Performance evaluation of frequency diverse Bayesian ultrasonic flaw detection

Jafar Saniie, Tao Wang, and Xiaome Jin
Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, Illinois 60616

(Received 26 October 1990; revised 26 August 1991; accepted 6 November 1991)

The performance of an ultrasonic flaw detection system is valued by the success of differentiating the flaw echoes from those echoes scattered by microstructures (e.g., grains in polycrystalline metals). Microstructure noise (clutter) is stationary and will only vary when the frequency of the ultrasonic beam is changed. In this paper, frequency diverse microstructure information is obtained through the split-spectrum processing of broadband ultrasonic backscattered echoes. The statistical difference between clutter and flaw echoes over different frequency bands has led to the development of a Bayes classifier that is quadratic and can incorporate the correlation properties of scattered echoes. The performance of the Bayes classifier has been examined for both experimental and computer simulated data, and compared to other commonly used techniques such as mean, minimum, median, and polarity detectors. The Bayes classifier shows superior performance, and the flaw-to-clutter visibility has been improved in the range of 5–15 dB when the measured flaw-to-clutter ratio is 0 dB or less. These results confirm the feasibility of grain noise suppression by frequency diversity with a subsequent statistical processor for flaw detection.

PACS numbers: 43.40.Le

INTRODUCTION

Defect detection by ultrasonic techniques has been proven to be an effective means to assure the quality of materials nondestructively. However, the performance of ultrasound is limited when the level of echoes from the surrounding unwanted reflectors (e.g., grains in polycrystalline metals) is comparable to or larger than that of the defect signal (i.e., target to clutter ratio is less than 0 dB). Existing conventional techniques, such as averaging, are only capable of removing time-varying echoes and are ineffective when the clutter is time invariant or coherent in nature. Frequency diversity based on shifting the frequency of the transducer can decorrelate coherent noise and proper postdetection processing will result in flaw-to-clutter ratio enhancement.[1–4]

A practical method in implementing frequency diversity is by transmitting a broadband echo through the materials and bandpass filtering the received echoes over many subbands of frequencies.[1,2] Since the microstructure in the detection cell consists of many unresolved and randomly distributed reflectors, the detected echoes exhibit randomness in amplitude and are sensitive to shifts in the transmitted frequency. In contrast, flaws are often larger in size and are less vulnerable to variation in the transmitted frequency. In general, flaw echoes exhibit different scattering distributions as a function of frequency when compared with microstructure scattering. Therefore, at any given time, the outputs of bandpass filters can be represented as a random feature vector that contains information related to flaw and microstructure echoes. The statistical approach for flaw detection, utilizing the statistical properties of this feature vector, is to design flaw detection algorithms.

In recent years, statistical methods have been used for ultrasonic flaw detection and improved ultrasonic medical imaging.[3–10] In particular, minimum detection and polarity thresholding have been studied extensively and have been shown to be effective for flaw detection in metals.[3,4] A suboptimal flaw detection algorithm, using flaw and microstructure scattering models, has been applied to a circular crack in a plastic test specimen.[5–8] In medical imaging related to breast tissue classification, ultrasound has been used to generate a large set of potentially useful features sensitive to benign and malignant tumors.[9,10]

In this paper, we focus on the use of the Bayes classifier designed for flaw detection and microstructure noise discrimination based on information collected from the output of bandpass filters of the split-spectrum processor (SSP). The performance of the Bayes classifier is examined using both computer simulation and experimental results.

I. THEORY OF OPTIMAL DETECTION

The performance of all ultrasonic flaw detection systems is limited by the presence of microstructure noise. Although the system output signal contains information related to flaw structure, this information is often masked by unwanted echoes caused by microstructure scattering. Since the output signals from the bandpass filters of the split-spectrum processor represent information related to the microstructure or flaw, in principle, it should be possible to extract the flaw echo from those random, unwanted reflecting clutter echoes (i.e., microstructure noise).

An effective method of obtaining frequency diverse information is through splitting the spectrum of broadband echoes Figures 1 and 2 summarize the steps involved in the

 © 1992 Acoustical Society of America

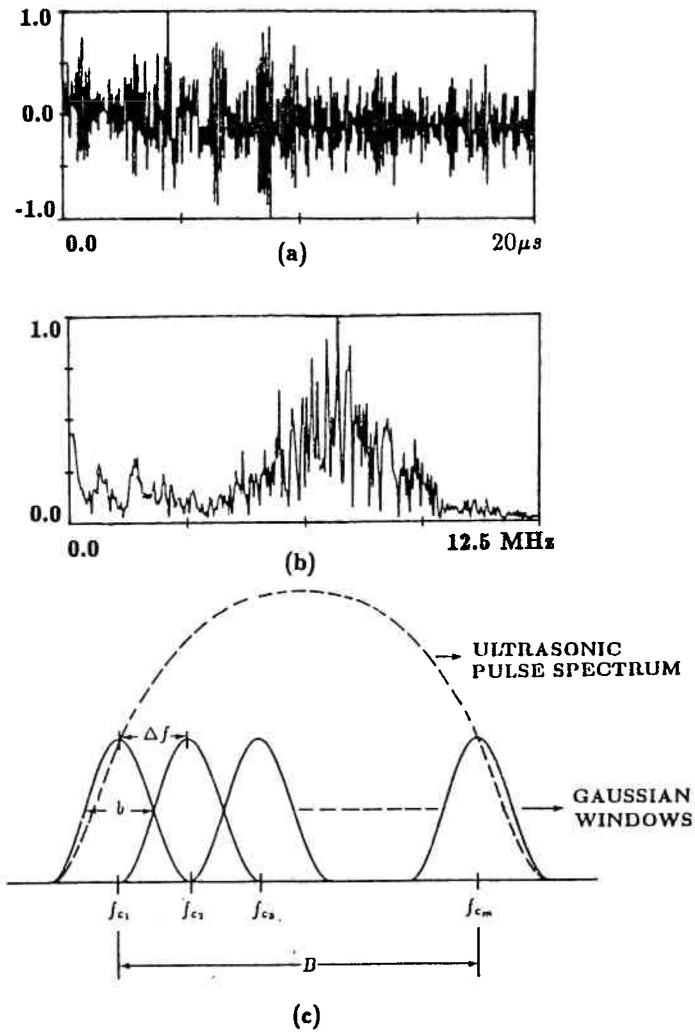

FIG. 1. Split-spectrum processing (SSP) technique; (a) an ultrasonic back-scattered microstructure signal, (b) broadband spectrum of the signal, and (c) split-spectrum windows.

frequency diverse Bayes flaw detection system. Figure 1(a) and (b) presents an example of a broadband grain signal and its power spectrum. Figure 1(c) displays graphically the split-spectrum schemes using Gaussian windows. Since the power spectrum amplitude of the received echoes is not uniformly distributed, we use a set of weighting coefficients α_1, $\alpha_2,\ldots,\alpha_K$, in order to obtain the equally powered output signals. Then, the weighted, filtered outputs are passed to the Bayes classifier for optimal detection. The entire split-spectrum Bayes classification procedure is illustrated in Fig. 2.

The design of the ultrasonic statistical flaw detection system involves feature selection and the formation of a discriminant function. The feature vector at any given time t can be represented as

$$Z = [z_1, z_2, \ldots, z_K]^T, \quad (1)$$

where $z_i = r_i(t)$ and is the ith filter's output signal after normalization, and K is the total number of bandpass filters used in the split-spectrum technique. The design of the discriminant function can be accomplished with the Bayes decision rule which is optimum in the sense of minimizing the probability of error.[11] The Bayes classifier in the flaw detection problem is concerned with the following hypotheses at a specific time t_n:

H_0: Flaw plus clutter echoes,

H_1: Clutter echoes.

Using the above hypothesis, the criterion for decision making can be represented by

$$\phi_{\text{op}}(Z) = \frac{p(Z/H_0)}{p(Z/H_1)} > \frac{P(H_1)}{P(H_0)} \rightarrow H_0, \quad (2)$$

$$\phi_{\text{op}}(Z) = \frac{p(Z/H_0)}{p(Z/H_1)} < \frac{P(H_1)}{P(H_0)} \rightarrow H_1, \quad (3)$$

where $\phi_{\text{op}}(Z)$ is the likelihood ratio which serves as the discriminant function for classification, $P(H_0)$ is the probability of the presence of flaw, $P(H_1)$ is the probability of the absence of flaw, and $P(H_1)/P(H_0)$ is the detection threshold. The term $p(Z/H_0)$ is the probability density function of flaw-plus-clutter echoes, and $p(Z/H_1)$ is the probability density function of the clutter echoes.

A major barrier faced in the design of the discriminant function is the estimation of the *a priori* probability density function of each class. We examined the histogram of the echoes at the output of each bandpass filter, and they were found to be Gaussian in shape with reasonable accuracy (see

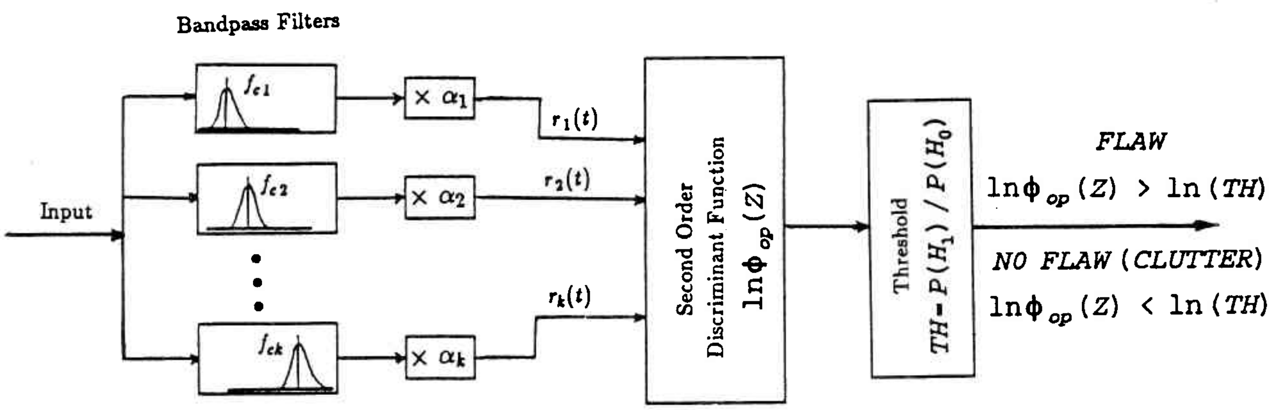

FIG. 2. The block diagram of the split-spectrum Bayes flaw detector.

Fig. 3). Furthermore, we have observed that an insignificant correlation exists among the elements of feature vectors. This has led to the assumption that the features from bandpass filters are jointly normal with different mean vectors and covariance matrices. Hence, the *a priori* probability density function of the feature vector becomes

$$p(Z,M_i,\Sigma_i) = (1/2\pi)^{K/2}|\Sigma_i|^{-1/2} \times \exp[-\tfrac{1}{2}(Z-M_i)^T\Sigma_i^{-1}(Z-M_i)]; \quad i=0, \text{ or } 1, \tag{4}$$

where M_i is the mean vector and Σ_i is the covariance matrix for the hypotheses H_0 or H_1, which can be represented as follows:

$$M_i = (m_{i_1}, m_{i_2}, \ldots, m_{i_K})^T \tag{5}$$

and

$$\Sigma_i = \begin{bmatrix} \sigma^2_{i_{11}} & \sigma^2_{i_{12}} & \cdots & \sigma^2_{i_{1K}} \\ \sigma^2_{i_{21}} & \sigma^2_{i_{22}} & \cdots & \sigma^2_{i_{2K}} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \sigma^2_{i_{K1}} & \sigma^2_{i_{K2}} & \cdots & \sigma^2_{i_{KK}} \end{bmatrix}. \tag{6}$$

For a normal distributed feature vector, it is more convenient to write the discriminant function in the log form,

$$\ln \phi_{op}(Z) = -\tfrac{1}{2}[(Z-M_0)^T\Sigma_0^{-1}(Z-M_0)] + \tfrac{1}{2}[(Z-M_1)^T\Sigma_1^{-1}(Z-M_1)] + \tfrac{1}{2}\ln(|\Sigma_0|/|\Sigma_1|). \tag{7}$$

The above equation is a second-order discriminant function. With the assumption that the hypotheses H_0 and H_1 are equally probable, the final decision rule is achieved by

$$\ln \phi_{op}(Z) > 0 \rightarrow H_0, \quad \ln \phi_{op}(Z) < 0 \rightarrow H_1. \tag{8}$$

It is important to point out that the hypotheses H_0 and H_1 may not be equally probable and the threshold value may not be a zero. In practice, careful examination of the discriminant function will reveal the optimal value for threshold selection.

II. COMPUTER SIMULATION AND DISCUSSION

The purpose of computer simulation is to form a pseudoclutter signal such that the performance of the Bayes classifier can be evaluated for different flaw-to-clutter ratios. The pseudo-clutter signal is generated by superimposing 512 Gaussian shape echoes with Gaussian-distributed amplitude at the uniformly distributed positions. These Gaussian echoes have a 5-MHz center frequency and a bandwidth of 2.5 MHz. The sampling rate for the simulated signal is 200 MHz. After the superposition of all returned echoes, the signal is normalized between 1 and -1. The flaw echo was simulated by one single Gaussian echo with a desired amplitude such that the ratio of the flaw echo to the largest possible clutter echo has a ratio less than or equal to unity [i.e., flaw-to-clutter (F/C) ratio is about 0 dB]. Finally, superimposing the pseudoclutter signal and the flaw echo at a known flaw location results in the simulated signal. The split-spectrum processor was implemented in frequency by using a set of bandpass filters having a Gaussian spectrum.

To estimate the statistical parameters of the feature vector for the probability density function, we use the following estimates:[11,12]

$$m_{i_j} = \frac{1}{N}\sum_{n=1}^{N} r_{i_j}(t_n); \quad \text{for } i=0,1; \quad j=1,2,\ldots,K \tag{9}$$

and

$$\sigma^2_{i_{jl}} = \frac{1}{N}\sum_{n=1}^{N} r_{i_j}(t_n) r_{i_l}(t_n) - m_{i_j} m_{i_l}; \quad \text{for } i=0,1; \quad j=1,2,\ldots K; \quad l=1,2,\ldots K, \tag{10}$$

where n is the time index, i is the class index, j or l is the filter index and N is the total number of observed samples. These statistical parameters for the mean vector M_i and covariance matrix Σ_i are estimated using simulated clutter and flaw echoes. Once these parameters are obtained, the classifier is designed and tested with different clutter signal patterns and flaw-to-clutter ratios. An example of the output signal from the Bayes detection system is presented in Fig. 4. The top trace is the measured, unprocessed data and the lower trace is the plot of the discriminant function. As shown in this figure, the discriminant function clearly shows the flaw position and makes detection possible.

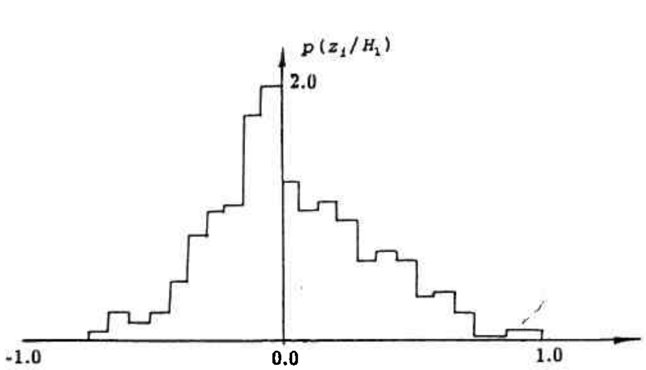

FIG. 3. An example of SSP filter output amplitude histogram.

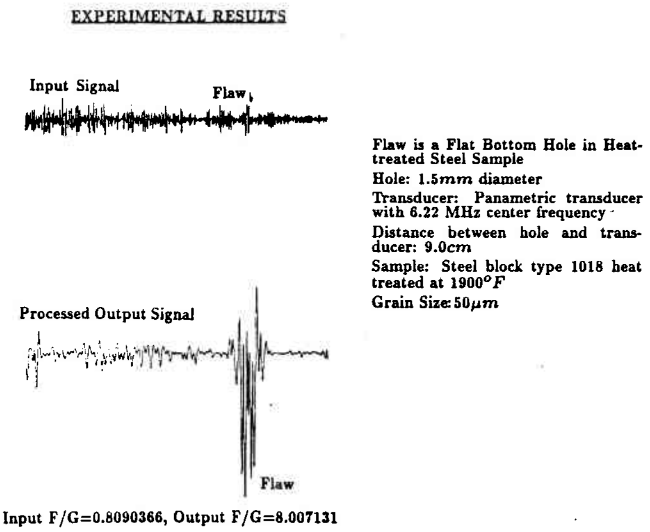

FIG. 4. An example of processed simulated output signal using Bayes flaw detector.

To observe the effect of processed results on the flaw-to-clutter (F/C) ratio, we use a window around a predicted position of flaws with the length of 0.6 μs (this width represents approximately three periods). Then, the flaw-to-clutter ratios are obtained by finding the ratio of the largest peak inside and outside the window. In the design of the Bayes classifier, careful evaluation of the split-spectrum parameters are necessary. In this study, we have examined these parameters in relation to correlation properties of the filter output. Increasing the number of bandpass filters will improve the performance of the Bayes classifier. This improvement is only possible to a certain extent when there is an insignificant correlation from band to band. Since the transmitted echoes cover only a finite band of frequency, a finite number of filters can be used for splitting the spectrum. We examined the correlation matrices of the clutter signal with and without flaws, and it has been observed that the output of the adjacent frequency bands are partially correlated. This correlation will increase if the number of filters is increased when the total frequency range covered by the filter banks remain unchanged. Increasing the number of frequency band demands higher computational time at slightly better, or sometimes no improvement, at the detector output. On the other hand, if the number of filters is too small, such that the filter bank cannot effectively cover the entire signal band, it will definitely lose some information related to clutter and flaw, and inevitably degrade the Bayes flaw detection performance. Figure 5 shows the comparison of Bayes detection performance via the number of filters. Figure 5(b)–5(d) shows the processed results using a filter bank with 18, 9, and 5 filters, respectively. In this filter design, the frequency range was from 0–14 MHz, with a 3-dB bandwidth of each filter is 1 MHz. Increasing the number of filters from 5 to 18 increases the number of observations, but, on the other hand, the correlations among the observations increase as well. Consequently, the performance of the discriminant function improves gradually. In this case, Fig. 5(c) (when $K = 9$, F/C = 20.80) is significantly better than Fig. 5(d) (when $K = 5$, F/C = 2.42). However, doubling the number of filters from 9 to 18 (when $K = 18$, F/C = 24.44) did not show much improvement in the processed output [compare Fig. 5(b) and 5(c)].

For a given number of filters in a filter bank, the bandwidth of each bandpass filter is kept to be the same, and this determines the degree of overlap in each frequency band. Large overlap introduces too much correlation into the outputs of the filter bank, and as a result, the sensitivity of clutter noise to the different frequency bands will be reduced, and consequently, the ability to detect flaw will be limited. If the overlap is small, or nonexistent (i.e., filter bandwidth is small compared to the step between the frequency bands), the correlation among the filters can be reduced to the very low level of a few percents. Improvement in the reduction of correlation is at the expense of reducing the number of observation bands that affects the detecting performance adversely. A comparison of the Bayes flaw detection via different bandwidths of filters is presented in Fig. 6. As shown in this figure, by increasing the bandwidth of the filters from 1 MHz [Fig. 6(b)] to 2 MHz [Fig. 6(c)], the performance of the Bayes detection is almost the same. However, by increasing the bandwidth, the width of output pulse will decrease, hence making detection more accurate.

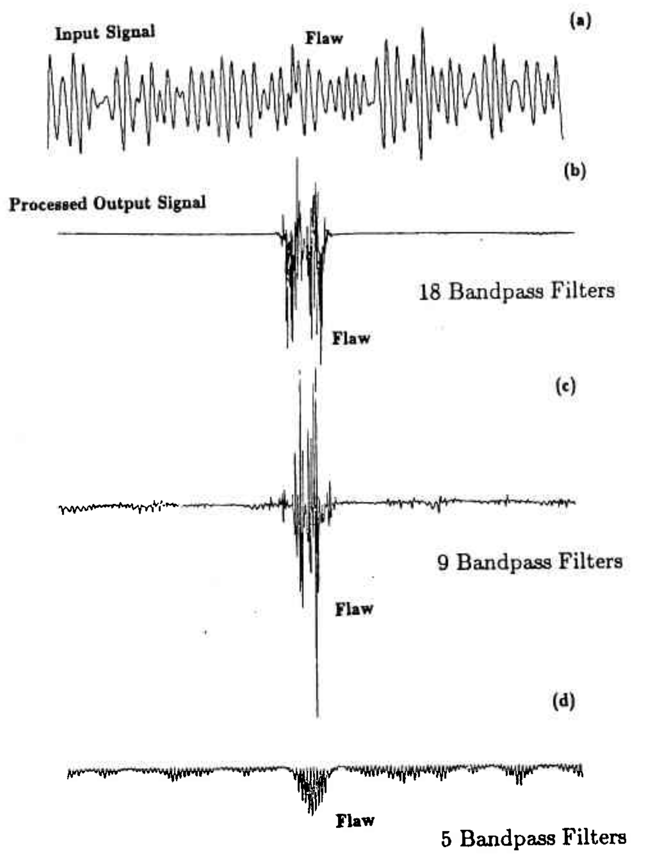

FIG. 5. A comparison of Bayes flaw detector performance via different number of filters (K) in SSP; (a) original signal, (b) $K = 18$, (c) $K = 9$, and (d) $K = 5$.

The choice of frequency range is a crucial factor in the performance of the Bayes detection algorithm. In detection

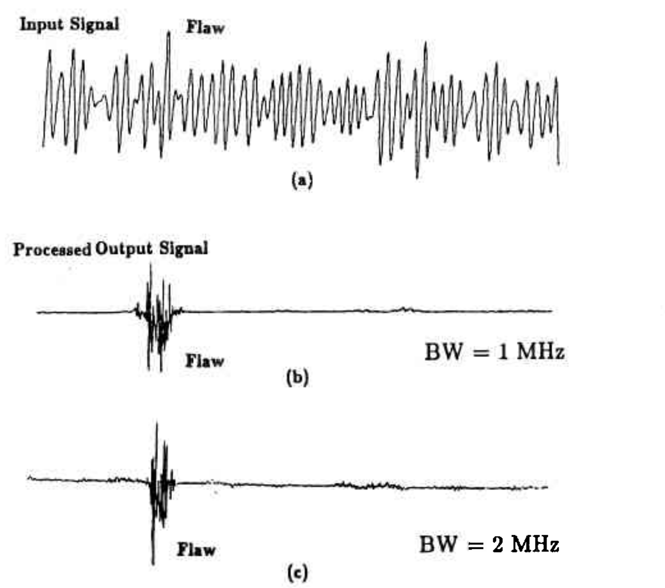

FIG. 6. A comparison of Bayes flaw detector performance via different bandwidth of SSP; (a) original signal, (b) BW = 1 MHz, and (c) BW = 2 MHz.

process, it is sensible to include those frequency ranges with a higher F/C ratio. In our previous investigation,[13,14] we found that the flaw echo has information mostly in the lower end of spectrum compared to clutter signal. Therefore, care must be taken to examine each frequency band using a calibration specimen prior to the application of the Bayes classifier. In order to demonstrate the importance of the frequency bands with a high F/C ratio, we have examined several frequency bands for flaw detection. Figure 7 shows the processed output for two different starting frequencies. Although the frequency range and the number of bandpass filters are the same, Fig. 7(c) has better performance than Fig. 7(b) since it covers the range of frequency where flaw-to-clutter ratio is the highest for the same number of filters.

III. COMPARISON OF FREQUENCY DIVERSE FLAW DETECTION ALGORITHMS

In this study, the Bayes classifier has been compared with other recently proposed frequency diverse flaw enhancement techniques such as averaging, the median detector, the minimization methods, and polarity thresholding.[1-3,15-19] The mathematical expressions of these techniques are as follows:

Average detector:

$$\text{Output signal: } \phi_{\text{av}}(t_n) = \frac{1}{K}\sum_{j=1}^{K} |r_j(t_n)|; \qquad (11)$$

Median detector:

Output signal:

$$\phi_{\text{med}}(t_n) = \text{median}[\,|r_j(t_n)|,\ j = 1,2,...,K\,]; \qquad (12)$$

Minimum detector:

Output signal:

$$\phi_{\min}(t_n) = \text{minimum}[\,|r_j(t_n)|,\ j = 1,2,...,K\,]; \qquad (13)$$

Polarity thresholding:

$$\phi_{pt}(t_n) = \begin{cases} r(t_n), & \text{if } r_j(t_n) > 0 \text{ or } r_j(t_n) < 0 \\ & \text{for all } j = 1,2,...,K \\ 0, & \text{otherwise,} \end{cases} \qquad (14)$$

where t_n are discrete time instants with $n = 1,2,...,N$.

The computer simulations were performed using signals with different clutter patterns and flaw positions. Furthermore, to best evaluate the performance of flaw detection algorithms, it has been assumed that the flaw signal covers the same frequency band as background grain echoes, although in certain experimental situations, flaw echoes may show different frequency content. If a flaw signal is present in certain frequency bands, this implies that for those frequency bands, a high flaw-to-clutter ratio exists. Consequently, simple bandpass filtering will improve the flaw visibility.[14] In addition, the application of any flaw enhancement algorithms will result in a satisfactory performance.

An example of the output signal using the above processing techniques applied to a simulated microstructure signal is shown in Fig. 8. This result shows that the Bayes flaw detector out performs other detectors. Since the performance of flaw enhancement algorithms is generally dependent on the clutter pattern, we have simulated a number of clutter signals with the same statistical properties. Furthermore, to achieve a more realistic picture of the performance, we have compared the Bayes classifier with other flaw detection algorithms mentioned above. All flaw detection algorithms are examined using the same simulated signals, and the processed results are presented in Table I. This table

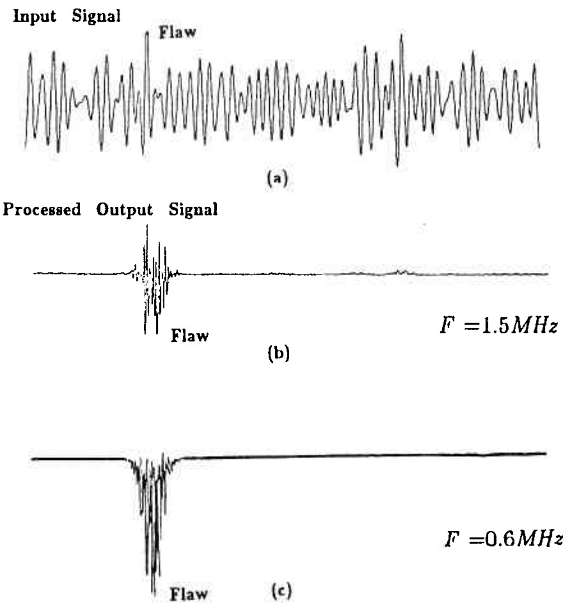

FIG. 7. A comparison of Bayes flaw detector performances via different starting frequency; (a) original signal, (b) $f = 1.5$ MHz, (c) $f = 0.6$ MHz.

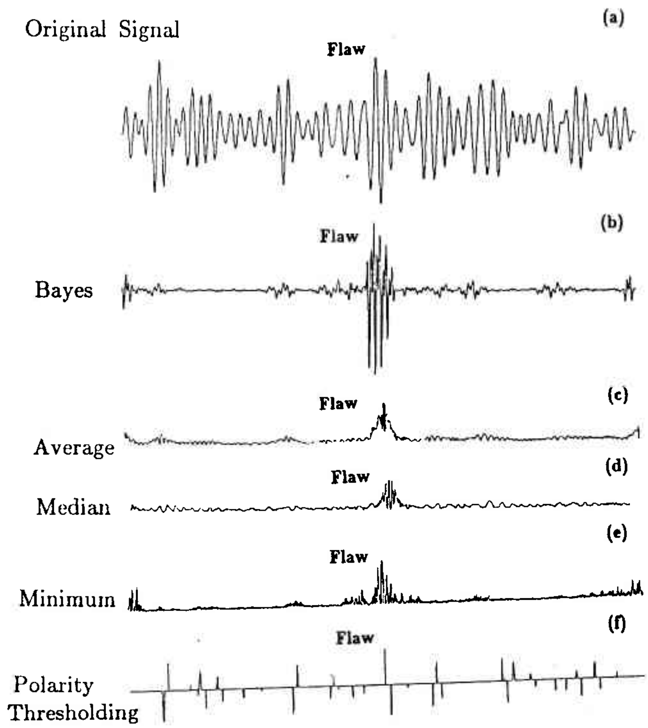

FIG. 8. A comparison of various signal processing outputs for a simulated microstructure signal; (a) original signal, (b) Bayes flaw detector, (c) average detector, (d) median detector, (e) minimum detector, and (f) polarity detector.

TABLE I. Flaw/clutter ratio enhancement of various processing techniques using simulated signals.

Trial No.	Bayes detector	Average detector	Median detector	Minimum detector	Polarity detector
1	2.62	1.07	1.09	2.39	0.42
2	4.40	1.37	1.34	1.80	0.98
3	6.17	1.21	1.82	2.20	0.93
4	10.83	1.03	1.22	3.24	0.56
5	1.79	1.08	1.56	1.13	0.39
6	3.82	1.57	1.46	1.60	0.82
7	3.19	1.06	0.97	2.63	0.20
8	6.09	1.02	1.05	2.36	0.29
9	7.49	1.06	1.08	2.27	0.13
10	7.88	0.97	1.01	1.99	0.29
11	3.34	1.59	1.51	1.45	0.68
12	4.86	1.07	1.28	1.81	0.28
13	7.54	1.04	1.24	1.83	0.22
14	8.04	1.02	1.08	2.08	0.19
15	7.99	1.00	1.19	1.55	0.24
16	7.98	1.05	1.17	1.75	0.31
17	8.41	1.06	1.24	1.96	0.19
18	3.28	1.08	1.20	2.20	0.11
19	2.68	1.04	1.08	2.20	0.11
20	3.25	1.02	1.07	2.19	0.03
Mean	5.57	1.12	1.23	2.03	0.36
SD	2.49	0.17	0.21	0.45	0.27

presents the flaw-to-clutter ratio enhancement results from 20 independent simulations where the flaw echoes were embedded in different locations. The means and standard deviation (SD) for each technique, i.e., Bayes detector, average detector, polarity thresholding, minimum and median detectors, is listed at the bottom of the table. A total of nine bandpass filters with a 2-MHz bandwidth, ranging from 0–14 MHz at frequency step 1.6 MHz, have been used for splitting the spectrum of returned echoes. For all simulated data, the input flaw-to-clutter ratio is kept close to unity where the flaw echo is not recognizable by a direct visual inspection of the backscattered signal. It is evident that the performance of the Bayes detector is superior when compared to other flaw detection algorithms. The averaging method, and minimum or median detectors provide only moderate improvement. The performance of polarity thresholding is unacceptable. This poor performance results from the fact that the method demands a change in the polarity of clutter echoes, but no change of polarity for flaw signals. In ultrasonic detection, polarity thresholding only works when the F/C is high[16,18] and, under this condition, many techniques such as a simple threshold detection will be equally as effective.

The performance of averaging techniques, and the minimum or median detector are highly dependent on the amplitude distribution function of returned echoes. Our previous analysis of these techniques indicates that these detectors are suboptimal and only perform well when certain conditions are satisfied.[19,20] As shown in Table I, F/C enhancement using the Bayes detector is 5.57, while other detectors can only improve detection up to twice. It should be noted that the variance of Bayes performance is also significantly larger than other detectors. This observation suggests that the performance of the Bayes detector is more sensitive to the parameters of the SSP or clutter patterns.

IV. EXPERIMENTAL RESULTS

This experiment was conducted using steel specimens with an average grain size about 50 μm and a Panametric transducer with a 6.22-MHz center frequency and a 3-dB bandwidth of 2.75 MHz. Flaws are formed by drilling two different flat-bottom holes, one with a 1.5-mm diameter and 2.5-cm depth and the other with a 2.0-mm diameter and 1-cm depth into the specimen. The complex flaw structure is formed by drilling two adjacent holes with 1.5-mm diameters, 2.5-cm depths, and a mutual distance about 3 mm into the specimen. The ultrasonic measurements were accomplished using the contact technique and data was acquired with a 100-MHz sampling frequency. Nine bandpass filters were used, with 3-dB bandwidth of 1 MHz and ranging from 0.6–9 MHz at frequency steps of 0.8 MHz. This range of frequency provides both flaw and microstructure noise in all nine bandpass channels. The test signals with different F/C ratios were obtained by slightly shifting the transducer beam path away from the flaw position at different directions in order to obtain a flaw-to-clutter ratio about 0 dB (i.e., signal has poor F/C ratio). The training process of the optimal flaw detector was accomplished by measuring the clutter signal with and without flaw under the same equipment setting.

An example of the output signal, from various signal processing techniques for an experimental measurement is displayed in Fig. 9. We repeated these observations for many microstructure signals and results are shown in Table II. Note that trials 1–8 represent the signal from simple flaw and trials 9 and 10 represent the signal from a complex target (i.e., two adjacent flaws). Because of the poor performance of polarity thresholding, the processed results for experiments are not included in Table II [see Fig. 9(f)]. Inspection of Table II and Fig. 9 suggest that the Bayes detector can improve the F/C ratio significantly when compared to other flaw enhancement algorithms. Furthermore, trials 9 and 10 given in Table II reveal that the presence of two flaw echoes falling within the resolution of the detection does not deteriorate the performance of the Bayes detector. Similar improvements in detection were observed using computer simulation of multiple flaw echoes, masked by different clutter patterns.

Overall, the performance of Bayes classifier flaw detection is dependent on statistical differences between microstructure scattering and defects. If defects are complex and, more importantly, if their sizes are comparable to the wavelength and grains (e.g., microcracks surrounded by large grains), the backscattered signal associated with cracks may not exhibit any statistical differences when compared with grain scattering, and, consequently, would go undetected. On the other hand, defects are often larger than ultrasonic wavelength and grain size, and will show statistical differences in their scattering echoes as a function of frequency which will be detected by the Baysian classifier. The Baysian method is particularly effective when the superposition of