Reprinted from January 1987, Vol. 109, Journal of Vibration, Stress, and Reliability in Design

Estimation of the Surface Roughness on the Race or Balls of Ball Bearings by Vibration Analysis

H. Kanai
Education Center for Information Processing,
Tohoku University,
Kawauchi, Sendai 980, Japan

M. Abe

K. Kido

Research Center for Applied
Information Sciences,
Tohoku University,
Katahira 2-1-1, Sendai 980, Japan

This paper describes a vibration-based diagnostic method by estimating the surface roughness on the rotating ring or balls in ball bearings. The surface roughness has been measured by a stylus that directly traverses the surface of the ring or balls obtained by taking apart the ball bearing. We developed a new method to estimate accurately the surface roughness by analyzing the short-length vibration signal that is excited when balls encounter flaws on the rotating ring or when races encounter flaws on the balls in a ball bearing. Our experimental results confirm that the roughness estimated by the proposed method agrees with that measured directly by using a stylus even in the case of crack μm wide. We applied this new method to the diagnosis of surface roughness in small-sized ball bearings and inferior samples were detected with a 95.3 percent accuracy rate.

1 Introduction

Vibration is induced when rings or balls in a ball bearing have a rough surface or flaws. For example, in the case of the ball bearings used to support the spindle of the rotating head of a video tape recorder, such defects debase the quality of the played-back picture. Therefore, at the final step in the manufacturing process, defects in ball bearings have been detected and classified aurally by inspectors listening to the vibration signals picked up using the Anderon meter [1] as described in Section 4. However, it requires a great deal of time to train a good inspector. Additionally, the physical and mental condition of inspector affects the results of detection and classification.

In response, several methods have recently been proposed for the automatic detection and classification of the flaw on the race or balls in a ball bearing [2–4]. These methods utilize the pulses excited when the balls encounter flaws on the race or when races encounter flaws on balls. Defects are detected and classified by checking the peaks observed at the expected intervals which are calculated from the position of the flaws, the shape of the bearing and the rotation speed of the inner ring. We proved by experiments that the defects can be classified with an accuracy rate of 97.9 percent [4]. However, the shape of the surface roughness or the number of flaws on the raceway or on the balls cannot be detected by this kind of periodicity analysis.

For these purposes, the surface roughness has been measured by a stylus which directly traverses the surface of the ring or the ball. However, it is necessary in this case to take apart the ball bearing. That is, it is impossible to measure the surface roughness on the race or the ball in the assembled rolling ball bearing.

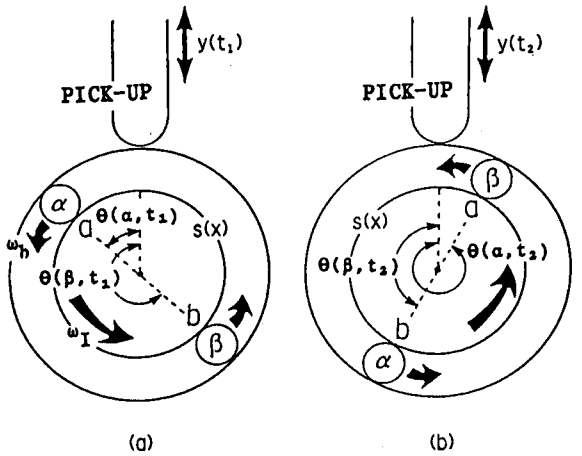

Fig. 1 Illustration of the simplified ball bearing showing the relation between the surface roughness s(x) and the picked up vibration signal y(t): (a) at a time $t = t_1$; (b) at a time $t = t_2$.

Therefore, we have developed a new diagnostic method to estimate accurately the surface roughness of the rotating ring and the balls by analyzing the short-length vibration signal excited by the surface defects. Our experimental results confirm that the roughness estimated by the proposed method agrees with that measured directly by using a stylus. Finally, we apply this new method to diagnosis of the surface roughness in small-sized ball bearings and inferior samples are detected with a 95.3 percent accuracy rate. The proposed method has

Contributed by the Noise Control and Acoustics Division and presented at the Winter Annual Meeting, Anaheim, California, December 7–12, 1986, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS. Manuscript received at ASME Headquarters, August 1, 1986. Paper No. 86-WA/NCA-8.

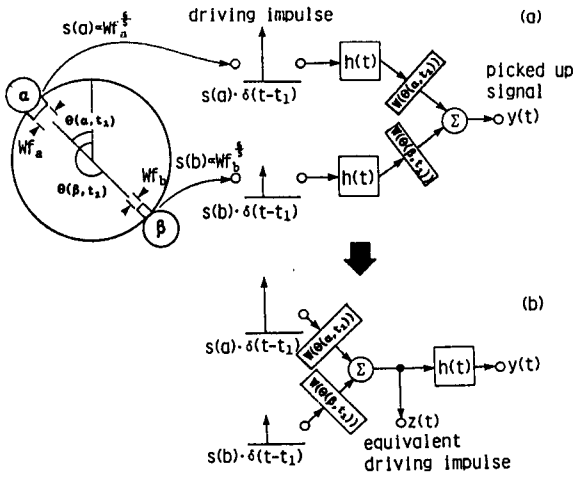

Fig. 2 (a) The picked up signal $y(n)$ is the weighted sum of the resonant vibrations which are driven by the impulses $s(a)\cdot\delta(t-t_1)$ and $s(b)\cdot\delta(t-t_1)$ caused by the balls in Fig. 1(a). (Wf_a and Wf_b: the width of the flaw, $h(t)$: the impulse response of the resonant vibration). (b) When the picked up signal $y(n)$ is deconvoluted by the inverse filter $h(n)^{-1}$ of the resonant vibration, the resultant driving impulse $z(t)$ denotes the weighted sum of the driving impulses due to the collisions.

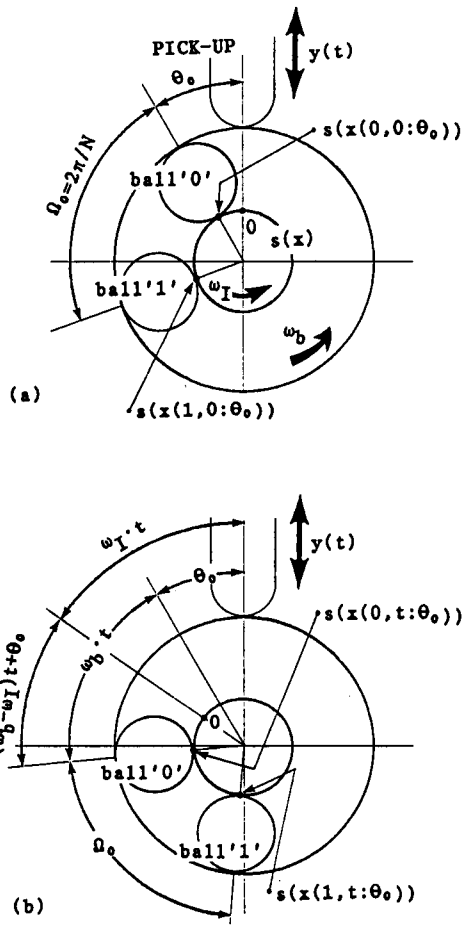

Fig. 3 Illustration of the ball bearing explaining the principle of the proposed method for the estimation of the surface roughness of the inner race. (In general case). (a) at a time $t=0$; (b) at a time t.

the following three advantages: (*a*) The degree of the surface roughness and the number of flaws on the raceway or balls, which cannot be detected by the periodicity analysis, is obtained. (*b*) The surface roughness is estimated from the vibration signal which length is so short that the every point on the surface contacts with a ball at least one time in the observation interval. (*c*) The accuracy rate of the estimated surface roughness is evaluated by comparing the observed vibration signal with the vibration signal synthesized from the estimated roughness.

2 Estimation of the Surface Roughness Based on the Vibration Analysis of Ball Bearings

2.1 Theoretical Explanation Using a Simplified Example. In this paper, we estimate the surface roughness of the rolling ring and the balls by analyzing the vibration signal detected by a vibration pick-up attached to the outer ring when the outer ring is fixed by imposing an axial pressure and the inner one is revolved at a constant speed. This section describes simply the principle for the estimation of the surface roughness of the rolling ring. Figure 1 shows a simplified ball bearing having only two balls, α and β, which are revolving π radian apart. Let the surface of the outer ring and every ball be smooth. At a time $t=t_1$, the balls α and β contact with the points a and b on the inner race, respectively, as shown in Fig. 1(*a*). When there is a flaw on the inner race, and a revolving ball arrives at the edge of the flaw on the raceway, the ball rotates around the edge, and the ball hits the other edge of the flaw. It is known that the amplitude of the impulse generated by the collision is in proportion to the 6/5th power of the width of the flaw [5, 7]. If there are flaws at the points a and b on the race, the impulses $s(a)\cdot\delta(t-t_1)$ and $s(b)\cdot\delta(t-t_1)$ are generated, where $s(x)$ denotes the amplitude of the impulse caused by the flaw at the point x on the race and $\delta(t)$ denotes the unit impulse

$$\delta(t-t_0) = \begin{cases} 0, & (t\neq t_0) \\ \infty, & (t=t_0) \end{cases}$$

and

$$\int_{-\infty}^{+\infty}\delta(t-t_0)dt=1.$$

As shown in Fig. 2(*a*), each impulse drives the resonant vibration and the picked up vibration $y(t)$ is the weighted sum of the resonant vibrations as follows:

$$y(t) = W(\theta(\alpha,t_1))\cdot\{h(t)^*s(a)\cdot\delta(t-t_1)\} + W(\theta(\beta,t_1))\cdot\{h(t)^*s(b)\cdot\delta(t-t_1)\}, \quad (1)$$

where $h(t)$ denotes the impulse response of the resonant vibration, * the convolution, $\theta(x,t)$ the angle between the pick-up and the ith ball at a time t, and $W(\theta(x,t))$ is the weighting function. Since the characteristics of the impulse response $h(t)$ of the resonant vibration are determined both by the size and by the material of the outer ring, the two resonant vibrations driven by the impulses $s(a)\cdot\delta(t-t_1)$ and $s(b)\cdot\delta(t-t_1)$ have the same resonant characteristics except for the amplitude of each signal. Therefore, the picked up signal $y(t)$ is represented as follows

$$y(t)=h(t)^*z(t), \quad (2)$$

where $z(t)$ denotes the equivalent driving impulse as shown in Fig. 2(*b*) and $z(t)$ is the weighted sum of the impulses as follows

$$z(t)=W(\theta(\alpha,t_1))\cdot s(a)\cdot\delta(t-t_1)+W(\theta(\beta,t_1))\cdot s(b)\cdot\delta(t-t_1). \quad (3)$$

That is, the observed vibration $y(t)$ is the resonant vibration driven by $z(t)$ caused by the collisions on the points a and b.

Next, at a time $t=t_2$, the balls α and β contact again with the points b and a on the race, respectively. The impulses are excited due to flaws, and the equivalent driving impulse $z(t)$ of equation (3) is modified as follows

$$z(t) = \{W(\theta(\alpha,t_1))\cdot s(a) + W(\theta(\beta,t_1))\cdot s(b)\}\delta(t-t_1) + \{W(\theta(\beta,t_2))\cdot s(a) + W(\theta(\alpha,t_2))\cdot s(b)\}\delta(t-t_2). \quad (4a)$$

Thus, the equivalent driving impulse $z(t)$ is expressed at $t=t_1$ and $t=t_2$ as follows

$$\begin{cases} z(t_1) = W(\theta(\alpha,t_1))\cdot s(a) + W(\theta(\beta,t_1))\cdot s(b), & (5a) \\ z(t_2) = W(\theta(\beta,t_2))\cdot s(a) + W(\theta(\alpha,t_2))\cdot s(b). & (5b) \end{cases}$$

All of the terms other than $s(a)$ and $s(b)$ involved in the above equations are determined when the following three conditions are satisfied:

(a) The values of the angular velocity ω_I and ω_b are determined previously and the initial angle θ_0 of a ball relative to the pick-up at a time $t=0$ is determined. By using these parameters, the following two values are calculated: the point $x(i,t)$ on the race with which the ith ball contacts, and the angle $\theta(i,t)$ of the ith ball relative to the pick-up at a time t.

(b) The weighting function $W(\theta)$ is determined previously.

(c) The inverse filter $h(t)^{-1}$ of the resonant vibration is determined, where $h(t)^{-1} * h(t) = \delta(t)$. Then, the equivalent driving impulse $z(t)$ is calculated from the convolution between $h(t)^{-1}$ and the observed vibration $y(t)$ as follows

$$z(t) = y(t)^* h(t)^{-1}. \quad (6)$$

Therefore, the surface roughness quantities $s(a)$ and $s(b)$ are determined by solving the set of simultaneous equations (5). Since the vibration $y(t)$ is generally corrupted by additive noise, more accurate surface roughness is estimated by using the least mean square fitting to the P equivalent driving impulses $\{z(t_j)\}$, $(j=1,2,\ldots,P)$, where P denotes the number of collisions between the points a and b and balls in the observation interval.

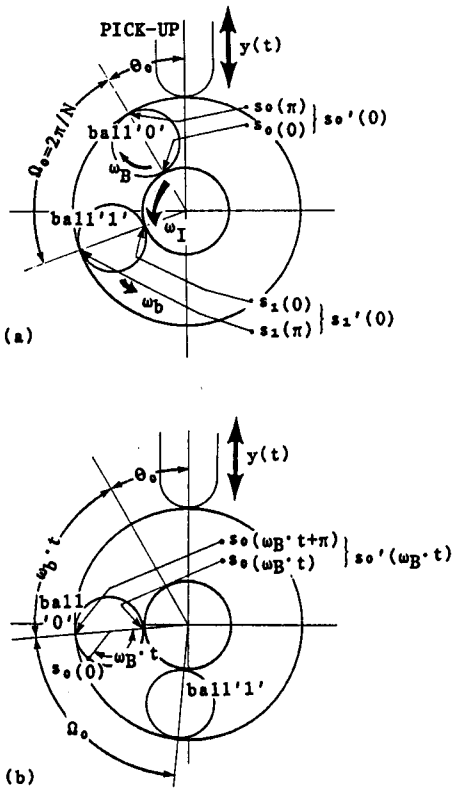

Fig. 4 Illustration of the ball bearing explaining the principle of the proposed method for the estimation of the surface roughness of the balls: (a) at a time $t=0$; (b) at a time t.

2.2 Estimation of the Surface Roughness of the Rolling Ring. This paragraph describes the principle of the proposed method when a ball bearing has N balls, each of which is $\Omega_0 = 2\pi/N$ [radian] apart. Let the surface of the outer race and the surface of every ball be smooth. When the initial angle of the first ball (0) relative to the pick-up is θ_0 [radian] at a time $t=0$ as shown in Fig. 3(a) and the balls revolve at a speed of ω_b [radian/s], the angle $\theta(i,t{:}\theta_0)$ of the ith ball relative to the pick-up at a time t is equal to $\omega_b t + \Omega_0 \cdot i + \theta_0$, $(i=0,1,\ldots,N-1)$ as shown in Fig. 3(b). Let the origin be the point 0 on the race just under the pick-up at a time $t=0$ as shown in Fig. 3(a). Since the inner ring is revolved at a speed of ω_I, the race point $x(i,t{:}\theta_0)$ coming in contact with the ith ball at a time t is equal to $(\omega_b - \omega_I)t + \Omega_0 \cdot i + \theta_0$ as shown in Fig. 3(b). Thus, the observed signal $y(t)$ of equations (2) and (4) is expressed as follows

$$y(t) = z(t)^* h(t) \quad (2)$$

and

$$z(t,\theta_0) = \sum_{i=0}^{N-1} W((\theta(i,t{:}\theta_0))_{2\pi}) \cdot s((x(i,t;\theta_0))_{2\pi}), \quad (4b)$$

where

$$\theta(i,t{:}\theta_0) = \omega_b t + \Omega_0 \cdot i + \theta_0,$$
$$x(i,t{:}\theta_0) = (\omega_b - \omega_I)t + \Omega_0 \cdot i + \theta_0,$$

and $((\Omega))_c$ denotes the remainder when Ω is divided by the modulus c so that the quantity Ω used in the function $W(\Omega)$ or $s(\Omega)$ is defined in the range $0 \le \Omega \le 2\pi$.

In the discrete time system, by letting the integer M be the number of points on the inner race in the range $0 \le \Omega < 2\pi$, the equivalent driving impulse of equation (4b) is represented as follows

$$z(n{:}\theta_0) = \sum_{i=0}^{N-1} W(([\theta_d(i,n{:}\theta_0)])_M) \cdot s(([x_d(i,n{:}\theta_0)])_M), \quad (4c)$$

where

$$\theta_d(i,n{:}\theta_0) = MF_b \cdot nT_s + Mi/N + M\theta_0/2\pi,$$
$$x_d(i,n{:}\theta_0) = M(F_b - F_I) \cdot nT_s + Mi/N + M\theta_0/2\pi,$$

T_s: sampling period,

$2\pi F_b = \omega_b$, $2\pi F_I = \omega_I$

and $[\Omega]$ represents the integer next to the number Ω.

By using the inverse filter $h(n)^{-1}$ of the resonant vibration, the equivalent driving impulse is obtained from the discrete observed signal $y(n)$. However, in the practical case, the equivalent driving impulse $\underline{z}(n)$ obtained from the observed signal $y(n)$ involves the error function $\Delta z(n{:}\theta_0)$. Then, $\underline{z}(n)$ is represented as follows:

$$\underline{z}(n) = y(n)^* h(n)^{-1} = z(n{:}\theta_0) + \Delta z(n{:}\theta_0). \quad (7)$$

If the above model shown in equation (4c) just represents the signal $z(n)$, the error $\Delta z(n{:}\theta_0)$ is equal to zero. Therefore, the surface roughness $s(x)$ is determined directly from the picked up vibration $y(n)$ by applying a least mean square criterion to the error $\{\Delta z(n{:}\theta_0)\}$. The total squared error $\alpha(\theta_0)$ is defined by

$$\alpha(\theta_0) = \sum_{n=0}^{L-1} \Delta z(n{:}\theta_0)^2 = \sum_{n=0}^{L-1} |z(n) - z(n{:}\theta_0)|^2, \quad (8a)$$

where L defines the interval where the error minimization oc-

curs. Here, the following two vectors and one matrix are introduced: $\mathbf{Z}$ and $\mathbf{S}$ are L-dimensional and M-dimensional vectors, respectively, whose transposes are given by

$$\mathbf{Z}^T = [z(0), z(1), \ldots, z(L-1)]$$

and

$$\mathbf{S}^T = [s(0), s(1), \ldots, s(M-1)]$$

$\mathbf{W}$ is the L-by-M matrix as

$$\mathbf{W} = \begin{bmatrix} w_{00} & w_{01} & \cdot & \cdot & \cdot & w_{0,M-1} \\ w_{10} & w_{11} & \cdot & \cdot & \cdot & w_{1,M-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ w_{L-1,0} & w_{L-1,1} & \cdot & \cdot & \cdot & w_{L-1,M-1} \end{bmatrix},$$

where w_{nm} is equal to $W(([\theta_d(i,n:\theta_0)]_M))$ of equation (4c) if the *race point* m contacts with the ith ball at a time $t = n \cdot T_s$: otherwise w_{nm} is equal to zero.

By using these vectors and the matrix, the total squared error of equation (8) to be minimized is represented as follows

$$\alpha(\theta_0) = (\mathbf{Z} - \mathbf{W}\,\mathbf{S})^T(\mathbf{Z} - \mathbf{W}\,\mathbf{S}). \tag{8b}$$

Minimization of $\alpha(\theta_0)$ is obtained by setting the partial derivatives of $\alpha(\theta_0)$ with respect to $s(m)$ to zero as follows

$$\frac{\partial \alpha(\theta_0)}{\partial s(m)} = 0. \quad (\text{for } m = 0,1,2, \ldots, M-1)$$

Then, the set of M linear simultaneous equations are obtained as:

$$\mathbf{W}^T\mathbf{W}\,\mathbf{S} = \mathbf{W}^T\mathbf{Z}. \tag{9}$$

When the length L of the interval for the error minimization is so long that every point on the race contacts with balls at least one time, the M-by-M matrix $\mathbf{W}^T\mathbf{W}$ is nonsingular due to the following reasons:

(*a*) Since the inner ring is revolved, the angle of a point coming in contact with a ball varies every moment.

(*b*) The weighting function $W(\theta)$ is expressed by a nonlinear function as shown in Fig. 5.

Therefore, the surface roughness $\mathbf{S}$ of the rolling ring is solved uniquely from equation (9) as follows:

$$\mathbf{S} = (\mathbf{W}^T\mathbf{W})^{-1} \cdot \mathbf{W}^T\mathbf{Z}, \tag{10}$$

where $(\mathbf{W}^T\mathbf{W})^{-1}$ denotes the inverse matrix of $\mathbf{W}^T\mathbf{W}$.

2.3 Estimation of the Surface Roughness of the Balls. This paragraph describes the principle of the estimation of the surface roughness of the ball (see Fig. 4). Let the surface of the outer and inner race be smooth. When the initial angle of the first ball (0) relative to the pick-up is θ_0 [radian] at a time $t = 0$ as shown in Fig. 4(*a*) and the balls revolve at a speed of ω_b, the angle $\theta(i,t:\theta_0)$ of the ith ball relative to the pick-up is equal to $\omega_b t + \Omega_0 \cdot i + \theta_0$ at a time t as shown in Fig. 4(*b*). Let the origin be the point 0 on each ball coming in contact with the inner race at a time $t = 0$ and the rotation speed of each ball be ω_B [radian/s], ($\omega_B \neq \omega_b$). Then, the ith ball surface point which contacts with the inner and the outer race are equal to $\omega_B t$ and $\omega_B t + \pi$, respectively, as shown in Fig. 4(*b*). Since the observed resonant vibration $y(t)$ is generated by the contacts between the ball and the inner race and by the contacts between the ball and the outer ring, $y(n)$ is expressed as follows

$$y(t) = z(t) * h(t) \tag{11}$$

and

$$z(t:\theta_0) = \sum_{i=0}^{N-1} W((\theta(i,t:\theta_0))_{2\pi}) \cdot \{s_i((\omega_B t)_{2\pi}) + \rho \cdot s_i((\omega_B + \pi)_{2\pi})\}, \tag{12a}$$

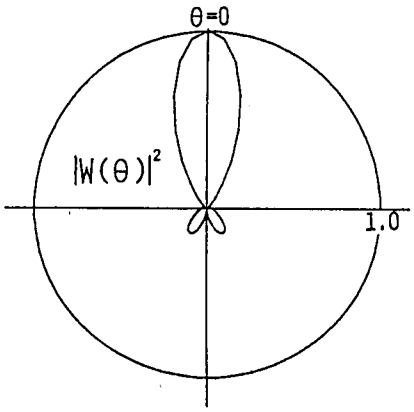

Fig. 5 The squared weighting function $|W(\theta)|^2$ representing the amplitude ratio of the excited impulse at a point to the equivalent driving impulse $z(t) = y(n) * h(n)^{-1}$, where the angle of the point relative to the point of the pick-up ($\theta = 0$) is θ.

where $s_i(x)$ denotes the surface roughness at the point x, $(0 \leq x < 2\pi)$ on the surface of the ith ball and ρ is the constant representing the amplitude ratio of the following two impulses: one excited by the collision between a ball and the inner race and the other excited by the collision between the ball and the outer race. Since it is difficult to determine the value of ρ, it is impossible to distinguish the surface roughness on the points π radian apart on the ball surface. Thus, the following new surface function $s_i'(x)$ is introduced

$$s_i'(x) = s_i(x) + \rho \cdot s_i((x + 2\pi)_{2\pi}),$$

and each ball surface $s_i'(x)$ is defined in the range $0 \leq x < \pi$. Then, the equivalent driving impulse $z(t:\theta_0)$ of equation (12) is expressed as follows

$$z(t:\theta_0) = \sum_{i=0}^{N-1} W((\theta(i,t:\theta_0))_{2\pi}) \cdot s_i'((\omega_B t)_{\pi}) \tag{12b}$$

In the discrete time system, by letting the integer M' be the number of points on the surface of a ball in the range $0 \leq x < \pi$, the equivalent driving impulse $z(n:\theta_0)$ of equation (12*b*) is represented as follows

$$z(n:\theta_0) = \sum_{i=0}^{N-1} W(([\theta_d(i,n:\theta_0)]_M)) \cdot s_i'((2M' \cdot F_B \cdot nT_s)_{M'}), \tag{12c}$$

where $2\pi F_B = \omega_B$.

The surface roughness $s_i'(x)$ on the ith rolling ball is determined directly from the vibration signal $y(n)$ by applying a least mean square technique as described previously in Section 2.2.

2.4 Determination of the Parameters Used in the Proposed Method. As described in Section 2.1, it is necessary to determine the following three parameters for the estimation of the surface roughness: (*a*) the values of the angular velocity ω_I and ω_b of the inner ring and balls, (*b*) the weighting function $W(\theta)$, and (*c*) the inverse filter $h(n)^{-1}$. By using the two-pulse model proposed in the separate paper [5, 6], the inverse filter $h(n)^{-1}$ is determined as follows. The vibration caused by a flaw with a simple configuration on the inner ring is represented by the sum of the two resonant vibrations driven by two impulses. The two impulses are excited on the following sequences, that is, the first impulse is excited when a ball comes in contact with the flaw on the inner ring and hits the inner ring. Then, the ball is reflected and the second impulses is excited when the ball hits the outer ring on the other side.

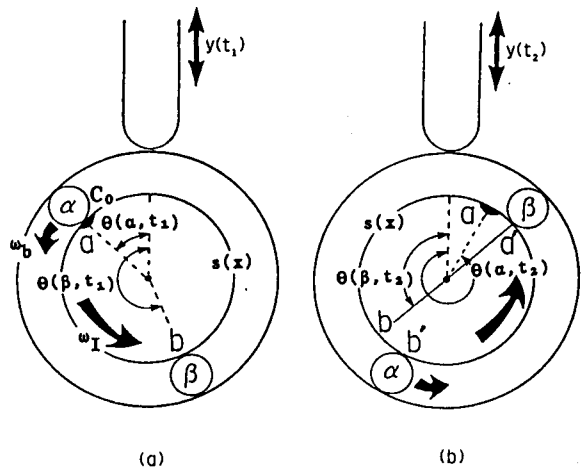

Fig. 6 Illustration of the simplified ball bearing showing the effect of the irregularity in the ball intervals on the estimation of the surface roughness: (a) at a time $t=t_1$; (b) at a time $t=t_2$.

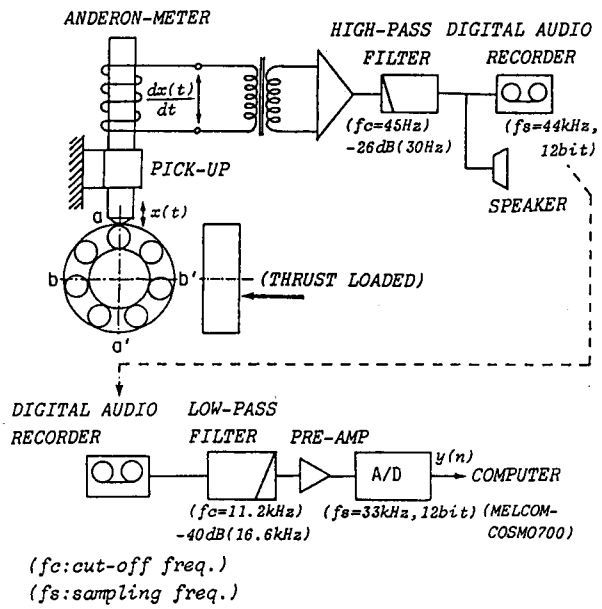

Fig. 7 The block diagram of the procedure for measuring the vibration signal of a ball bearing.

Therefore, the modeled vibration signal is expressed as a linear combination of complex exponentials driven by the two pulses. By applying the linear least mean square fitting technique to the minimization of the difference between the model signal and the observed vibration $y(n)$ for various lags, the parameters of the poles of the vibration are estimated. From these parameters, the inverse filter $h(n)^{-1}$ is calculated.

When the inner ring is revolved at a constant speed of $\omega_I=2\pi F_I$, the rotation speed $\omega_b=2\pi F_b$ and the revolution speed $\omega_B=2\pi F_B$ of the balls are theoretically derived from the following equations [2, 3]

$$F_b = \frac{F_I}{2}\left\{1-\frac{d}{D}\cos\delta\right\}, \tag{13a}$$

$$F_B = \frac{F_I}{2}\cdot\frac{D}{d}\left\{1-\left(\frac{d}{D}\right)^2\cos^2\delta\right\}, \tag{13b}$$

where d is the ball diameter, D the pitch diameter, and δ the contact angle.

The weighting function $W(\theta)$ is determined from the vibration signal. It is difficult to measure the weighting function $W(\theta)$ using the vibration caused by the flaw on the rolling ring because it is necessary to know the exact location of the flaw and the time when the flaw contacts with a ball. Therefore, by using the vibration due to a flaw on the outer ring, the weighting function $W(\theta)$ is determined as follows: When the outer ring is fixed in an Anderon meter so that the flaw on the race is set at one of M_0 angles $\{k\cdot\theta=2\pi k/M_0\}$, $(k=0,1,2,\ldots,M_0-1)$ relative to the pick-up, the vibration signal $y(n:k)$ caused by the flaw is detected as described in Section 4. The vibration $y(n:k)$ is band-limited so that the significant frequency band around the central frequency of the resonant vibration is selected [4]. The envelope signal is obtained by squaring the narrow band-passed signal and passing it through a low-pass filter. The periodicity of the vibration $y(n:k)$ caused by the flaw at the angle $k\cdot\theta$ is detected from the envelope signal. The magnitude $P(k)$ of the peak at the repetition frequency is calculated from the power spectrum of the vibration signal $y(n:k)$. By varying the angle $k\cdot\theta$ of the flaw relative to the pick-up, the magnitudes $\{P(k)\}$, $(k=0,1,2,\ldots,M_0-1)$ of the peaks are obtained from the above processes $(M_0=16)$. Then, the squared weighting function $|W(k)|^2$ is calculated from the peaks $\{P(k)\}$ divided by $P(0)$ as follows

$$|W(k)|^2=P(k)/P(0). \text{ (for } k=0,1,2,\ldots,M_0-1) \tag{14}$$

By using the digital interpolation process [8] to increase the number M_0 of the points, the interpolated weighting function $|W(k)|^2$, $(k=0,1,2,\ldots,M-1)$ is obtained at M points $(M_0<M=64)$ on the circumference as shown in Fig. 5.

3 Reduction of the Effect of the Irregularity in the Ball Intervals

3.1 The Effect of the Irregularity in the Ball Intervals on the Estimation of the Surface Roughness. It is known that there is scattering in the repetition intervals of the flaw pulses generated by many balls [4] due to the following two reasons: (*a*) the diameter of the ball is a little smaller than that of the space in the cage and (*b*) the intervals of the balls on the race vary every moment due to the slipping of the balls in the cage. In this section, by using the simple model of the ball bearing, the effect of the irregularity in the ball intervals on the estimation of the surface roughness is explained as follows:

Figure 6 shows the simplified ball bearing having two balls α and β and they are not π radian apart on the race. At a time $t=t_1$, the balls α and β contact with the points a and b on the race, respectively (see Fig. 6(*a*)). When the angle between α and β is π, the balls α and β contact with the points b and a as described in the preceding with Fig. 1(*b*). However, in the practical case, the balls α and β do not contact with the points b and a, respectively, at the time $t=t_2$ as shown in Fig. 6(*b*) since the two balls are not π radian apart.

Let the inner race have a flaw only on the point a as shown in Fig. 6. Then, the surface roughness $s(x)$ is described as

$$s(x)=\begin{cases}C_0 & (x=a)\\ 0 & (x\neq a).\end{cases} \tag{15}$$

The equivalent driving impulses $z(t_1)$ and $z(t_2)$ of the left-hand side of the set of simultaneous equation (5) are expressed as follows

$$W(\theta(\alpha,t_1))\cdot C_0 = W(\theta(\alpha,t_1))\cdot s(a)+W(\theta(\beta,t_1))\cdot s(b), \tag{16a}$$

$$0 = W(\theta(\beta,t_2))\cdot s(a)+W(\theta(\alpha,t_2))\cdot s(b). \tag{16b}$$

From these equations, the surface roughness quantities $s(a)$ and $s(b)$ are estimated as follows

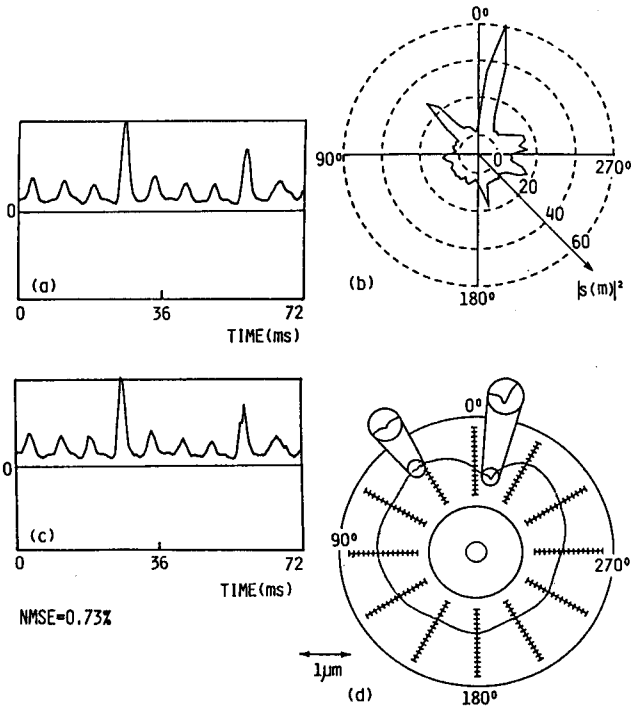

Fig. 8 Experimental results of the estimated surface roughness, the synthesized vibration signal and the surface roughness measured directly by a stylus: (a) the envelope signal $z_L(n)$ of the vibration signal $y(n)$ generated by flaws on the inner ring; (b) estimated surface roughness $s(m)$, ($m = 0,1,2,\ldots,63$) from the envelope signal $z_L(n)$; (c) the envelope signal $\underline{z}_L(n)$ synthesized from the estimated surface roughness $s(m)$, the weighting function $|W(\theta)|^2$ and the initial angle θ_0; (d) the profile of the surface measured using a stylus (cross-sectional view of the inner ring of the sample used).

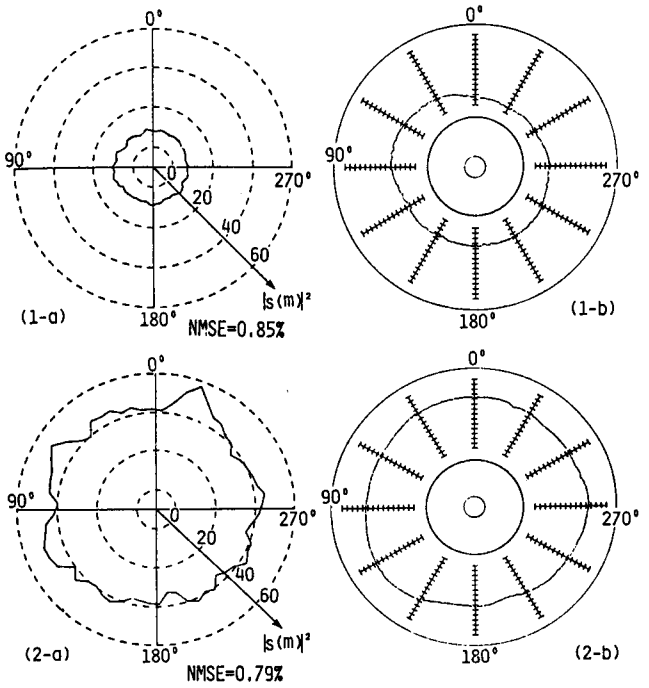

Fig. 9 Experimental results of the estimated surface roughness and surface profiles of inner races in ball bearings measured by a stylus: (1) for a normal bearing having smooth races; (2) for a bearing having a rough inner race; (a) the surface roughness estimated by the proposed method; (b) the surface profile measured by a stylus.

$$\underline{s}(a) = W(\theta(\alpha,t_1))\cdot W(\theta(\alpha,t_2))\cdot C_0/D$$

and

$$\underline{s}(b) = -W(\theta(\alpha,t_1))\cdot W(\theta(\beta,t_2))\cdot C_0/D, \tag{17}$$

where

$$D = W(\theta(\alpha,t_1))\cdot W(\theta(\alpha,t_2)) - W(\theta(\beta,t_1))\cdot W(\theta(\beta,t_2)),$$

and $\underline{s}(a)$ and $\underline{s}(b)$ are the estimated values of $s(a)$ and $s(b)$, respectively. Thus, the true value of the surface roughness expressed by equation (15) cannot be obtained due to the irregularity of the ball intervals. Therefore, it is necessary to decrease the sampling rate to reduce the effects of the irregularity, as described in the following proposed method.

3.2 Estimation of the Surface Roughness Based on the Envelope Signal of the Squared Magnitude of the Narrow Band Signal. To reduce the effect of the irregularity in the ball intervals, we propose a new reliable method. The signal $z(n)$ is band limited. Then, the envelope signal $z_L(n)$ is obtained by squaring the narrow band-passed signal $z_B(n)$ and passing it through a low-pass filter. The surface roughness is estimated from the envelope signal $z_L(n)$. The details of this processing are explained below.

1 The signal $z(n)$ is band limited using a band-pass filter. If there is no scattering in the contact intervals, the equivalent driving impulse of equation (4*b*) at a time t_0 is described as follows:

$$z(n{:}\theta_0) = \sum_{i=0}^{N-1} W(([\theta_d(i,n{:}\theta_0)]))_M \cdot s(([x_d(i,n{:}\theta_0)]))_M \cdot \delta(nT_s - t_0), \tag{18a}$$

where the unit impulse $\delta(n)$ is defined in the discrete time system as follows

$$\delta(n) = \begin{cases} 1 & (n=0) \\ 0 & (n\neq 0). \end{cases}$$

However, when there is time lag $\Delta(i,t_0)$ for the impulse generated by the collision between the ith ball and the race at a time t_0, the above impulse $z(n)$ is modified as follows

$$z(n{:}\theta_0) = \sum_{i=0}^{N-1} W(([\theta_d(i,n{:}\theta_0)]))_M \cdot s(([x_d(i,n{:}\theta_0)]))_M \cdot \delta(nT_s - t_0 - \Delta(i,t_0)) \tag{18b}$$

Since the band-pass filter $h_B(n)$ is described by the product of the low-pass filter $h_L(n)$ and the cosine signal of the frequency F_0 as follows

$$h_B(n) = h_L(nT_s)\cdot\cos(2\pi F_0 \cdot nT_s),$$

the band-passed signal $z_B(n)$ around t_0 is obtained as follows

$$z_B(n) = \sum_{i=0}^{N-1} W(\theta_i)\cdot s(x_i)\cdot h_L(nT_s - t_0) - \Delta(i,t_0)) \cdot\cos\{2\pi F_0(nT_s - t_0 - \Delta(i,t_0))\}, \tag{19}$$

where

$$W(\theta_i) = W(([\theta_d(i,n{:}\theta_0)]))_M,$$

and

$$s(x_i) = s(([x_d(i,n{:}\theta_0)]))_M.$$

2 The squared signal $|z_B(n)|^2$ of the band-passed signal $z_B(n)$ is calculated.

3 The envelope signal $z_L(n)$ is obtained by passing the squared signal through a low-pass filter as follows

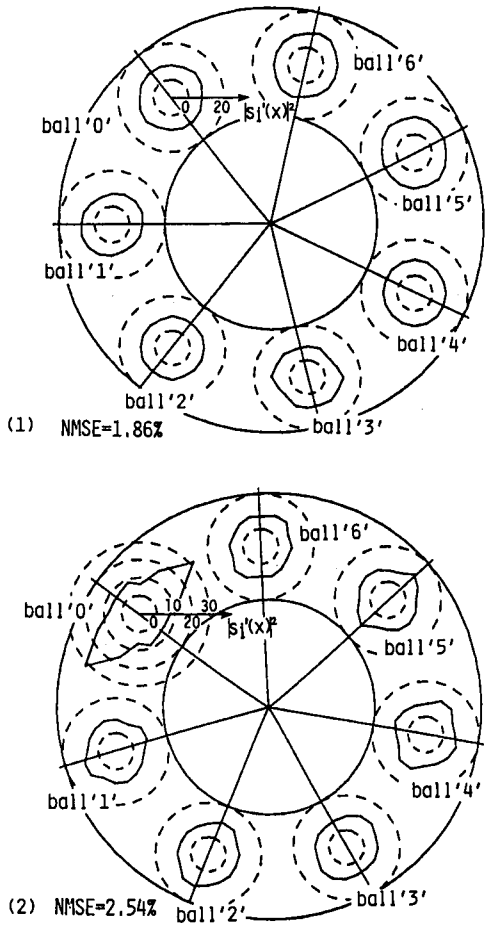

Fig. 10 Experimental results of the estimated surface roughness of the balls of ball bearings. (1) for a normal bearing. (2) for a bearing having a flaw on the ball.

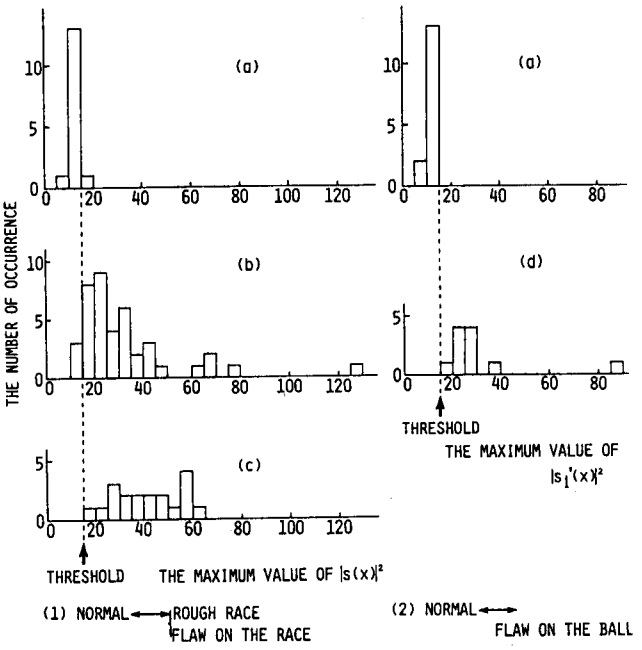

Fig. 11 The distribution of the maximum values of the estimated surface roughness $|s(x)|^2$ of the inner race and the estimated surface roughness $|s_l'(x)|^2$ of the balls and thresholds for sorting bearing samples: (a) 15 normal samples; (b) 41 samples having flaws on the inner races, (c) 19 samples having rough races; (d) 11 samples having flaws on the balls.

$$z_L(n) = [\text{low frequency components of } |z_B(n)|^2]$$

$$= \sum_{i=0}^{N-1} \{W(\theta_i)\cdot s(x_i)\cdot h_L(nT_s - t_0 - \Delta(i,t_0))\}^2/2$$

$$+ \sum_{i=0}^{N-1}\sum_{\substack{j=0 \\ (i<j)}}^{N-1} W(\theta_i)\cdot W(\theta_j)\cdot s(x_i)\cdot s(x_j)$$

$$\cdot\cos\{2\pi F_0\,|\Delta(i,t_0) - \Delta(j,t_0)|\}$$

$$\cdot h_L(nT_s - t_0 - \Delta(i,t_0))\cdot h_L(nT_s - t_0 - \Delta(j,t_0)) \qquad (20)$$

In this equation, the envelope signal $z_L(n)$ has two fluctuation terms: (*a*), the time lag $\Delta(i,t_0)$ involved in the low-passed signal, and (*b*), the constant $\cos\{2\pi F_0\,|\Delta(i,t_0) - \Delta(j,t_0)|\}$ involved in the second term. However, the effects of both the terms can be disregarded due to the following reasons:

(*a*) Since the low-passed signal does not involve the high-frequency components, the phase shift due to the time lag $\Delta(i,t_0)$ in the first term is negligible.

(*b*) Since the central frequency F_0 of the band-pass filter is high, $2\pi F_0\cdot|\Delta(i,t_0) - \Delta(j,t_0)|$, $(i \neq j)$ in the second term takes a large value. However, when a ball bearing has N balls, there are a number of combinations of the term in the second term. Since there is no correlation between $\Delta(i,t)$ and $\Delta(j,t)$ $(i \neq j)$ because the irregularity of the ball intervals is generated by the slipping of the balls in the cage, the constant $\cos\{2\pi F_0\,|\Delta(i,t_0) - \Delta(j,t_0)|\}$ $(i \neq j)$ in the second term takes a random value in the range of $-1 \sim +1$. Then, the component value of the second term is smaller than that of the first term.

Therefore, the low frequency component of equation (20) is expressed approximately as follows

$$z_L(n) \simeq |h_L(nT_s - t_0)|^2 \sum_{i=0}^{N-1} |W(\theta_i)|^2\cdot|s(x_i)|^2/2$$

$$(\text{for } nT_s \text{ around } t_0) \qquad (21)$$

From equation (21), the envelope signal $z_L(n)$ denotes the sum of the squared signal of the impulses generated by N balls. Equation (21) is the same as equation (4*b*) except for the squared low-pass filter $|h_L(n)|^2$: that is, $(|h_L(n)|^2 \cdot |W(\theta)|^2/2)$ and $|s(x)|^2$ are used instead of $w(\theta)$ and $s(x)$, respectively.

Thus, by neglecting the characteristic of the low-pass filter, the squared surface roughness $|s(x)|^2$ is estimated accurately from the envelope signal $z_L(n)$ of the squared magnitude of the narrow band signal. The squared surface roughness $|s(x)|^2$ obtained here is in proportion to the $(6/5)^2$th power of the width of the flaw on the surface as described in Section 2.1.

4 Apparatus for the Experiments

Figure 7 shows a block diagram of the procedure for measuring the vibration signal of a ball bearing [4, 5]. The outer ring is fixed by imposing an axial pressure, and the inner one is revolved at a constant speed of 1,800 rpm in an Anderon meter [1]. Under such the conditions, surface defect causes a radial movement of the outer ring and the signal resulting from this movement is picked up by a vibration pick-up attached to the outer ring. The signal is amplified and filtered through a high-pass filter to attenuate the primary frequency component ($F_l = 30$Hz) corresponding to the rotation of the inner ring. The filtered signal is stored in a high-fidelity digital audio tape. In the laboratory, the signal is played back and is A/D converted with a 12 bit A/D converter at a sampling period T_s of 30 μs.

Table 1 Sorting results of ball bearings

(1) Sorting results based on the surface roughness of the inner race estimated by the proposed method in Section 2.2.

IN \ OUT	The number of samples recognized to be normal	The number of samples recognized to be inferior	Total
Normal	14	1	15
Flaw on the inner ring	3	38	41
Rough race	0	19	19

(2) Sorting results based on the surface roughness of the balls estimated by the proposed method in Section 2.3.

IN \ OUT	The number of samples recognized to be normal	The number of samples recognized to be inferior	Total
Normal	15	0	15
Flaw on the ball	0	11	11

Total number of used samples = 86

Recognition rate = $(14+38+19+11)/(15+41+19+11)*100=95.3\%$

5 Experiments and the Results

To estimate the surface roughness based on the proposed method in the practical case, the following processes are carried out in the following sequences:

1 The squared weighting function $|W(k)|^2$, $(k=0,1, \ldots,M-1)$ is obtained at M points on the race ($M=64$) by the method described in Section 2.4.

2 The inverse filter $h(n)^{-1}$ is calculated from the impulse response of the resonant vibration estimated using the two-pulse model as described shortly in Section 3.4. The equivalent driving impulse $z(n)$ is obtained by convoluting $h(n)^{-1}$ with the observed vibration signal $y(n)$.

3 The signal $z(n)$ is band-limited by using the FIR narrow-band pass filter (16 points), and the narrow-band signal $z_B(n)$ having the components around the central frequency of the resonant vibration is obtained [4]. The band-limited signal $z_B(n)$ is squared and moving-averaged using a 16-point Hanning window so that only low-frequency components are extracted. Then, an envelope signal $z_L(n)$ is generated by picking up one sample from every third segment.

4 The squared surface roughness $|s(m)|^2$ of equation (21) is estimated from the envelope signal $z_L(n)$ using the squared weighting function $|W(k)|^2$.

5 The unknown value of the initial angle θ_0 of equation (4c) is decided as follows: Since the initial angle θ_0 has a value in the range of $0 \leq \theta_0 < 2\pi/N$, where N denotes the number of ball, the total squared error $\alpha(\theta_0)$ in equation (8) is calculated for various θ_0 within the range. The optimum angle of θ_0 is decided so that the squared error $\alpha(\theta_0)$ is minimum. Using the optimum angle of θ_0, the surface roughness is obtained.

Figure 8 shows the experimental results obtained by the proposed method for a ball bearing (JIS696) having flaws on the inner ring. Figure 8(*a*) shows the envelope signal $z_L(n)$ calculated from the vibration signal. The squared surface roughness $|s(m)|^2$, $(m=0,1,2, \ldots,63)$ estimated from the envelope signal is shown in Fig. 8(*b*). The two main flaws are observed in this figure in the direction of 40 deg and 350 deg.

By using the estimated surface roughness $|s(m)|^2$ in Fig. 8(*b*), the weighting function $|W(k)|^2$ in Fig. 5 and initial angle $\theta_0(\sim 5.1$ deg), the envelope signal $\underline{z}_L(n)$ of the vibration is synthesized as shown in Fig. 8(*c*). Since the synthesized envelope signal $\underline{z}_L(n)$ just coincides with the envelope signal $z_L(n)$ shown in Fig. 8(*a*), it is found that the surface roughness is estimated accurately. The synthesized envelope signal $\underline{z}_L(n)$ obtained by the proposed method is evaluated by the following normalized mean square error (NMSE) as

$$\text{NMSE}=\frac{\sum_n (z_L(n)-\underline{z}_L(n))^2}{\sum_n z_L(n)^2} \tag{22}$$

and we find that the value of the NMSE for the above synthesized vibration is about 0.73 percent (−21dB). Figure 8(*d*) shows the surface roughness of the inner race measured directly by a stylus. The positions of the two main flaws in Fig. 8(*d*) agree with those of the two main flaws in Fig. 8(*b*).

Figure 9 shows the estimated surface roughness and the measured surface profile of a normal bearing (1) and a ball bearing with a rough inner race (2). The result (2−*a*) estimated by the proposed method shows that the surface is rough on the all over the race. However, such roughness is not detected clearly using a stylus as shown in Fig. 9(2*b*) because the radius of the measured profile in Fig. 9(1*b*) or 9(2*b*) does not denote the surface roughness.

Figure 10 shows the ball surface roughness $|s_i'(x)|^2$ estimated by the proposed method as described in Section 2.3. The position and the rotation angle of each ball at a time $t=0$ is obtained by the proposed method as shown in these figures. Figure 10(2) shows that the 0th ball of the ball bearing has a large flaw on the surface.

6 Sorting of Ball Bearings

Figure 11 shows the distributions of the maximum values of the squared surface roughness $|s(m)|^2$ of the inner race and $|s_i'(m)|^2$ of the balls, which are estimated from the vibration of 86 ball bearings. At first, these ball bearings were classified into the following three categories aurally by a skilled inspector: 15 normal bearings, 41 ball bearings having flaws on the inner race, 19 ball bearings having rough races, and 11 ball bearings having flaws on the balls. Table 1 shows the results of sorting the ball bearings using the thresholds shown in Figs. 11(1) and 11(2). The number of correctly sorted samples and that of incorrectly sorted samples are 82 and 4, respectively. The ball bearings are sorted correctly at the rate of 95.3 percent comparing with the results by human auditory sense.

7 Conclusions

A new method, utilizing vibration signals, is proposed for the estimation of the surface roughness of the races of the rolling ring or the surface roughness of the balls in ball bearings.

The degree of the surface roughness and the number and the width of the main flaws on the races are obtained from the estimated surface roughness. It is confirmed with the experimental results presented herein that the roughness estimated by the proposed method agrees with that measured directly by using a stylus. Based on the estimated surface roughness obtained from the short length vibration signal, ball bearings having inferior races are detected by the proposed method with a 95.3 percent accuracy rate. The surface estimation of a ball bearing having flaws on the inner race, on the outer race and on the balls is the subject for a future study.

References

1 "Anderon meter," Technical Bulletin, The Bendix Corporation.

2 Igarashi, T., "Sound of Rolling Bearings," (in Japanese), *Lubrication*, Vol. 22, No. 12, 1977, pp. 751-756.

3 Braun, S., and Datner, B., "Analysis of Roller/Ball Bearing Vibrations," ASME *J. Mech. Design*, Vol. 101, 1979, pp. 118-125.

4 Kanai, H., Abe, M., and Kido, K., "Detection and Discrimination of Flaws in Ball Bearings by Vibration Analysis," *J. Acoust. Soc. Jpn.*, Vol. (E)-7, No. 2, 1986, pp. 121-131.

5 Kanai, H., Abe, M., and Kido, K., "Detection of Slight Defects in Ball Bearings by Non-periodic Analysis," *J. Acoust. Soc. Jpn.*, Vol. (E)-7, No. 4, 1986, pp. 219-228.

6 Kanai, H., and Abe, M., and Kido, K., "Detection of Slight Defects in Ball Bearings Using the Resonant Vibration Estimated by the Two-Pulse Model," *Proc. of 1986 International Conference on Noise Control Engineering*, July 1986, Cambridge, Ma., pp. 1191-1196.

7 Nishimura, G., and Takahashi, K., "Ball Bearing Noise," (in Japanese), *J. Jpn. Soc. Precision Eng.*, Vol. 30, 1964, pp. 475-489.

8 Rabiner, L. R., and Schafer, R. W., *Digital Processing of Speech Signal*, Prentice-Hall, Englewood Cliffs, 1978, Ch. 2.4.