Abstract—Biological macromolecules are highly dynamic. Their functions result from dynamic structural changes and dynamic interactions with other molecules. The visualization of the dynamic processes at high spatial and temporal resolution appears to be straightforward to understanding the functional mechanisms of biological macromolecules. Such visualization is only possible by the atomic force microscope (AFM) with a high-speed imaging capability, since dynamic biomolecular processes generally occur on a millisecond timescale. Therefore, we have been improving both the scan speed as well as the tip force exerted on the sample by developing various devices and techniques. The latest version of our high-speed AFM can capture biological processes at an imaging rate of 30-100 ms/frame without disturbing their functions. Here, we report the devices and techniques that realized this performance.

I. INTRODUCTION

AFM was invented in 1986 by Binnig et al. [1] and opened the door to the visualization of nanometer-scale worlds in liquids [2]. This unique capability was received with excitement by researchers of biological sciences as biomolecules only show vital activities in aqueous solutions. Before the AFM era, the high-resolution visualization of individual biopolymers (proteins, DNA) was only possible by electron microscopy in a vacuum environment. However, the AFM’s unique capability does not seem to have contributed significantly to answering many biological questions.

One of the essential features of biological systems is “dynamics”. The functions of biological systems are produced through dynamic processes that occur in biopolymers, biosupramolecules, organelles, and cells. Therefore, what required of AFM for biological sciences is the ability to rapidly acquire successive high-resolution images of individual biomolecules at work. This is solely because this type of imaging is impossible to perform using other techniques. However, the imaging rate of conventional AFM is too slow to observe dynamic behavior of active biomolecules. Thus, endowing AFM with high-speed imaging capability is expected to have a revolutionary impact on biological sciences.

The low imaging rate of AFM mainly originates in that AFM uses the mechanical detection of the sample surface and the mechanical scanning of the sample stage. The sample surface is detected by its interaction with a sharp tip attached to the free end of a soft cantilever, and this interaction is reflected on the cantilever’s mechanical behavior. The relative position between the tip and the sample is changed by scanning the sample stage in the three dimensions. To increase the surface detection speed, we first need a scanner with high resonant frequencies as well as small cantilevers with high resonant frequencies and small spring constants. After having these devices, we need control techniques to realize fast and accurate positioning of the scanner. Since biological samples are fragile and dynamic biomolecular interactions are more delicate, we further need control techniques to maintain the tip-sample interaction force at a small level without sacrificing the scan speed.

At present, digital mode controllers are not sufficiently fast for realizing high-speed AFM capable of imaging at a video rate of 30 ms/frame. Therefore, recently reported sophisticated control algorithms [3]-[5] are not applicable to the video-rate high-speed AFM. We have been exploring simple control techniques [6]-[8] that can be implemented in analog circuits. These controllers have enabled the maximum imaging rate of 30-50 ms/frame, together with other devices optimized for high-speed scanning. The tip force acting on the sample has been reduced greatly. These capacities allowed us to observe biological processes that involved weak protein-protein interactions [9]-[11]. In this article, I review our studies that realized the high-speed AFM with this capability, focusing on the devises and techniques.

II. THEORETICAL CONSIDERATIONS

In the development of high-speed AFM apparatuses, it is important to have practical guidelines that can quantitatively indicate how each device performance affects the scan speed and the imaging rate. An early theoretical consideration of the scan speed limit in contact mode AFM was given in [12]. Concerning tapping mode AFM, the dependence of feedback bandwidth on various factors has been quantitatively described [13]. Numerical simulations have also been performed for this purpose, including the effect of the dynamics of the tip-sample interaction [14]. However, they are not sufficient as practical guidelines. Here, we derive the
quantitative relationship between the feedback bandwidth and the various factors involved in AFM devices and the scanning conditions, based on an idea previously presented for the derivation [7].

A. Acquisition Time and Feedback Bandwidth

Suppose that an image is acquired in time period $T$ over scan range $W \times W$ with $N$ scan lines, then the scan velocity $V_s$ in the x-direction is given by $V_s = 2WN/T$. Assuming that a sample has a sinusoidal shape with periodicity $\lambda$, the scan velocity $V_s$ requires a feedback operation at frequency $f = V_s/\lambda$, to maintain the tip-sample distance. The feedback bandwidth $f_B$ should be greater than or equal to $f$ and therefore be expressed as

$$f_B \geq 2WN/\lambda T \quad (1).$$

Eq.1 gives a relationship between the image acquisition time $T$ and the feedback bandwidth $f_B$. For example, for $T = 30$ ms with $W = 240$ nm and $N = 100$, the scan velocity is 1.6 mm/s. When $\lambda = 10$ nm, $f_B \geq 160$ kHz is required for this scan velocity.

B. Feedback Bandwidth as a Function of Various Factors

Although not described here, the relationship between the open-loop phase delay $\Phi(\omega)$ and the closed-loop phase delay $\phi(\omega)$ is given by $\phi(\omega) = 2\Phi(\omega)$, provided the feedback gain is maintained $\approx$ 1. The phase delay $\Phi$ is given by $2\pi f_{\Delta t}$, where $\Delta t$ is the total time delay in the open-loop and $f$ is the feedback frequency. In tapping mode AFM, the main delays are the reading time of the cantilever oscillation amplitude ($\tau$), the cantilever response time ($\tau_c$), the z-scanner response time ($\tau_z$), the integral time ($\tau_I$) of error signals in the feedback controller, and the parachuting time ($\tau_p$). Here, “parachuting” means that the cantilever tip completely detaches from the sample surface at a steeply inclined region of the sample, and thereafter, time elapses until it lands on the surface again. It takes at least a time of $(1/2f)$ to measure the amplitude of a cantilever that is oscillating at its resonant frequency $f_c$. The response time of second-order resonant systems such as cantilevers and piezoelectric actuators is expressed as $Q/\pi f_0$, where $Q$ and $f_0$ are the quality factor and resonant frequency, respectively. The feedback bandwidth is usually defined by the feedback frequency that results in a phase delay of $\pi/4$. On the basis of this definition, the feedback bandwidth $f_B$ is approximately expressed as

$$f_B = \frac{f_c}{8} \left( 1 + \frac{2Q_c}{\pi} + \frac{2Q_z f_c}{f_c^2} + 2f_c (\tau_{\Delta t} + \tau_1) \right) \quad (2),$$

where $f_c$ is the z-scanner’s resonant frequency; $Q_c$ and $Q_z$ are the quality factors of the cantilever and z-scanner, respectively. $\alpha$ represents a factor related to the phase compensation effect given by the D component in the proportional-integral-differential (PID) feedback controller.

The parachuting time is a function of various parameters such as the sample height $h_0$, the free oscillation amplitude $A_0$ of the cantilever, the amplitude set point $r$, the phase delay $\beta$, and the cantilever’s resonant frequency. Its analytical expression was previously obtained under rough approximation [7] as

$$\tau_p \approx \left[ (\tan \beta)/\beta - 1 \right] / f_c \quad (3),$$

where $\beta$ is $\cos^{-1} \left( 2A_0 (1 - r)/h_0 \sin(\beta/2) \right) \right]$. However, experimental estimations of the feedback bandwidth imposes a modification on $\beta$ as

$$\beta \approx \cos^{-1} \left( 2A_0 (1 - r)/\left[ 5h_0 \sin(\beta/2) \right] \right) \quad (4).$$

It is difficult to theoretically estimate the integral time ($\tau_I$) with which the optimum feedback control is attained. Its experimental estimation approximately resulted in $\tau_I \approx 4h_0 \sin(\beta/2)/(A_0 f_c)$. The feedback bandwidths obtained theoretically and experimentally are shown in Fig.1 as a function of $r$ and $2A_0/h_0$.

![Feedback Bandwidth as a Function of Various Factors](image)

III. MECHANICAL DEVICES

A. Small Cantilevers

The small cantilevers recently developed by Olympus are made of silicon nitride and are coated with gold of thickness of ~20 nm [15]. They have a length of ~6 μm, a width ~2 μm, and a thickness ~90 nm, which results in the resonant frequencies of ~3.5 MHz in air and ~1.2 MHz in water, a spring constant $k_c$ ~0.2 N/m, and $Q_c$ ~2.5 in water. Therefore, their response time $\tau_c$ = $Q_c/(\pi f_c)$ is 0.66 μs in water. The tip is not sufficiently small (the apex radius, ~17 nm) for the high-resolution imaging of biological samples. We usually attach a sharp tip on the original tip by electron beam deposition in naphthalene gas, which can be sharpen to ~4 nm by plasma etching in argon or oxygen gas. To detect the cantilever’s oscillation amplitude at every half oscillation period, we developed a peak-hold method; the peak and bottom voltages are captured and then their difference is output as the amplitude [16]-[17]. The amplitude detection
time $\tau_d (=1/2f_c)$ is 0.42 $\mu$s for our small cantilevers with a resonant frequency of ~1.2 MHz in water.

B. Sample-stage Scanner

We use a flexure-based stage for the x- and y-scanners [10]. The flexure stage is produced by monolithic processing to minimize the number of resonant elements. The slowest y-scanner displaces the x-scanner, and the x-scanner displaces the z-scanner, on top of which a sample stage is placed. The maximum displacements at 100 V of the x- and y-scanners are 1 $\mu$m and 3 $\mu$m, respectively. The x-scanner shows resonant peaks at 45 kHz, 65 kHz, and higher frequencies, as shown in Fig. 2. The peak at ~20 kHz does not originate from vibrations of the x-scanner but from the measurement method used.

Two z-piezoelectric actuators (maximum displacement, 2 $\mu$m at 100 V; self-resonant frequency, 360 kHz) are used to alleviate the z-scanner vibrations; impulsive forces are countered by the simultaneous displacements of the z-piezoelectric actuators of the same length in the counter direction. In this arrangement, the counterbalance works effectively below the resonant frequencies. The vibration phase changes sharply around the resonant frequencies, and therefore, a slight difference in the mechanical properties of the two actuators disturbs the counterbalance. To solve this problem, we are now attempting to use a different design for the z-scanner, as mentioned later.

IV. ACTIVE DAMPING

It is impossible to manufacture a scanner that exhibits no resonance around the harmonic frequencies contained in its driving signals. Therefore, the elimination of unwanted vibrations from the scanner is a key to realizing high-speed AFM. Furthermore, the bandwidth of any mechanical systems is limited by its dimensions and materials, and hence, some manipulation techniques are also required to overcome this limitation and expand the scanner bandwidth.

A. Active Damping of X-scanner

The active damping of the x-scanner is easy because their scan speed is not high, and their scan waves are known beforehand and periodic, and therefore, feedforward control for active damping can be implemented in a digital mode [18]-[20]. When the line scan was performed at 3.3 kHz without damping, its displacement exhibited vibrations [Fig. 3(a)]. When it was driven with inverse compensation damping, the x-scanner moved approximately in a triangle waveform [Fig. 3(b)]. This damping was also verified by imaging mica surfaces at 33 frames/s over a scan range of 240 $\times$ 240 nm$^2$ with 100 scan lines (this imaging speed corresponds to an x-scan frequency of 3.3 kHz). When no active damping was applied, striation patterns parallel to the y-axis appeared [Fig. 3(c)]. The periodicity of the patterns indicates that the vibrations had a frequency of ~60 kHz, which corresponded with the frequency of the main resonant peak shown in Fig. 2. When active damping was applied, the striation patterns disappeared [Fig. 3(d)]. Here, note that the resonant vibrations in the x-scanner also produce vibrations in the z-direction.

B. Feedback Type of Active Damping for Z-scanner

For the active damping of the z-scanner, we cannot use feedforward control in digital modes as its scan waves are unpredictable and the scan speed with high-speed AFM is too fast for digital driving signals to be calculated in real time. The active Q-control is well known as an active damping technique and has been often used to control the quality factor of cantilevers [21]-[24]. When this control is applied to the z-scanner, its displacement or speed has to be detected. However, it is difficult to do this. This problem is solved by using a mock scanner characterized by the same transfer function as the z-scanner (Fig. 4) [6]. Using this technique,
we achieved a bandwidth of 150 kHz (Fig. 5) and a quality factor of 0.5, which resulted in a response time of 1.1 μs. However, the higher resonant vibrations are not damped well. This is because the two main elemental resonators are connected in parallel judging from the phase spectra [Fig. 5(b)].

When the elemental resonators are connected in series, we can use multiple Q-controllers connected in series, each of which contain a mock scanner characterized by the transfer function representing one of the elemental resonators. However, when they are connected in parallel, active damping becomes more difficult. Although we have not yet examined experimentally, there is a solution to this problem. Here, we consider only the case where two elemental resonators are connected in parallel. The feedback Q-controller $H(s)$ is expressed as

$$H(s) = \frac{1}{G_1(s)+G_2(s)} \frac{1}{r(s)} = \frac{1}{a_1\omega_1^2 + \frac{b_1\omega_1^2}{s^2 + \frac{a_1\omega_1}{Q_1}s + \omega_1^2 + \frac{b_1\omega_1}{Q_2}s + \omega_2^2}} - \frac{s^2 + \frac{a_2\omega_2}{Q_2}s + \omega_2^2}{\omega_2^2} + 1 $$

where $a$ and $b$ are the weight factors of two resonators (hence, $a+b=1$). To eliminate the highest order term from $H(s)$, $\omega_0^2$ should equal $a_1\omega_1^2+b_1\omega_2^2$, which approximately results in

$$H(s) \sim [(\text{band-pass filter}) \times g - 1] \times s/\tilde{\omega}$$

(6).

The band-pass filter $BP(s)$, $\tilde{\omega}$, and the gain parameter $g$ are respectively expressed as

$$BP(s) = \frac{a}{Q_2\omega_2 + \frac{b}{Q_1\omega_1}}s + \frac{\omega_0^2}{\omega_1^2\omega_2^2} \frac{Q_2\omega_2}{Q_1\omega_1}$$

and

$$\tilde{\omega} = 1\left\{ \frac{1}{Q_2\omega_0} - \frac{a}{Q_2\omega_1} - \frac{b}{Q_1\omega_2} \right\} $$

(8).

$$g = c\tilde{\omega} \left\{ \frac{a}{Q_2\omega_2} + \frac{b}{Q_1\omega_1} \right\} + 1$$

(9),

where $c$ is given by

$$c = \frac{1}{\omega_1^2} + \frac{1}{\omega_2^2} - \frac{\omega_0^2}{\omega_1^2} - \frac{\omega_0^2}{\omega_2^2} + \frac{1}{\omega_0^2} \left[ \frac{a_1\omega_1}{Q_2(Q_0 - 1)} - \frac{a_2\omega_2}{Q_1(Q_2 - 1)} \right]$$

(10).

The effect of this controller $H(s)$ is depicted in Fig. 6.

![Fig. 5. Frequency spectra of mechanical response of the z-scanner. (a) Gain spectra, (b) phase spectra. The gray-line curves represent the response without feedback Q-control, and the black-line curves represent the response with feedback Q-control.](image)

![Fig. 6. Effect of feedback control for active damping on the resonant system that consists of two elemental resonators connected in parallel.](image)

C. Feedforward Type of Active Damping for Z-scanner

In the feedforward type of active damping, the z-scanner is driven through a circuit with a transfer function $1/G(s)$, where $G(s)$ is the transfer function of the z-scanner. However, for a complicated $G(s)$, it is very difficult to design an electric circuit characterized by $1/G(s)$. We invented a circuit that can automatically produce the inverse transfer function (APTIF) for a given transfer function. [8, 25]. The transfer function $K(s)$ of the circuit shown in Fig. 7(a) is given by $K(s) = 1/[1 + g(M(s) - 1)]$; here, $M(s)$ represents a mock scanner. In the case of $g = 1$, a complete inverse transfer function, $1/M(s)$, can be realized. However, delays occur in the electronic components such as the operational amplifiers used in the circuit. Therefore, the gain factor $g$ must be less than 1, and consequently, the complete inverse transfer function cannot be realized. In fact, $K(s)M(s)$ has a large resonant peak at a frequency higher than the peak frequencies of $M(s)$. However, this peak can be eliminated by a Notch filter. When the disagreement between $K(s)$ and $1/M(s)$ is significant, we can reduce the difference by placing a phase compensator ($1 + \text{differential operation}$) immediately before or after the mock scanner and by using a multiple-loop circuit, as indicated in Fig. 7(b). The difference is more reduced by increasing the number of the nested loops. We examined the effectiveness of the APTIF method by applying it to a
z-scanner under development. In this z-scanner, a single piezoelectric actuator is used, and is only held at the rims and corners of a plane perpendicular to the displacement direction. The piezoelectric actuator used is similar to that used for the scanner mentioned above. The z-scanner exhibited large resonant peaks at 370 kHz and 540 kHz, as indicated with a black-line curve in Fig. 8(a). The resonant frequency of 370 kHz is similar to that of the free oscillation of the piezoelectric actuator (410 kHz). Judging from the phase spectrum [the black-line curve in Fig. 8(b)], the two resonators are connected in parallel. Thus, we applied feedforward active damping using a double-loop APITF circuit containing a mock scanner composed of two LRC circuits connected in parallel. The resulting gain and phase spectra are shown with gray-line curves in Fig. 8(a) and 8(b), respectively. The peak at 370 kHz was completely removed and the frequency that gave a 90° phase delay was increased from 370 kHz to 500 kHz.

Fig. 7. Circuit diagram for APITF. G(s) represents a given transfer function for a real resonant system. M(s) represents a transfer function of a mock resonator for G(s). (a) Single-loop circuit type. K(s) approximately gives 1/M(s). (b) Nested circuits for improving the approximation.

Fig. 8. Frequency spectra of mechanical response of the z-scanner. (a) Gain spectra, (b) phase spectra. Black-line curves represent the responses without feedforward active damping, and gray-line curves represent the responses with feedforward active damping.

V. DYNAMIC PID CONTROL

To observe biological processes containing delicate protein-protein interactions, the tip-sample interaction force has to be minimized. The tapping force exerted on the sample by the oscillating cantilever tip can be reduced by using an amplitude set point \( r \) close to 1. However, this set point condition induces “parachuting”, during which the error signal (difference between the cantilever’s peak-to-peak oscillation amplitude \( A_{pp} \) and the set point \( A_s \)) saturates at \( (2A_s - A_p) = 2A_p(1 - r) \). The saturated error signal becomes smaller as \( r \) increases, and hence, the parachuting time \( (\tau_p) \) is prolonged. The feedback gain cannot be increased to shorten \( \tau_p \), since a large gain induces overshoot at uphill regions of the sample, resulting in the instability of the feedback operation. To solve this problem, we devised a new PID controller named “dynamic PID controller” in which the gain parameters are automatically altered depending on the oscillation amplitude [7]. A threshold level \( A_{upper} \) is set between \( A_s \) and \( 2A_s \) (or is set to \( A_s \)). We usually set a value of \( A_{upper} \) that is much closer to \( A_s \) than \( 2A_s \). When \( A_{pp} \) exceeds \( A_{upper} \), the gain parameters are increased. This method significantly enhances the feedback bandwidth, particularly when the set point \( r \) is close to 1 (see dotted-line curves in Fig. 9). The feedback bandwidth is independent of \( r \), provided \( r \) is less than ~0.9, indicating that parachuting does not occur. The superiority of the dynamic PID control is also clear from captured images [Fig. 10(b) and 10(c)] of a mock sample with steep slopes [Fig. 10(a)].

Fig. 9. Feedback bandwidth as a function of the set point \( r \). Solid-line curves: feedback bandwidths measured with a conventional PID controller; dotted-line curves: feedback bandwidths measured with the dynamic PID controller. These curves are aligned from top to bottom according to the ratio \( 2A_s/\delta_0 = 5, 2, 1, \) and 0.5.

Fig. 10. Pseudo-AFM images of a sample with rectangles with two different heights. (a) A mock AFM sample. The images were obtained using a conventional PID controller (b) or using the dynamic PID controller (c). These simulations with the mock AFM system were made under the following conditions: the cantilever resonant frequency, 1.2 MHz; quality factor of the cantilever oscillation, 3; the resonant frequency of the z scanner, 150 kHz; and quality factor of the z scanner, 0.5; the line scan speed, 1 mm/ s; the line scan frequency, 1 kHz; the frame rate, 100 ms/ frame, the ratio of 2A_s to the total sample height = 1; and \( r = 1 \).

VI. PRESENT CAPACITY AND LIMITATIONS

Our most recent high-speed AFM can capture an image in 30-50 ms/frame, without significantly disturbing weak protein-protein interactions. Some biological processes were successfully captured on video. For example, the dynamic
interaction process of chaperonin GroEL-GroES, which is regulated by the ATPase reaction, was clearly visualized. The alternate association and dissociation of GroES at the two rings of GroEL were evident. Surprisingly, before the alternate switching took place, a GroES-GroEL-GroES complex often appeared whose existence had been a controversial issue for a long time. As demonstrated by this observation, high-speed AFM imaging can quickly answer to biological questions.

To expand the scope of biological samples that can be studied by high-speed AFM, the imaging rate must be enhanced further and the tip-sample interaction force must be reduced significantly. Presently, the cantilever is the factor that prevents the achievement of both of these improvements. Generally, to improve the resonant frequency, one must compromise the stiffness and vice versa. The most advanced small cantilevers developed by Olympus appear to have almost achieved the ultimate goal of balancing these two mechanical quantities. Considering their practical use, doubling the resonant frequency to ~2.4 MHz in water seems to be the upper resonant frequency limit, provided the cantilever compliance is not sacrificed.

VII. FURTHER REDUCTION OF TAPPING FORCE

Reduction of the tapping force must be achieved by increasing the detection sensitivity of the tip-sample interaction. The nonlinear impulsive tip-sample interaction induces higher-harmonic vibrations of the cantilever. In the amplitude detection, these vibrations are neglected. Since the impulsive force is exerted transiently in a short time, its peak force is relatively large. This means that the peak force must be a highly sensitive quantity. The force \( F(t) \) cannot be detected directly because the cantilever oscillation gain is lower at higher harmonic frequencies. \( F(t) \) can be calculated from the cantilever’s oscillation wave \( z(t) \) by substituting \( z(t) \) into the equation of cantilever motion and then subtracting the excitation signal (i.e., inverse determination problem) [26]-[27]. Figure 11(b) shows force signals \( F(t) \) that were obtained by off-line calculation using an oscillation signal of a small cantilever weakly interacting with a mica surface in water. Since the time width of the impulsive force is narrow, it appears to be difficult to capture the peak force using a sample/hold circuit. Instead of capturing the peak force, we can calculate it from the time when the cantilever oscillation reaches the end of the bottom swing.

Our ultimate goal is the realization of noncontact high-speed AFM in which we can use a much stiffer cantilever with a significantly high resonant frequency. Therefore, an ultra-fast scan speed must be achieved without damaging the sample. The discussion on this issue is presented in [28]-[29].

REFERENCES


