parameters, the average motion along the biopolymer is approximately 6 µm/sec. Given the 400 nm size of the image, this corresponds to approximately 10 frames/sec. This rate with the scan size shown above is able to track a macromolecule with a walking speed up to 2 µm/sec. It is currently limited by the detection bandwidth which in turn is limited by the need for the cantilever to return to steady state before the next crossing. Much higher speeds, therefore, can be achieved through, for example,
active $Q$ control [33] to drive the cantilever back to its steady state as quickly as possible. The switching time of the $Q$ control would rely on the prior knowledge of the sample from previous scans. Note, however, that this speed is sufficient even for the dynein motor discussed in Sec. I.

V. CONCLUSIONS

In this paper, we have developed a width detector primarily designed for tracking single macromolecules moving along string-like biopolymers. The results are promising with respect to using the scheme to locate a moving macromolecule. A reference visual feedback was generated and a video rate of 10 frames/sec was achieved in a simulation study.

REFERENCES