AFRL-SR-BL-TR-00-

0117

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	January 2000	Final Technical Report 07/15/98-12/31/99

4. TITLE AND SUBTITLE
Fractographic Analysis of High-Cycle Fatigue in Aircraft Engines

5. FUNDING NUMBERS

6. AUTHOR(S)
Donald A. Shockey, Takao Kobayashi, Naoki Saito, Jean-Marie Aubry, Alberto Grünbaum

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025-3493

8. PERFORMING ORGANIZATION REPORT NUMBER
2478

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
AFOSR/NM
801 North Randolph Street
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
F49620-98-C-0041

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

12b. DISTRIBUTION CODE

13 ABSTRACT (*Maximum 200 words*)

A seedling program was conducted to determine whether load information is stored in the fracture surface features of failed aircraft components and, if so, to explore methods for extracting this information.

Fracture surfaces produced under systematically varied cyclic load conditions in laboratory specimens of titanium turbine blade alloy were provided to the program by an aircraft engine manufacturer. Three-dimensional topographs of the surfaces were obtained via confocal optics and miscroscopy and analyzed with Fourier and wavelet techniques. Both analysis techniques confirmed the existence of load information in the fracture surface topography. Moreover, correlation was found between Fourier results and fatigue load parameters.

The topographic characterization and analysis method was then applied to an F-118 engine component that had cracked in service to determine whether the load history that caused the cracking could be extracted. Using the results from the laboratory specimens as a reference database, the method provided a quantitative estimate of an influential fatigue load parameter and indicated its variation with crack growth distance.

14. SUBJECT TERMS			15. NUMBER OF PAGES
Fractography Failure Analysis	Fourier Fracture Surface	High-Cycle Fatigue Wavelet	16. PRICE CODE 123

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	Unlimited

SUMMARY

Cracks that form in aircraft engine turbine blades, disks, vanes, and rotors under high frequency vibrational loads during flights have caused many Class A incidents (loss of aircraft) in the USAF fighter fleet. These high-cycle fatigue failures are responsible for unacceptable loss of life, costly redesigns and retrofits, and huge maintenance costs. More important, high-cycle fatigue failures result in grounding of the fleet for days to weeks. The USAF has identified high-cycle fatigue as the primary air readiness problem.

The high-cycle fatigue problem exists because the load spectrum experienced by engine components is not known. In the damage tolerance philosophy used to manage aircraft engine integrity, this means crack growth rates cannot be adequately predicted, which means components cannot be reliably designed and inspection intervals cannot be set with confidence.

Thus, a way is needed to determine the load history of in-service components. The load spectra deduced by finite element analysis of engine vibration characteristics are not pertinent because of aberrations from mission uniqueness, unforeseen pilot maneuvers, wind gusts, and the like. Pilot reports are subjective and on-board sensors are not usually fielded. A possible solution is to extract load spectra information from the markings on the crack surfaces. This seedling program investigated whether parameters of the fatigue load could be obtained by analyzing the crack face topography.

We obtained laboratory specimens of Ti-6Al-4V that had been tested to failure under prescribed fatigue loading conditions and selected areas of the fracture surfaces that had been produced at various values of K_{max}, stress range, and stress ratio—the three important load parameters for fatigue. Using a confocal optics microscope, we made topographic maps of these areas and analyzed the topographs with Fourier and wavelet techniques.

The energy power spectral density (EPSD) curves from the Fourier analysis showed clear systematic variations with ΔK, stress ratio, and hence K_{max}. The curves varied in height with ΔK; the shape of the curves varied with stress ratio. For any given wavelength between 0.5 and 10 μm, a linear relationship existed between the square root of EPSD and ΔK. Moreover, a stress ratio of 0.8 produced a linear EPSD curve in this wavelength range, whereas a stress ratio of 0.1 produced a concave shape.

These results show that load parameter information is encrypted in the roughness details of the fracture surfaces. Wavelet results confirmed these findings, but were less definitive because the fracture areas examined were small relative to the grain size of the alloy and because of noise in the elevation data. Nevertheless, the Fourier results reveal the opportunity, when fully explored and properly analyzed, to determine quantitative load spectra parameters from the fracture surface topography.

We then examined an aircraft component that had acquired a crack during service to determine if the load history that caused the crack could be identified. The topographies of areas along the crack growth direction were characterized, then analyzed with the Fourier method. The EPSD curves from the components were interpreted, using the results from the laboratory specimens to estimate quantitatively the load conditions responsible for the failure and their variation with crack depth. Strictly from the fracture surface elevation data, we determined that the crack grew to a length of 1.5 mm under rather constant ΔK level of about 5 $MPa\sqrt{m}$ after initiating at a higher value. At this crack depth, the ΔK increased strongly and linearly to 10 $MPa\sqrt{m}$ in the next millimeter of crack advance. We further estimated that the stress ratio was about 0.8.

Thus, a route has been demonstrated for extracting load information from fracture surfaces. The next steps are to develop topography characterization techniques capable of resolving finer fracture surface features, develop size scaling procedures for roughness features ranging from nanometers to millimeters, develop wavelet analysis techniques to obtain fracture-site-specific signals, and develop a computational model of microstructural deformation and microfailure. The resulting technology will enable load conditions and crack evolution kinetics to be extracted from the fracture surfaces of a failed part and will constitute a leapfrog advance in the state of practice of failure analysis.

CONTENTS

FIGURES

ACKNOWLEDGMENTS

The authors thank Dr. Dennis Healy of the Defense Advanced Research Projects Agency and Dr. Anna Tsau of the Center for Computing Sciences at the Institute of Defense Analysis for their interest and encouragement. Drs. Michael Gehron, Joe Bateman, and Jay Littles and their colleagues at Pratt & Whitney consulted frequently, provided crack growth specimens for analysis, and hosted a two-day meeting of the project team at their West Palm Beach facility in which they described how they used fractography and discussed their needs. Dr. Theodore Nicholas of the Wright-Patterson Air Force Laboratory provided the user's perspective of failure analysis, the engine components that failed in service, and very valuable technical advice throughout this program.

INTRODUCTION

Fractography, the art of interpreting markings on fracture surfaces to deduce how and why a failure occurred, is an important aspect of many failure investigations.[1] Fractography is often relied on to determine whether the material was defective, the component design was inadequate, or the allowable operating conditions were exceeded. Hence, fractography implicates what was responsible for the failure and provides the basis for modifications to reduce the likelihood of future failures.

Generally, fractographic results are qualitative and consist of the failure mode, the load conditions, the fracture initiation site, and the propagation direction. Fractures are classified as ductile, brittle, or intergranular; the loads that caused the failure are identified as cyclic, overload, or environment-assisted; the crack is determined to have initiated at a material defect, a machining or damage mark, or at a stress concentrating feature of the component geometry. When the origin of the crack can be pinpointed, the initiation site can be examined with microscopes and chemical analysis instruments to determine the nature of the defect.

Although all this information is very useful, the failure analyst would often like to know additional and, if possible, quantitative information. For example, what were the loads (their magnitude, duration, frequency) that operated and led to the fracture? How long did it take for the crack to form? And once formed, what was its growth history? Did it propagate at a uniform speed, or did it experience periods of acceleration, deceleration, or perhaps hesitation? A way to quantitatively extract the load parameters and the crack kinetics from the fracture surfaces would be of great utility.

Several years ago, we began thinking whether such information might be encrypted in the fracture surfaces and, if so, how it might be extracted. We participated in a Navy workshop[2] where the topic of interest was fatigue failures of helicopters and fixed-wing aircraft.

An unacceptable frequency of accidents was causing safety concerns, incurring huge maintenance costs, and compromising the air readiness of the fleets.* Component designers at the workshop told us that fatigue problems could be reduced if the fatigue load spectra were known. Components were designed on load spectra deduced by finite element analysis of engine vibration characteristics. In reality, however, the load spectra experienced by a component during service can be quite different because of mission uniqueness, unforeseen pilot maneuvers, wind gusts, and the like. Overloads are not reliably obtained from pilot reports, and on-board sensors are not usually fielded.

It occurred to us at this workshop that load spectra might be obtainable from the fracture surfaces. We reasoned that the important loads in the spectrum were those that affected the fatigue process, and since these loads affected crack growth, they might manifest themselves on the fracture surfaces. We considered that the extent and nature of out-of-plane deformation should be a function of the loads and hence should be accessible from the topography. Thus the fracture surface may act as a load spectrum gauge!

---

* It is difficult to estimate the total annual cost of aircraft fatigue failure to the military. When we asked the question to representatives of the military and industry, we learned that the costs fall into several categories, failure investigations, redesign and retrofitting, and maintenance, but the greatest concern of aircraft fatigue is air readiness. A dollar value cannot be placed on having the fleet grounded.

Failure investigations. "The costs of engineering root cause investigations frequently run into the multiple millions of dollars. These investigations require large cross-disciplinary groups of investigators (from 20 to 100 people), laboratory verification of failure scenarios, and costly engine tests (about $40,000 per day). (One specific accident investigation cost $23M. Otha Davenport, SPO USAF.) These investigations also can take many months to complete, during which time the unresolved problem can lead to additional field failures, possibly resulting in loss of aircraft and life." Dr. Jerrol W. Littles, Jr., Pratt & Whitney

Redesign and retrofitting. "The potential for saving a costly redesign and retrofitting of the fleet if the cause were known and another solution such as damping could be used would be enormous. For example, the fourth stage blade failure in a single crystal material, which was never fully understood, ended up costing the government probably over 20 million dollars. If we knew the failure mode, we might have prevented failure by adding damping or redirecting air flow." Dr. Theodore Nicholas, AFRL, Program Manager, Materials Characterization, HCF Science and Technology Program. The Navy experienced "extreme trouble with cracking in helicopter rotor heads caused by uncertainty in fatigue crack growth rates, because the load spectrum was not known." The problem was addressed at great expense by redesigning the component and changing the material. John T. Cammett, Naval Aviation Depot, Cherry Point, NC.

Maintenance costs. Annual costs for just one specific failure mode and component—high-cycle fatigue of aircraft engines—were $200M in FY94, $150M in FY92-3, and $220M in FY88-91. Robert J. May, Jr., USAF Propulsion Product Group Manager. The Navy, with a fleet of 100's of helicopters, requires each to be inspected at intervals as short as every 10 hours of flight. John T. Cammett, Naval Aviation Depot, Cherry Point, NC.

Air readiness. High-cycle fatigue is the No. 1 air readiness concern of the USAF. Dr. Deutch USAF. If just one recurring failure mode could be eliminated, countless days or weeks of not being in a readiness posture in the field could be avoided.

This idea may not be so far-fetched. For example, in certain situations a procedure for deducing ΔK from fracture surfaces already exists.[3,4] When striations are visible and a correlation with load cycle and striation can be established, the rate of change of crack length with load application can be obtained and correlated with data from controlled fatigue crack growth tests in the laboratory to deduce ΔK. Even in the absence of striations, skilled fractographers can often estimate the crack growth rate to an accuracy of half an order of magnitude[5] and, using the Paris law, obtain an estimate of ΔK. Furthermore, recent work by fracture scientists in Japan[6,7] suggests a relationship between striation height-to-width ratio and the stress ratio, R.

To be more useful and generally applicable, however, the load spectra extraction procedure needs to be more systematic, more quantitative, and more precise. The fact that engineers experienced at assessing fracture surfaces can reach consensus on approximate load conditions suggests that an objective method is possible. If the eye and brain can perform this function, it must be possible to develop an instrument and software to do the same. A preliminary effort was conducted by SRI to test whether fracture surface topography was sensitive to load conditions,[8,9] and a seedling program was initiated by DARPA to conceive an approach and investigate this possibility. This report describes the results of the seedling program and presents a plan for elevating fractography from a qualitative, subjective, two-dimensional art to a quantitative, objective, three-dimensional science.

BACKGROUND

We first consider how a fracture surface forms. Figure 1(a) shows the slip bands that formed in the grains of a nickel alloy after 1200 stress cycles.[10] These bands of localized dislocations lie on crystallographic planes of low shear resistance and hence are oriented differently in adjacent grains. Where the bands encounter a grain boundary, dislocation motion stops, because of the misalignment of the slip planes in the adjacent grain. Stresses at such intersections are created because the displacement produced by the band is not transferred to the neighboring grain and the strain discontinuity is accommodated in the grain boundary.

Figure 1(b) shows the same area after 4000 cycles. The density of the slip bands has increased, and the resulting increase in stress at the grain boundaries has caused the boundaries to separate. The failed grain boundaries become islands on the fracture surface. Moreover, the grain boundary facets may contain steps where the slip bands intersected.

As cycling continues, slip band activity continues, additional grain boundaries fail, and ligaments between grains stretch, neck, and rupture, resulting in a macrocrack and two opposing fracture surfaces containing roughness features on a range of size scales. The tortuous topography belies the deformation and failure processes that took place in the microstructure. Second phase particles, inclusions, pores, and inherent flaws (when they are present) also influence the failure process and hence the fracture surface.

In the case of fatigue failure, an additional phenomenon can affect the fracture surface. In the unloading phase of the load cycle, the crack faces relax toward each other. If they make contact, they can modify the roughness features. Especially for conditions where the lower loads are a small fraction of the maximum loads or even compressive, the features standing up in relief can be deformed to different shapes. This phenomenon is called closure and has an important effect on fatigue crack growth rate.

Thus, the topography of a fracture surface is dictated by the microstructure and the deformation and microfailure processes exacted on the microstructure by the applied loads. This suggests that the load history experienced by a failed component is recorded in the fracture surface and encrypted in the roughness details. The challenge is to develop a procedure for extracting this information.

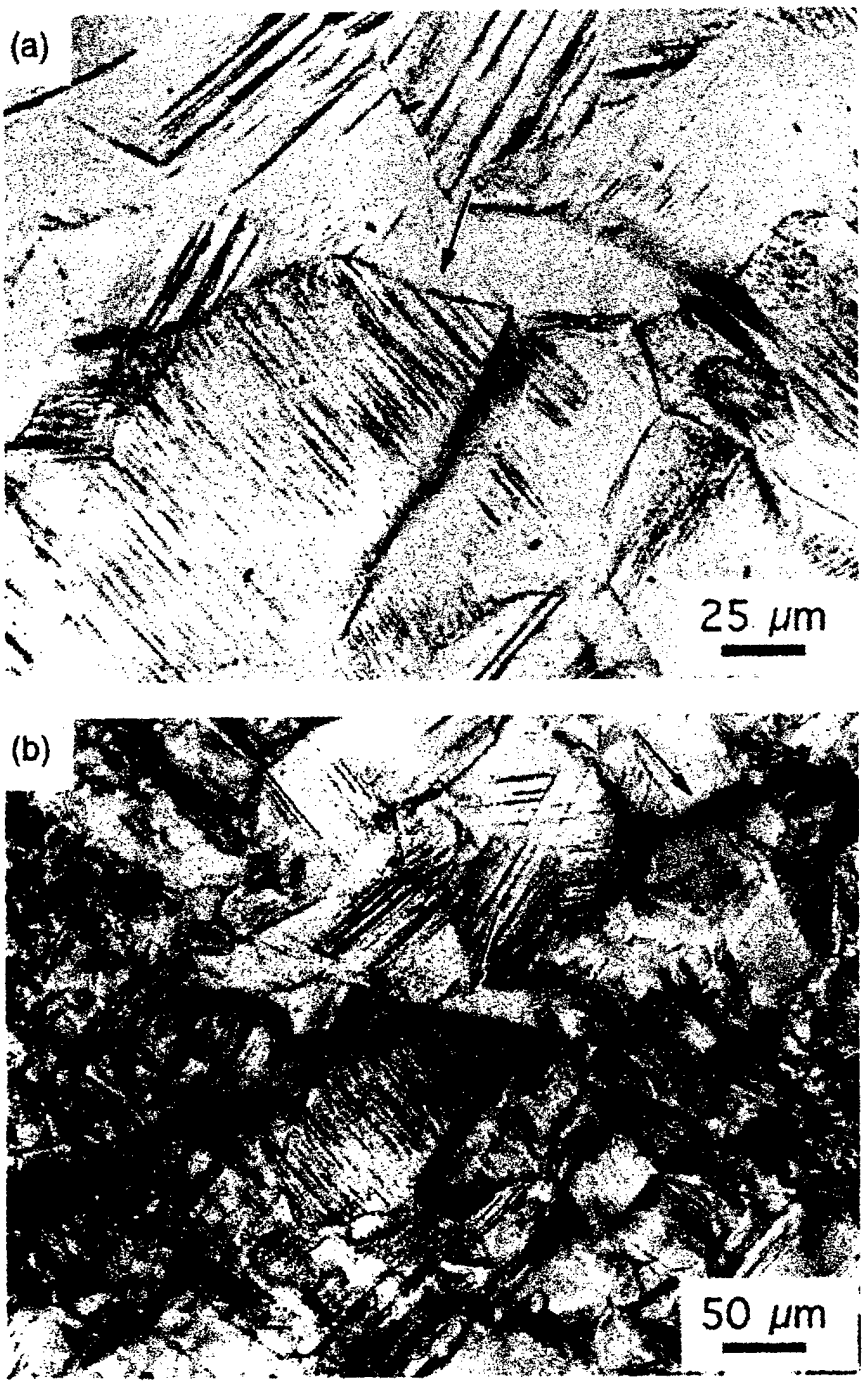


Figure 1. Fatigue cracks in nickel after 1200 (a) and 4000 (b) tension-compression load cycles. (Ref. 1)

The first steps were taken in this seedling program. To determine whether fracture surfaces were sensitive to loading conditions, we obtained fracture specimens tested under cyclic loads under controlled conditions of different ΔK, quantified the topographies of their fracture surfaces, and applied Fourier and wavelet analysis techniques to the elevation data. We looked for correlations in roughness signatures and features of the analyses with load parameters. To demonstrate a route for extracting load parameters from fracture surfaces, we then performed a similar analysis of a field failure and applied the modest database from the laboratory specimens to estimate ΔK as a function of crack length.

ANALYSIS OF LABORATORY FRACTURE SPECIMENS

Pratt & Whitney, which has conducted many fracture tests over the years to determine the fatigue properties of turbine engine material, made its large inventory of tested specimens available to this program. We selected for analysis two compact-tension specimens of Ti-6Al-4V that had been tested to examine fatigue crack growth behavior in the near-threshold regime. A crack had been driven through the specimens under a range of well-controlled and documented loading conditions according to the ASTM E647 requirements for load shedding.[11] The fracture surfaces were produced at constant stress ratio, $R = K_{min}/K_{max}$, and varying stress range, $\Delta K = K_{max} - K_{min}$, with successive areas produced under a slightly lower or slightly higher ΔK. Two values of R were examined: 0.1 and 0.8. The variation of load conditions for each specimen is shown in Figure 2. In total, 13 load conditions and 119 fracture surface areas were examined.

SURFACE TOPOGRAPHY CHARACTERIZATION

Small areas on the fracture surface in each fatigue test section were selected at a certain spacing along the crack growth direction, and their topography was characterized by a system known as a FRASTAscope.[12,13] This system was developed at SRI and consists of a confocal-optics-based scanning laser microscope; a precision, computer controlled x-y-θ stage; and computer programs that control data acquisition, manipulation, and display. Tables 1 and 2 summarize the locations of the areas we chose for examination and the corresponding ΔK values. The size of each area was 140.3 μm in the crack growth direction and 466.5 μm in the orthogonal direction. The number of elevation data points in the crack propagation direction was 600 and in the orthogonal direction 2000. Data spacing, thus, was 0.233 μm in both directions. Figure 3 shows the contrast image and gray-scale elevation image of a typical area.

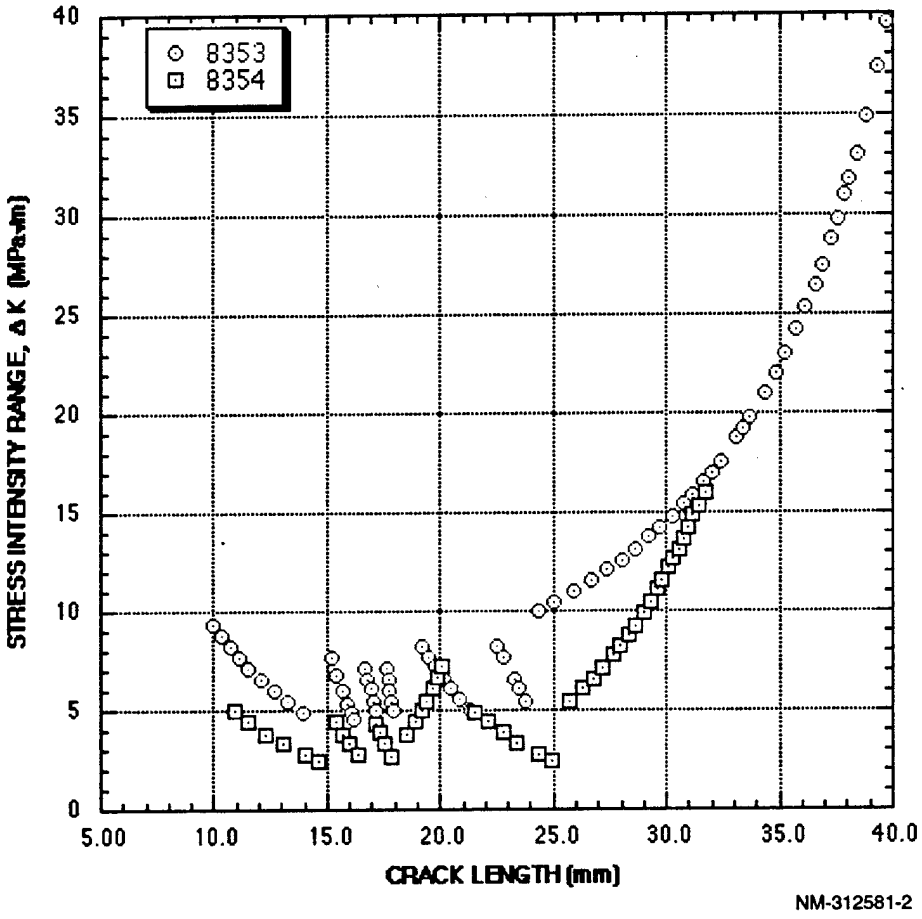


Figure 2. Stress intensity range, ΔK, as a function of crack length during multiple fatigue tests in Specimens 8353 and 8354.

Table 1. Crack Length and Stress Intensity Range for Several Fatigue Tests in Specimen No. 8353 (R = 0.1)

File No.	Test Run Number	Crack Length (mm)	Stress Intensity Range, ΔK, (MPa$\sqrt{m}$)
1	Run No. 1	9.9238	8.52
2		10.320	7.99
3		10.739	7.45
4		11.148	6.96
5		11.557	6.53
6		12.128	5.96
7		12.670	5.46
8		13.236	5.00
9		13.927	4.48
10	Run No. 2	15.222	6.94
11		15.448	6.14
12		15.659	5.46
13		15.878	4.86
14		16.050	4.47
15		16.165	4.15
16	Run No. 3	16.685	6.50
17		16.800	5.98
18		16.919	5.53
19		17.064	5.01
20		17.203	4.54
21	Run No. 4	17.633	6.53
22		17.696	6.01
23		17.770	5.50
24		17.849	5.01
25		17.923	4.51
26	Run No. 5	19.228	7.46
27		19.489	7.03
28		19.789	6.54
29		20.147	6.01
30		20.508	5.53
31		20.858	5.08
32		21.364	4.51
33	Run No. 6	22.545	7.50
34		22.784	6.96
35		23.256	5.97
36		23.487	5.53
37		23.815	4.97

Table 1. Crack Length and Stress Intensity Range for Several Fatigue Tests in Specimen No. 8353 (R = 0.1) (Concluded)

File No.	Test Run Number	Crack Length (mm)	Stress Intensity Range, ΔK, (MPa$\sqrt{m}$)
38	Run No. 7	24.326	9.11
39		25.065	9.52
40		25.880	10.00
41		26.695	10.53
42		27.404	11.02
43		28.026	11.49
44		28.628	11.97
45		29.246	12.51
46		29.746	12.97
47		30.259	13.49
48		30.770	14.03
49		31.173	14.48
50		31.684	15.10
51		31.989	15.49
52		32.390	16.03
53		33.109	17.06
54		33.409	17.53
55		33.708	18.03
56		34.326	19.10
57		34.834	20.08
58		35.243	20.92
59		35.756	22.06
60		36.159	23.05
61		36.568	24.12
62		36.871	24.96
63		37.274	26.18
64		37.584	27.15
65		37.899	28.22
66		38.105	28.94
67		38.407	30.07
68		38.824	31.77
69		39.327	34.02
70		39.736	36.04
71		40.472	40.21
72		41.118	44.68

Table 2. Crack Length and Stress Intensity Range for Several Fatigue Tests in Specimen No. 8354 (R = 0.8)

File No.	Test Run Number	Crack Length (mm)	Stress Intensity Range, ΔK, (MPa$\sqrt{m}$)
1	Run No. 1	10.927	4.52
2		11.554	4.01
3		12.266	3.49
4		13.063	2.99
5		14.008	2.50
6		14.661	2.20
7	Run No. 2	15.400	4.05
8		15.705	3.49
9		16.017	3.01
10		16.391	2.53
11	Run No. 3	17.140	3.97
12		17.315	3.53
13		17.562	3.01
14		17.866	2.47
15	Run No. 4	18.519	3.45
16		18.908	4.05
17		19.195	4.57
18		19.393	4.95
19		19.682	5.58
20		19.870	6.07
21		20.076	6.58
22	Run No. 5	21.580	4.47
23		22.075	4.03
24		22.748	3.51
25		23.406	3.05
26		24.323	2.52
27		24.930	2.22
28	Run No. 6	25.756	4.98
29		26.299	5.52
30		26.794	6.02
31		27.181	6.47
32		27.645	7.05
33		27.950	7.46
34		28.329	8.00
35		28.628	8.45
36		29.002	9.06
37		29.294	9.56
38		29.583	10.10
39		29.804	10.50
40		30.099	11.09
41		30.307	11.54
42		30.526	11.98
43		30.736	12.45
44		30.963	12.93
45		31.166	13.58
46		31.440	13.99
47		31.722	14.60

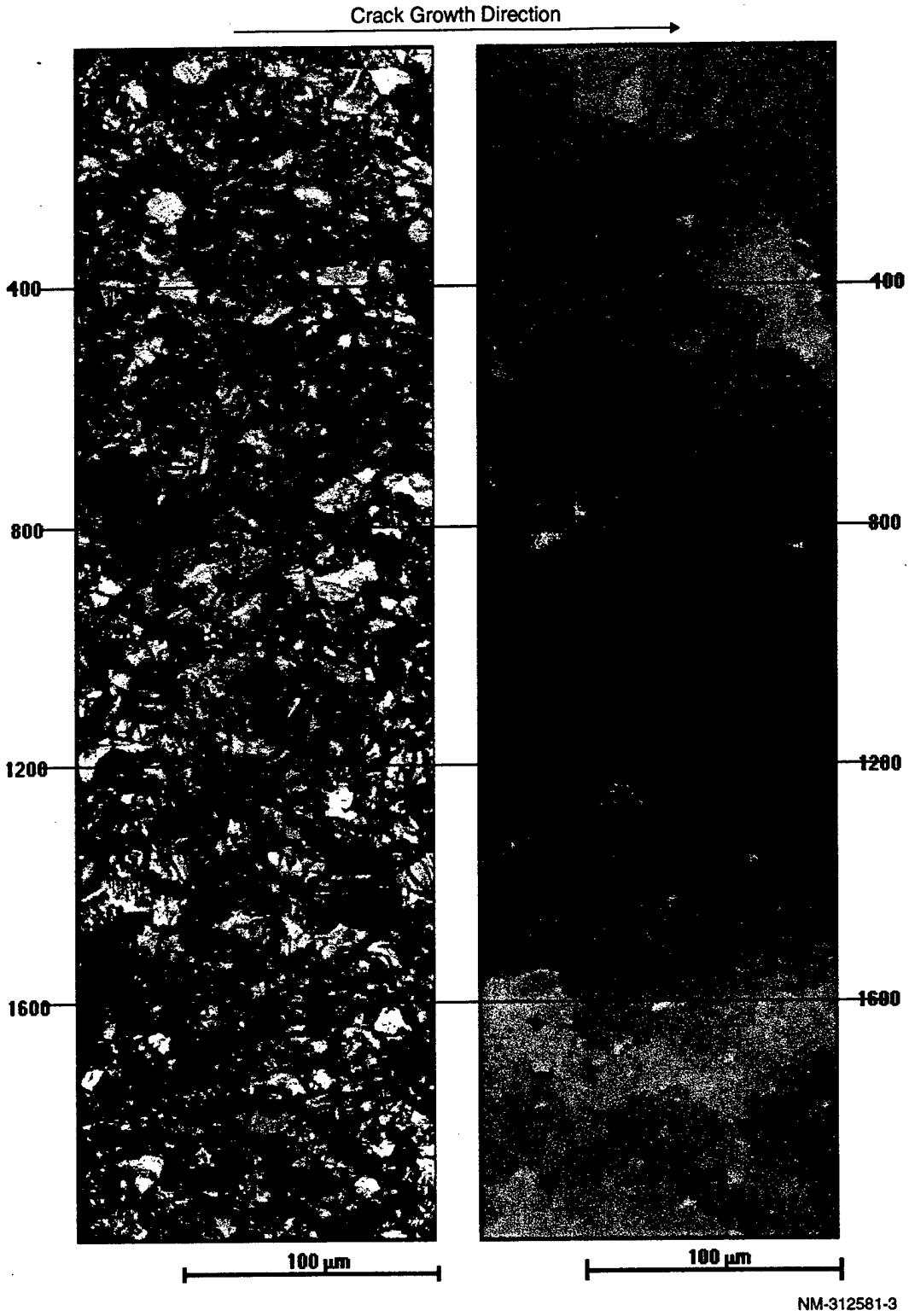


Figure 3. Contrast and gray-scale elevation images of an area on the fracture surface of Specimen 8353.

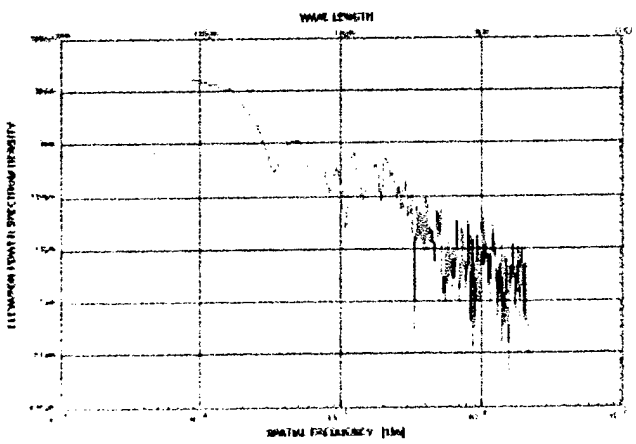
FAST FOURIER TRANSFORM (FFT) ANALYSIS

A fast Fourier transform (FFT) analysis was applied to a row of data points parallel to the crack growth direction, and the elevation power spectrum density distribution was calculated. Each row contained 600 data points; however, only the central 512 ($=2^9$) points were analyzed. The EPSD curves for individual rows exhibit significant differences, as shown in Figure 4, due to local differences in microstructure. Such microstructural effects, although important from a materials science viewpoint, obscure the effect of loading. Noting that the stress intensity range, ΔK, used to represent the loading condition is an average value over the length of the crack front and not a local ΔK acting on a single grain or a few grains, we depressed local microstructural effects and emphasized the effects of loading condition by averaging the EPSD results over 2000 lines wavelength-by-wavelength.

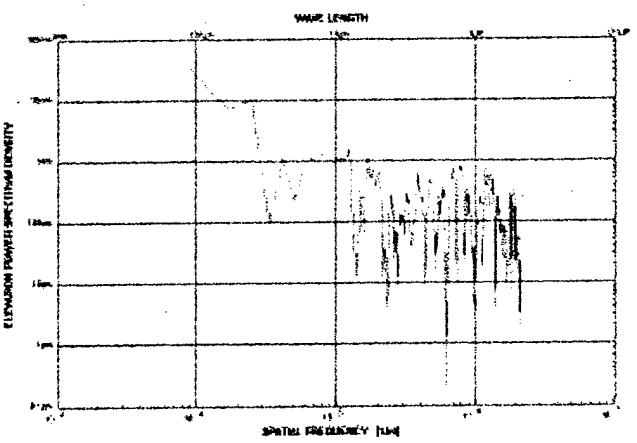
The averaged EPSD curves from the different locations on the fracture surface are closely parallel between wavelengths of 0.5 μm and 20 μm (see Figure 5, for example). Choosing the square root of the EPSD [SQRT(EPSD)] value at a wavelength of 5 μm as a representative index of the roughness of each area and plotting these values as a function of position on the fracture surface, we obtain the results shown in Figures 6(a) and 6(b) for Specimens 8353 (R = 0.1) and 8354 (R = 0.8), respectively. Also plotted are the ΔK values corresponding to the fracture surface locations or crack lengths.

At first inspection, the behaviors of SQRT(EPSD) and ΔK as a function of crack length seem quite different, and a correlation between load and roughness is not clear. For example, when the load shedding tests were initiated at a higher ΔK, the SQRT(EPSD) values did not start at a higher value, but instead gradually increased to a higher value before decreasing with decreasing ΔK. In the monotonically increasing ΔK fatigue test, the SQRT(EPSD) values initially decreased before increasing.

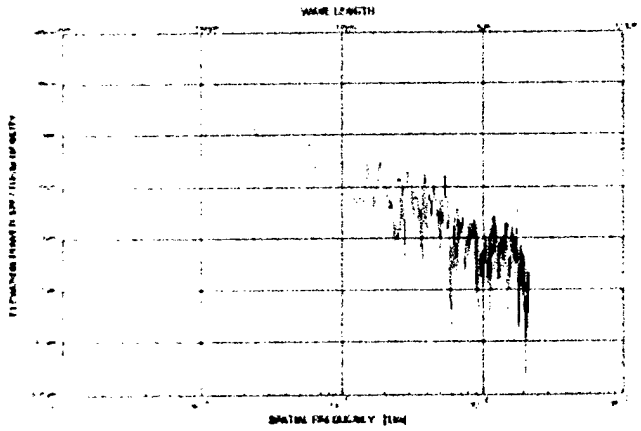
In seeking to unify these results, we plotted all of the SQRT(EPSD) values in Figure 6 as a function of ΔK. The results in Figure 7 show three regions for each specimen: (1) a linear relationship in the low ΔK region (less than about 8 $MPa\sqrt{m}$ for 8353 and about 6 $MPa\sqrt{m}$ for 8354), although a significant scatter exists; (2) a transition region (between 8 and 16 $MPa\sqrt{m}$ for 8353 and between 6 and 9 $MPa\sqrt{m}$ for 8354); and (3) another linear relationship region at higher ΔK (above 16 $MPa\sqrt{m}$ for 8353 and 9 $MPa\sqrt{m}$ for 8354).



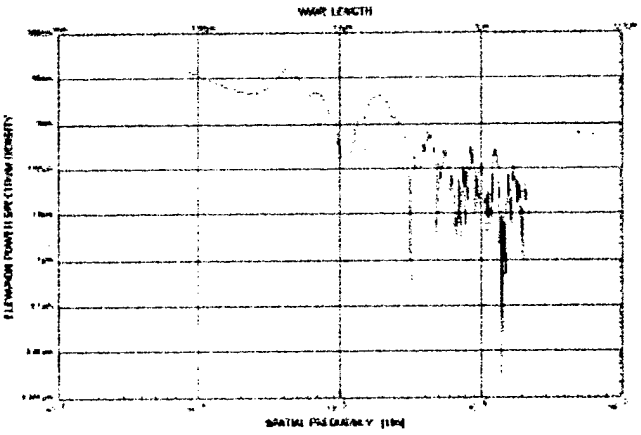
(a) EPSD Curve of Data along Line 400



(b) EPSD Curve of Data along Line 800



(c) EPSD Curve of Data along Line 1200



(d) EPSD Curve of Data along Line 1600

NM-312581-4

Figure 4. EPSD Curves for the roughness profiles along several lines shown in Figure 3. (Specimen 8353)

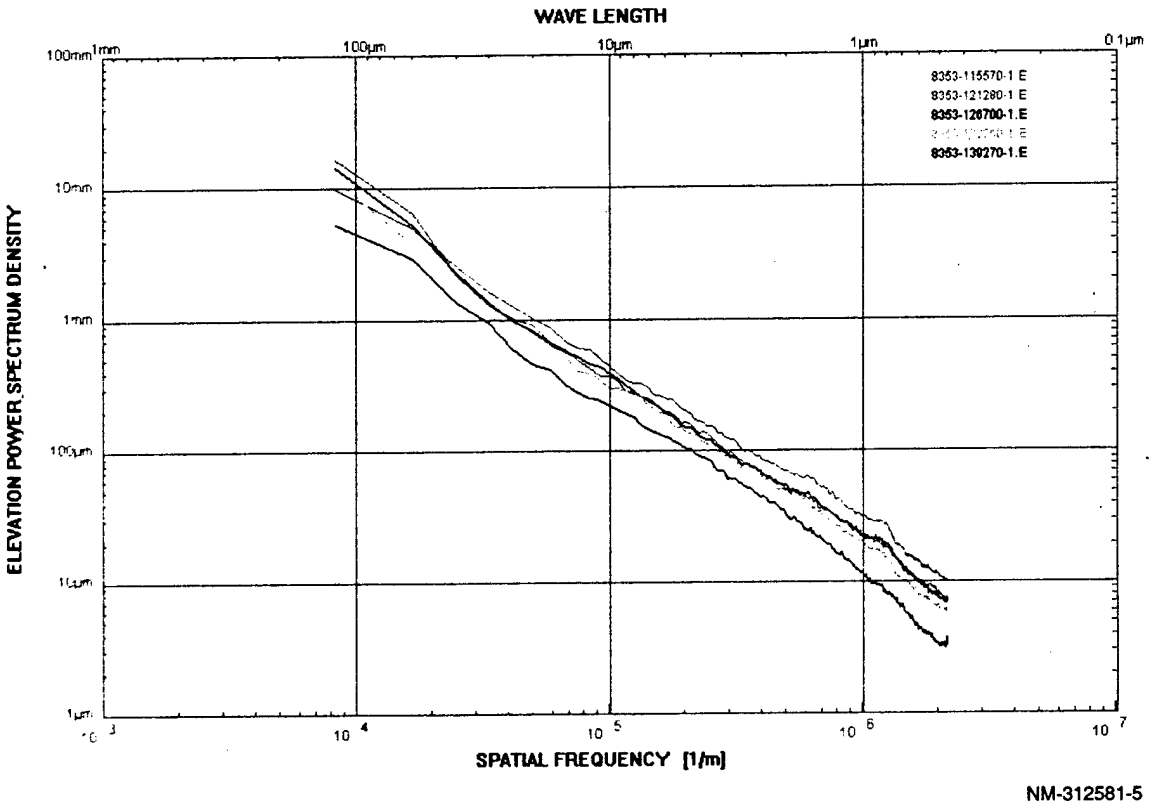
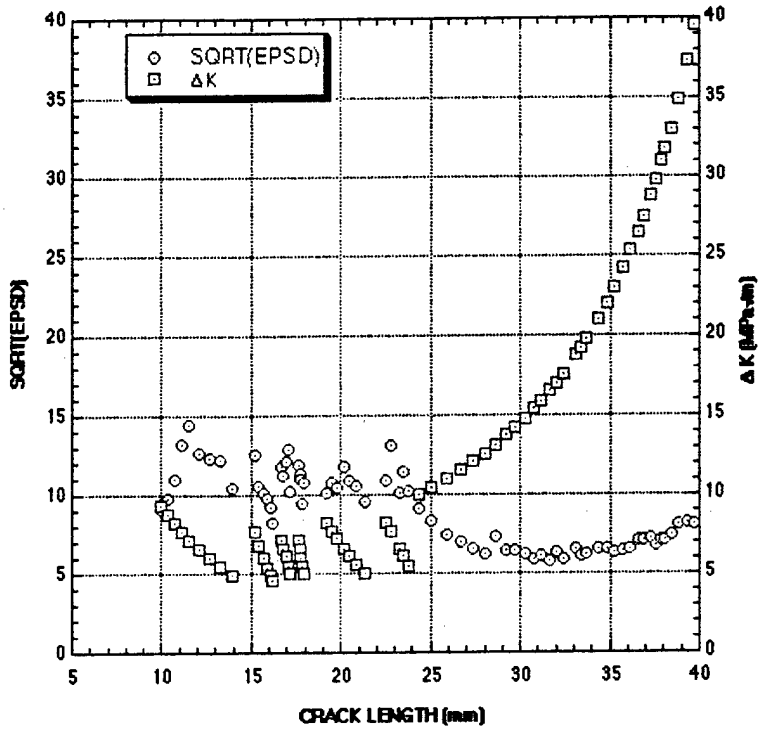
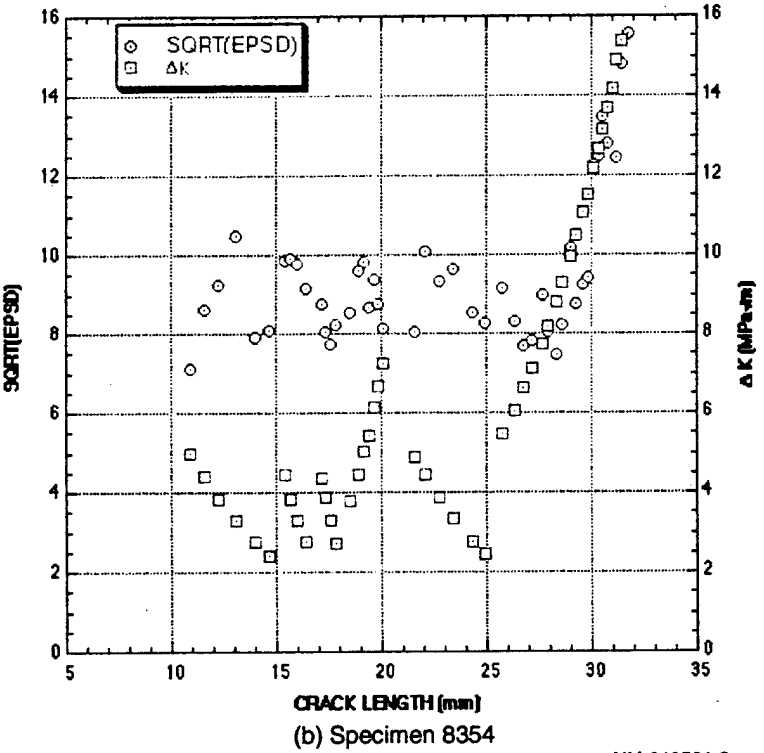


Figure 5. Averaged EPSD curves from five locations on the surface in Test Run 1 for Specimen 8353.



(a) Specimen 8353



(b) Specimen 8354

NM-312581-6

Figure 6. Square root of EPSD values and stress intensity range, ΔK, as a function of crack length for Specimens 8353 and 8354.

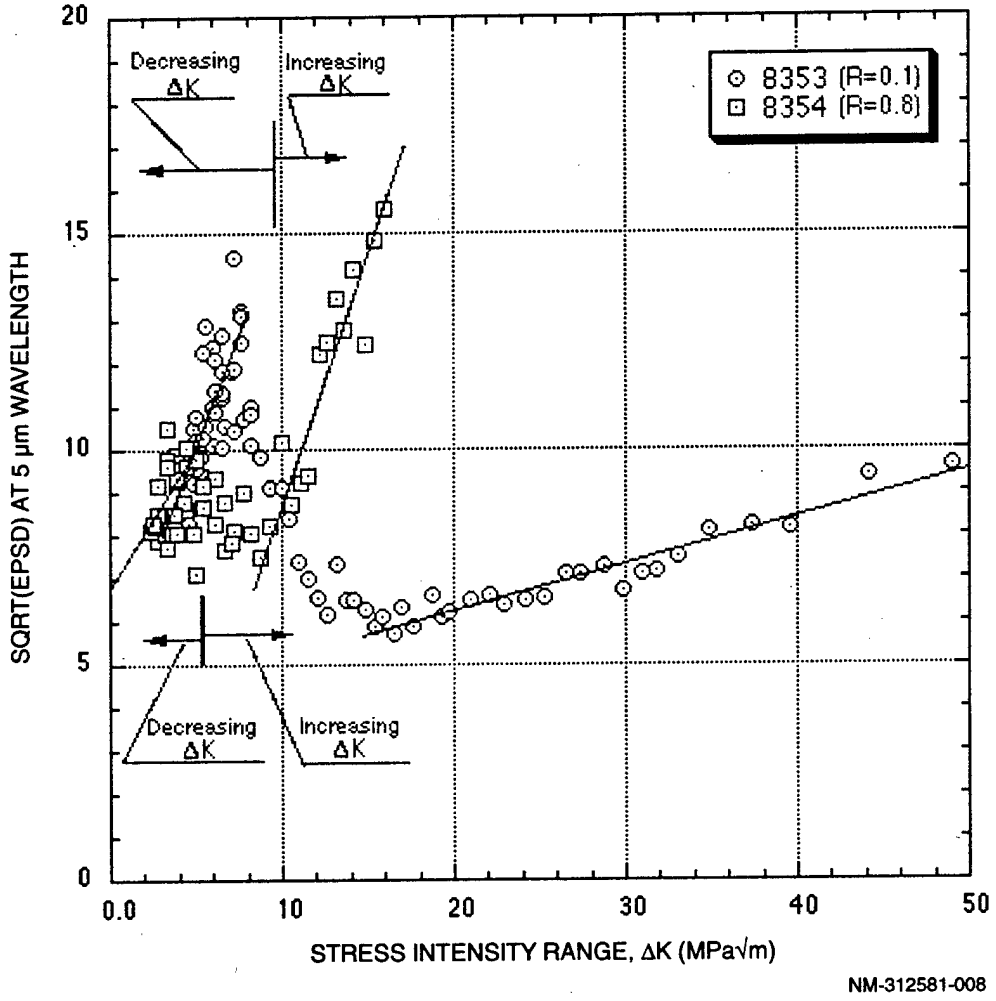


Figure 7. Summary of the relationship between SQRT(EPSD) and ΔK for Specimens 8353 and 8354.

The equations for these linear relationships are given below:

8353 low ΔK region: SQRT(EPSD) = 4.710 + 1.067* ΔK
8353 higher ΔK region: SQRT(EPSD) = 4.116 + 0.1077* ΔK

8354 low ΔK region: SQRT(EPSD) = 6.776 + 0.6553* ΔK
8354 higher ΔK region: SQRT(EPSD) = -2.957 + 1.156* ΔK

The relationship between SQRT(EPSD) and ΔK is not simple, but can be explained in terms of crack tip plastic deformation and crack closure, as discussed below.

CRACK CLOSURE EFFECT ON THE FRACTURE SURFACE TOPOGRAPHY

Two competing processes are responsible for the deformation topography of the fracture surfaces: (1) the roughening of the surface due to increasing ΔK or K_{max} and (2) the flattening of the surface due to crack closure.* When the ΔK values are low, Figure 7 indicates that the SQRT(EPSD) value increases with increasing ΔK. In this region, the effect of roughening of the surface with increasing ΔK is dominant.

However, when ΔK reaches a certain value, the SQRT(EPSD) starts to decrease as ΔK continues to increase (transition region in Figure 7). In this region, crack closure begins to have an effect, flattening asperities on the fracture surfaces and gradually overcoming the effect of surface roughening with increasing ΔK. The surface flattening effect outpaces surface roughening as ΔK increases. Thus, the SQRT(EPSD) values decrease despite the increase in ΔK.

As ΔK continues to increase, the effect of surface roughening starts to overcome the effect of surface flattening, and the SQRT(EPSD) values start to increase again as ΔK increases. This is the third region in Figure 7.

With the above explanation in mind, we examined and compared the results in Figure 6. Recall that Specimen 8353 was tested at a stress ratio R = 0.1 and 8354 at R = 0.8, and that ΔK and K_{max} were higher for Specimen 8353 than for Specimen 8354. Since the ΔK in the first region is low and crack closure is unlikely for either specimen, the SQRT(EPSD) in the first region of the specimen tested under higher ΔK (Specimen 8353) should show higher values. Figure 7 confirms this. Conversely, the test condition with lower stress ratio should result in a higher degree of crack

* Crack closure is the phenomenon of the faces of a crack contacting and perhaps deforming each other when the load is reduced toward K_{min} during load cycling.

closure, and so the drop in the SQRT(EPSD) of Specimen 8353 (with R = 0.1) should be more significant than that of Specimen 8354 (with R = 0.8). Figure 7 also shows this to be the case.

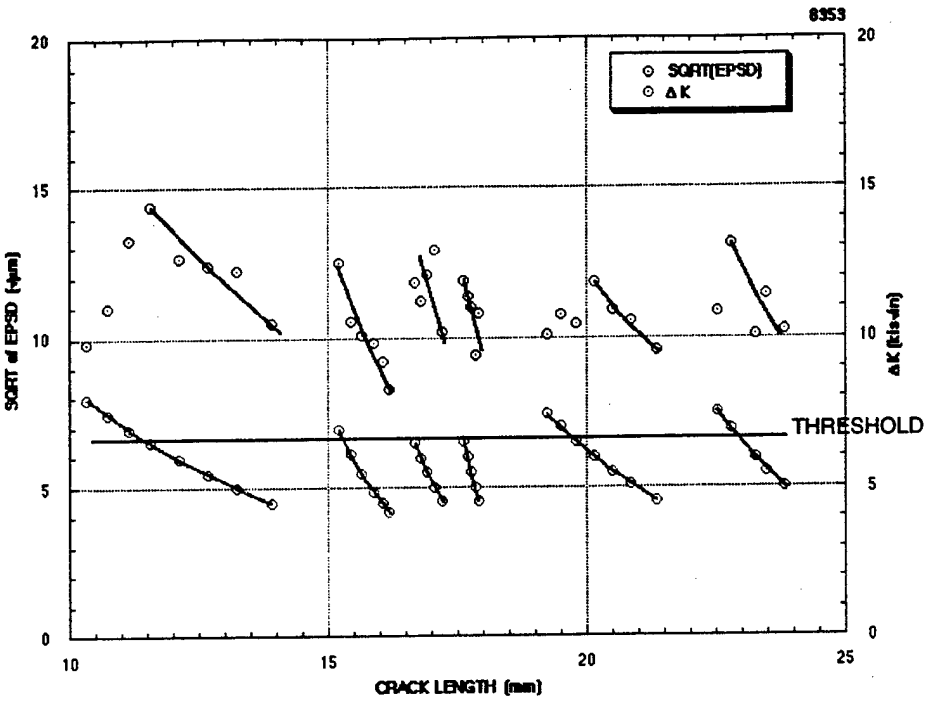
Thus crack closure can explain the results from the two specimens, so we examined the individual fatigue test results more closely. Figure 8 is a magnified view of the results shown in Figure 6. In these load shedding tests, the ΔK was decreased as the crack extended. When ΔK is above a certain threshold as shown in Figure 8, the SQRT(EPSD) exhibits low values and the higher the value of ΔK, the lower the SQRT(EPSD) value. If the ΔK values are below the threshold, the trend in SQRT(EPSD) change is similar to that of ΔK.

This threshold trend is observed also for tests in which ΔK was increased as the crack grew, Figure 8(b). Here the changes in SQRT(EPSD) above the threshold line are the same as for the load shedding test; i.e., the higher ΔK is above the threshold line, the lower the SQRT(EPSD) value. This trend is consistent with crack closure modifying the fracture surfaces at ΔK's above the threshold value.

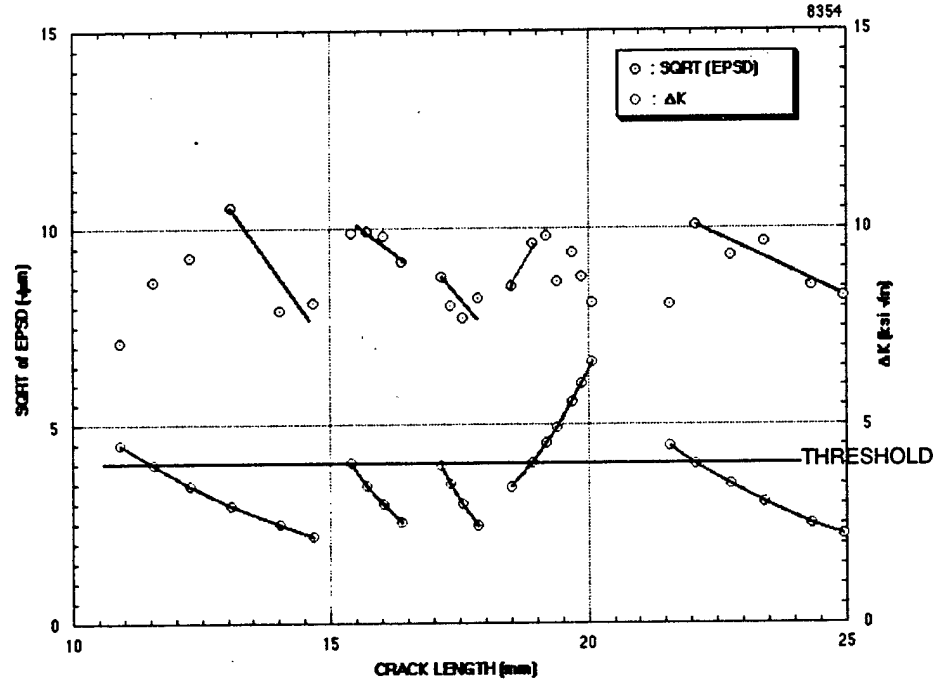
ADDITIONAL EFFECTS OF CRACK CLOSURE

To delineate the effect of crack closure on the surface roughness, we compared the EPSD curves from the fracture surface areas in the third region in Figure 7 at similar ΔK's of about 17 $MPa\sqrt{m}$ in the samples 8353 (R = 0.1) and 8354 (R = 0.8). The four curves in Figure 9 (two for R = 0.1 and the other two for R = 0.8) show that, within the same R-value, the EPSD curves have the same shape and are parallel each other; however, at different R-values, the curves have different shapes. The curves for R = 0.8 are nearly straight between wavelengths of 0.5 μm and 20 μm; however, the curves for R = 0.1 deviate from linear below 1 μm and above 4 μm. This suggests that crack closure more significantly affects the features below the scale of 1 μm and above 4 μm. In the transition region, the EPSD values decrease and the curves bow more as ΔK increases, as shown in Figure 10. This suggests that as ΔK increases, crack closure effects become more important.

The EPSD changes in the load shedding tests were also examined. We examined the first fatigue test in each sample where crack growth was initiated at ΔK well above the threshold line shown in Figure 8. EPSD values were initially low, but gradually increased. The shape of EPSD curves during this change is shown in Figure 11. The EPSD curves corresponding to initial high ΔK exhibit bowing; however, as ΔK decreases below the threshold ΔK, the value the EPSD curves become straighter. Again these curves suggest that the higher ΔK loading induces crack closure effects that cause bowing of the EPSD curves.



(a) Results for 8353



(b) Results for 8354

NM-312581-7

Figure 8. Possible correlation of SQRT(EPSD) behavior with ΔK behavior.

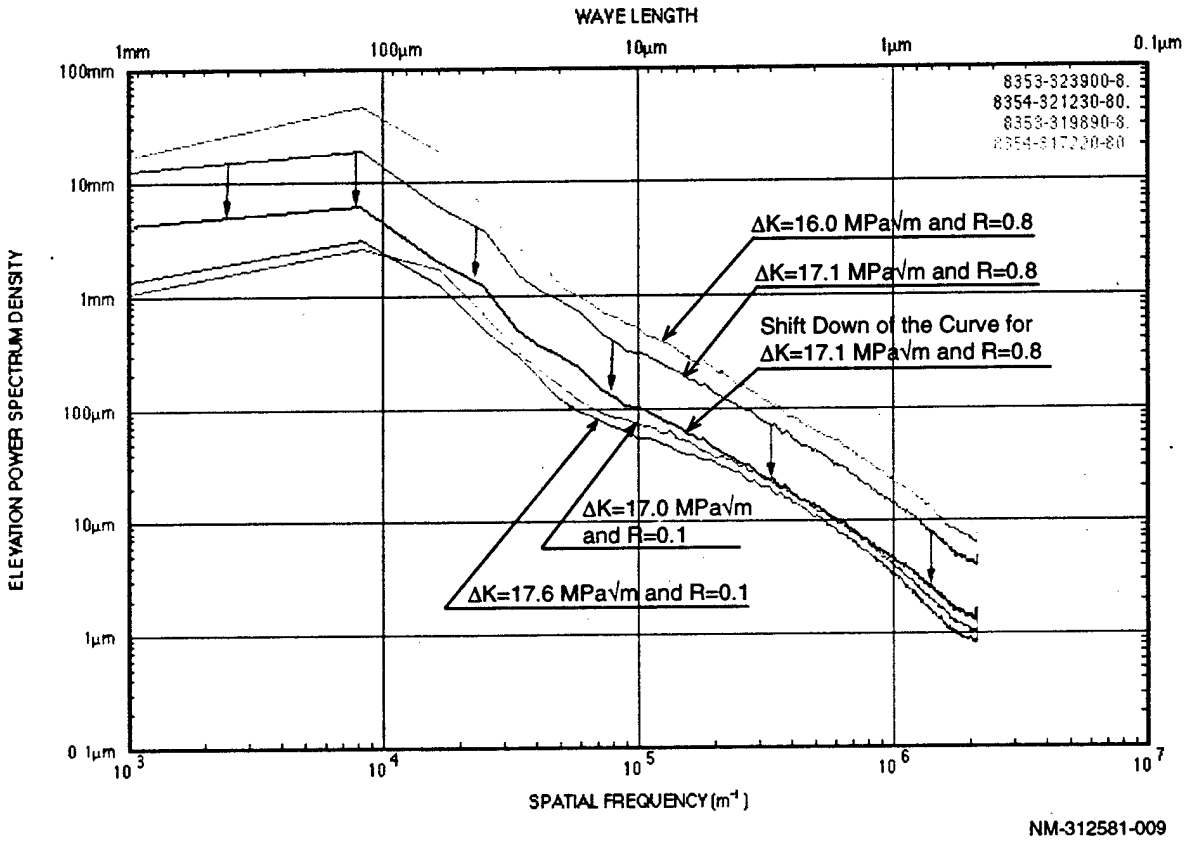


Figure 9. Comparison of EPSD curves for tests under similar stress intensity range but different stress ratios.

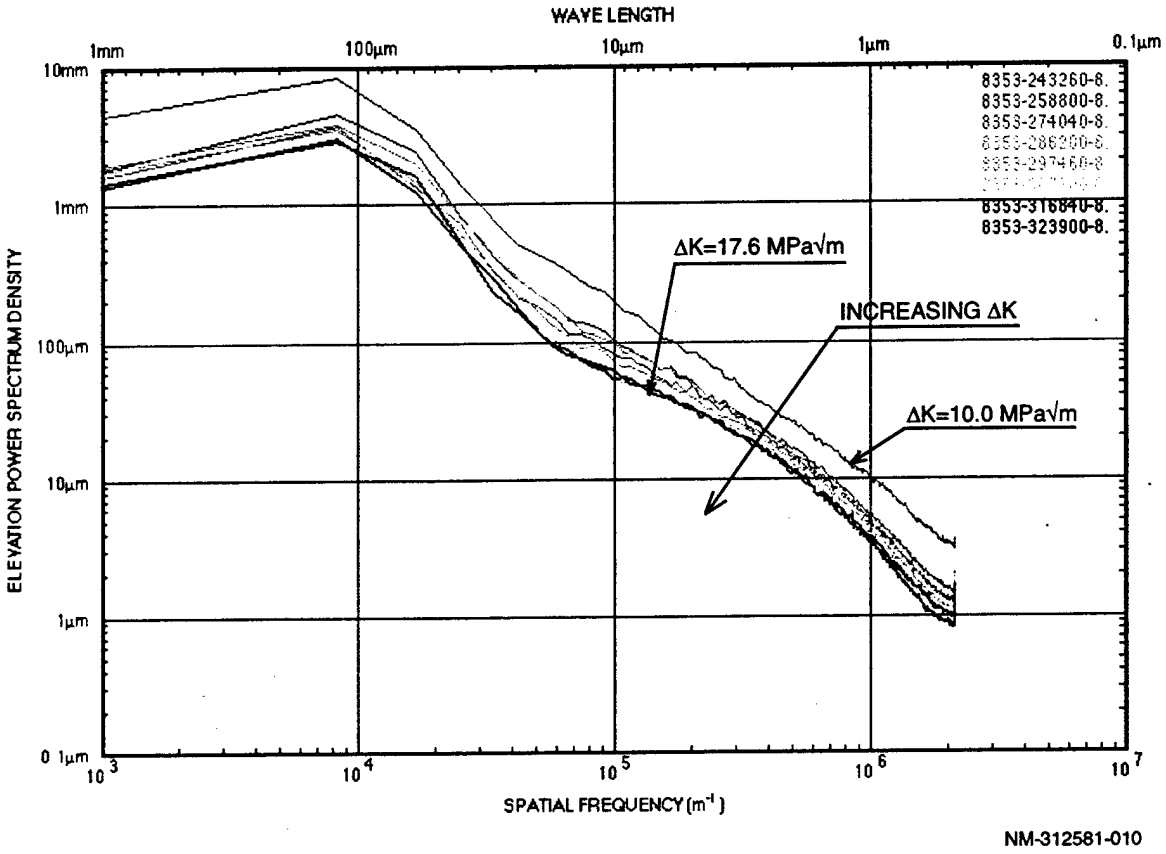
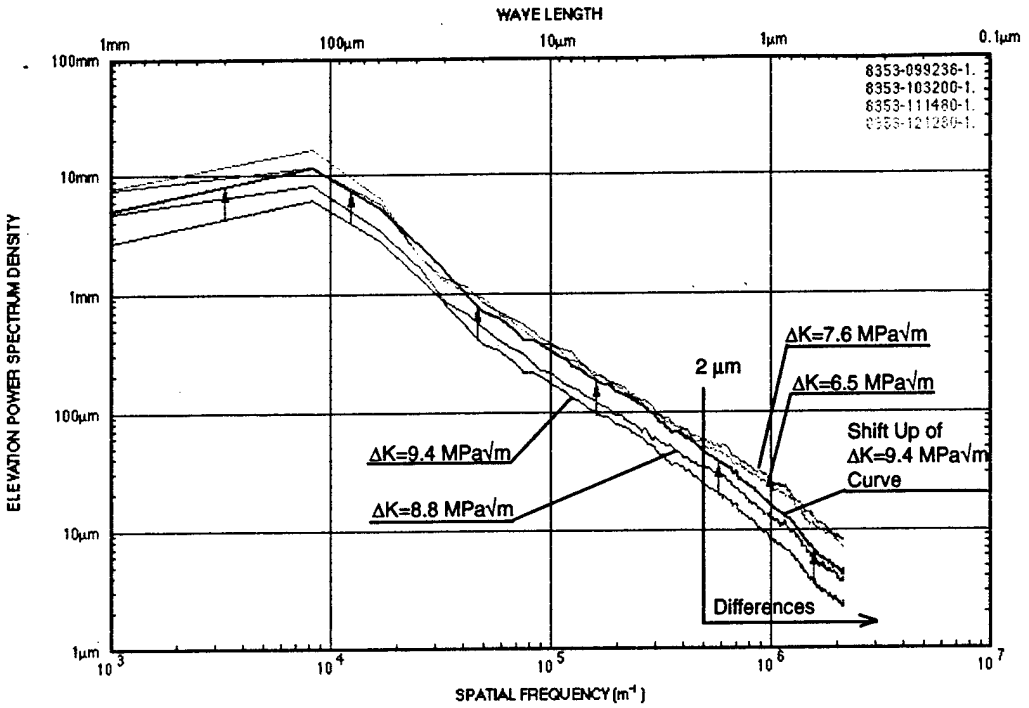
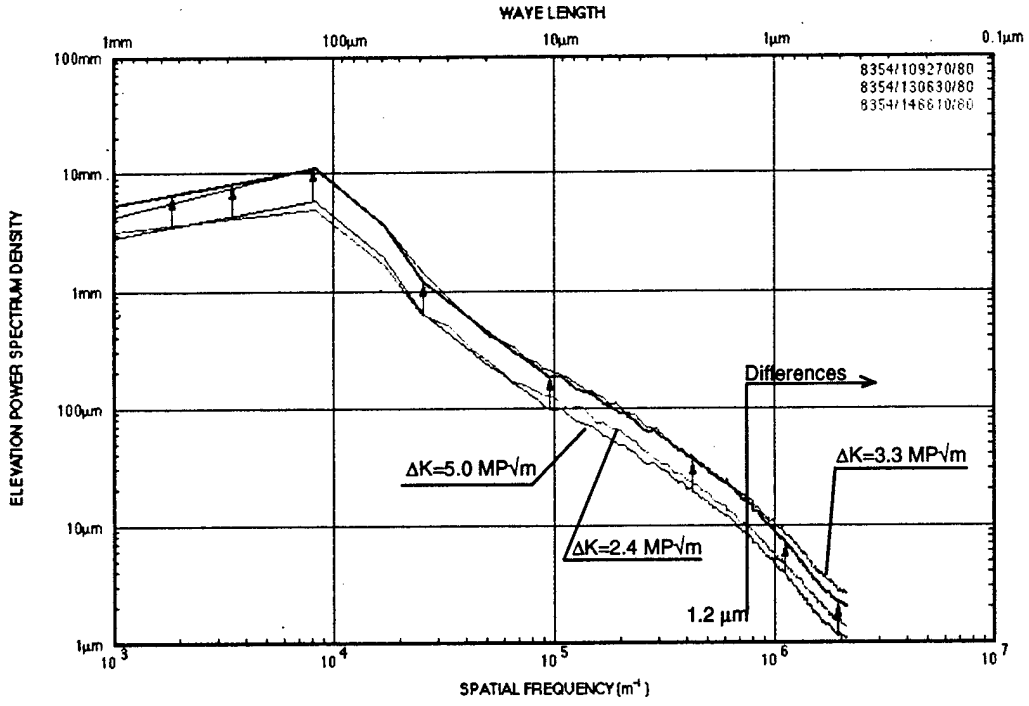


Figure 10. Changes observed in the shape of EPSD curves as ΔK increases in the transition zone of R = 0.1 specimen.



(a) EPSD Curves of Run 1 in Specimen 8353 (R = 0.1)



(b) EPSD Curves of Run 1 in Specimen 8354 (R = 0.8)

NM-312581-011

Figure 11. EPSD curves showing changes in their shape with ΔK for two specimens.

We looked in the technical literature for results from other investigators that might support these findings. Figure 12 shows an observation by Newman of crack opening stress ratios for 2024-T3 aluminum alloy at low and high stress ratios.[14] A striking similarity exists in the shape of the curve for R = 0.7 with our SQRT(EPSD) curve for R = 0.8 shown in Figure 8. The curve in Figure 12 suggests that, due to the crack closure and re-deformation of the fracture topography, less stress is necessary to open the crack. In other words, the lower the opening stress, the higher the effect of surface deformation and flattening of the surface.

WAVELET ANALYSIS

Wavelet analysis was applied to the topographic data files by Professor Naoki Saito and Dr. Jean-Marie Aubry of the Mathematics Department of the University of California at Davis (Appendix A). The objectives were (1) to classify the fracture surfaces according to stress intensity range, ΔK, and (2) to identify signatures that specify the loading condition and can be used to discriminate others.

To conduct the classification study, the available data set needed to be split into training and test data sets. The training data set was used to construct the classification rules, and the test data set was used to access and evaluate whether such rules can predict the classes of the fracture surfaces.

The Local Discriminate Basis[15] (LDB) algorithm first decomposes the training signals into a basis dictionary, which is a large collection of the basis functions (such as wavelet packets and local cosine functions) localized in both space and spatial-frequency. Then, it computes discriminant information at each coordinate in the dictionary. The original LDB algorithm uses the difference of energy distributions as the discriminant measure, whereas the modified LDB[16] (we call this LDB+) uses the probability density function (pdf) in each coordinate, i.e., the way the coefficients are distributed. Then, a complete orthonormal basis called LDB (or LDB+), which "can see" the distinguishing signal features among signal classes, is selected from the dictionary by maximizing such discriminant information. After the basis is determined, expansion coefficients in the most important several coordinates (features) are fed into a traditional classifier such as the Linear Discriminant Analysis (LDA). Finally, the corresponding coefficients of the test signals are computed and fed to the classifier to predict their classes.

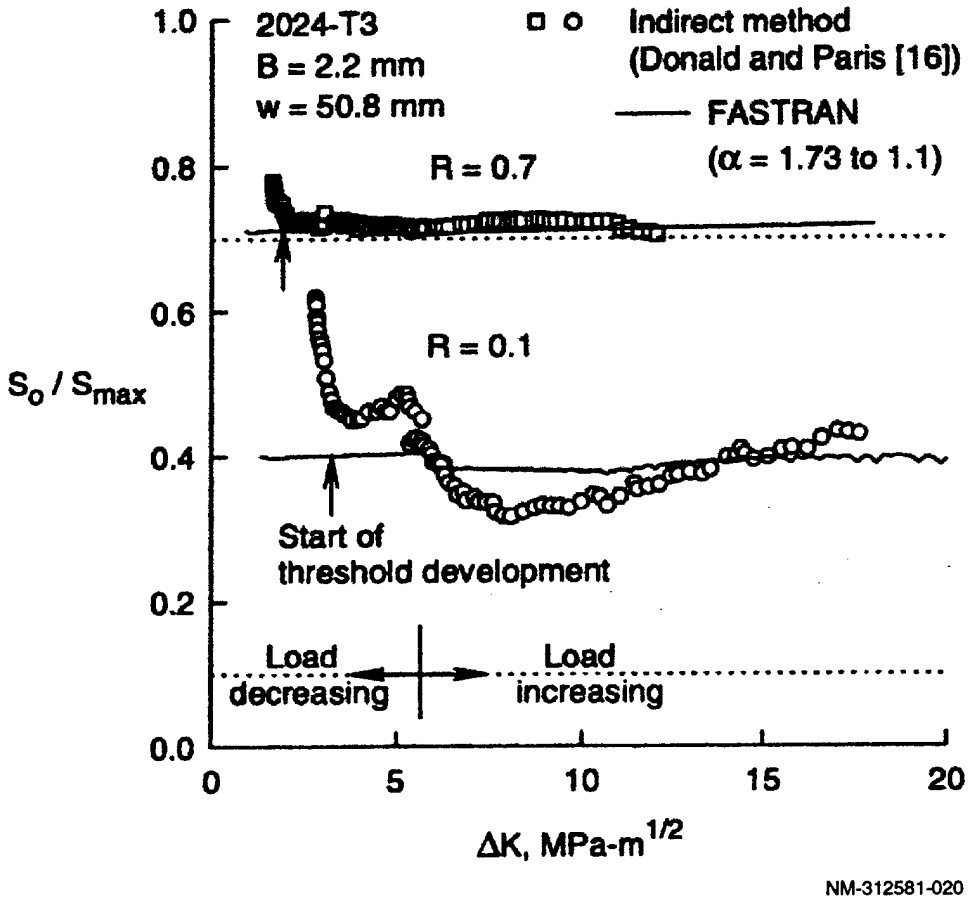


Figure 12. Crack-opening stress ratios for 2024-T3 aluminum alloy at low and high stress ratios. (Ref. 14)

Our experiments can be grouped into two categories:

(1) Application of 1D LDB algorithms along either a row or column or along a circular path with varying parameters (normalizations, filters, discriminant measures, number of features, space-domain data, or frequency-domain data).

(2) Application of 2D LDB algorithms to the same data sets viewed as images.

Summary of 1D Classification Results

- In general, the greater the difference in stress intensity range, the lower the misclassification rates.
- The misclassification rate varies with the number of LDB features used.
- If both the training and test data sets are selected from the same region randomly, then the misclassification rates are very small (0.4% ~ 17%).
- If the training and test data sets are selected from two different regions (no overlapping) without randomization, then the misclassification rates are much higher (8.4% ~ 41%).

Summary of 2D Classification Results

- We computed the 2D power spectrum (modulus of the 2D Fourier transform) of each image as the basic data set.
- The best and most robust results were obtained by the LDB based on the local cosine dictionary.
- Again, the more different the stress intensity factors, the lower the misclassification rates.
- If both the training and test data sets are selected from the same region without overlap (checker board pattern), then the misclassification rates are relatively small (20% ~ 35%).
- If the training and test data sets are selected from two different regions, then the misclassification rates are slightly higher (30% ~ 35%).
- A misclassification rate < 50% for test data implies that a majority rule based on many independent samples yields a low probability of error of that region.
- The variance of misclassification rates for the 2D results is much smaller than for the 1D case.

These preliminary results suggest that wavelet analysis principles are well suited to overcome the challenges of extracting and discriminating fracture surface features. However, the technique must be extended and adapted to address issues of noise in the data and the linking of information over a range of size scales. Moreover, more decisive results will likely be obtained by increasing the size of the area chosen for LDB application. The results are discussed in a later section, and suggestions for modifying the procedure are discussed in the section entitled Next Steps.

The success achieved in correlating Fourier results with load parameters, however, encouraged us to attempt an analysis of a field failure. In the next section, we describe how we applied the findings from the laboratory specimens to interpret the fracture surfaces of a crack in an aircraft turbine component. Appendix B describes how we applied fracture surface topographic analysis to assist a team investigating the failure of a space shuttle component.

ANALYSIS OF FIELD FAILURES

Two aircraft engine components that had acquired cracks during service (Figure 13) were provided to SRI for analysis.[17] Component 39 had a part-through crack along its central axis; Component 20 had a through-crack. Figure 14 shows the fracture surfaces. The discolored regions on both components are probably oxide films that formed during service. The crack in Component 39 apparently initiated at a single site near the midpoint, whereas the crack in Component 20 originated from several small cracks, which nucleated near the midpoint and merged. The concentric elliptical markings indicate the crack fronts at various stages of their advance.

TOPOGRAPHY CHARACTERIZATION

We chose to analyze small areas of the crack faces at regular distances (indicated by the tick marks in Figure 15) from the crack origins along a line A-A normal to the component surface. A confocal optics scanning laser microscope was used to produce elevation maps of small areas of the fracture surface 140 x 89 μm at 250-μm intervals along a line from the crack origin to the opposite component surface. This area was divided into a 600 x 400 pixel array, and the elevation of each pixel was measured. The data spacing was 0.233 μm in each direction.

Figures 16 and 17 show a series of contrast images and gray-scale topography images. The shades of gray indicate relative height: lighter areas are higher, and darker areas are lower. The areas near the component exteriors, which were covered by the film, are smoother than elsewhere and hence were omitted in the evaluation.

FOURIER ANALYSIS AND FRACTURE PARAMETER CORRELATION

Differences in the topographies arising from differences in loading conditions are not easily discerned from Figures 16 and 17. To obtain an alternative representation of the topography, we applied a one-dimensional fast Fourier transform operation, row by row, to 400 rows of 512 (2^9) elevation data points. Figures 18 and 19 show the energy power spectral density (EPSD) curve for each of the areas in Figures 16 and 17 obtained by averaging the 400 rows. In Figures 18 and 19, the square root of the sum of the squares of the real and imaginary components of the elevation amplitudes normalized by a crack extension distance is presented as a function of

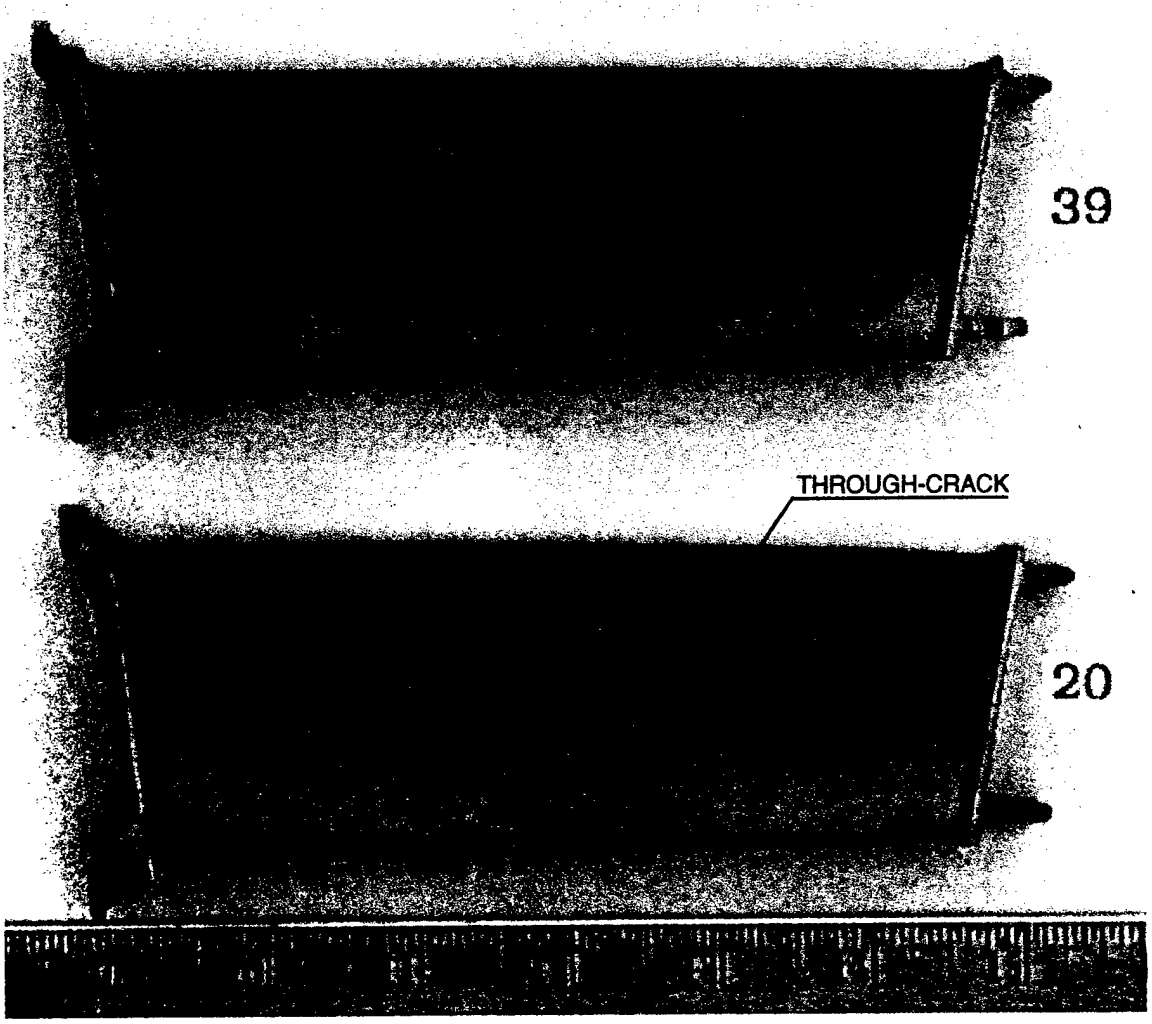
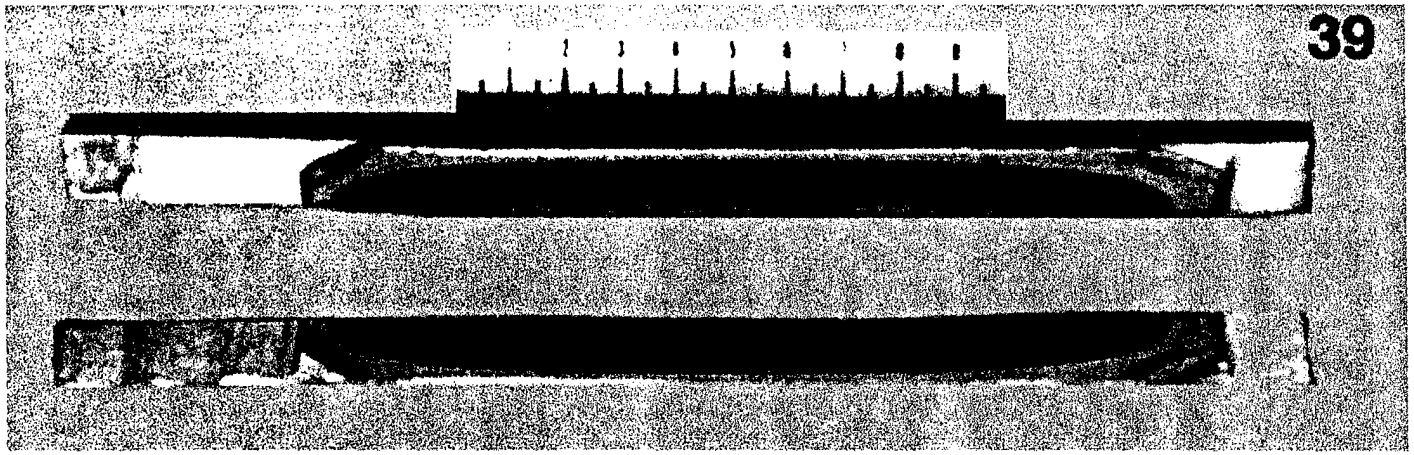
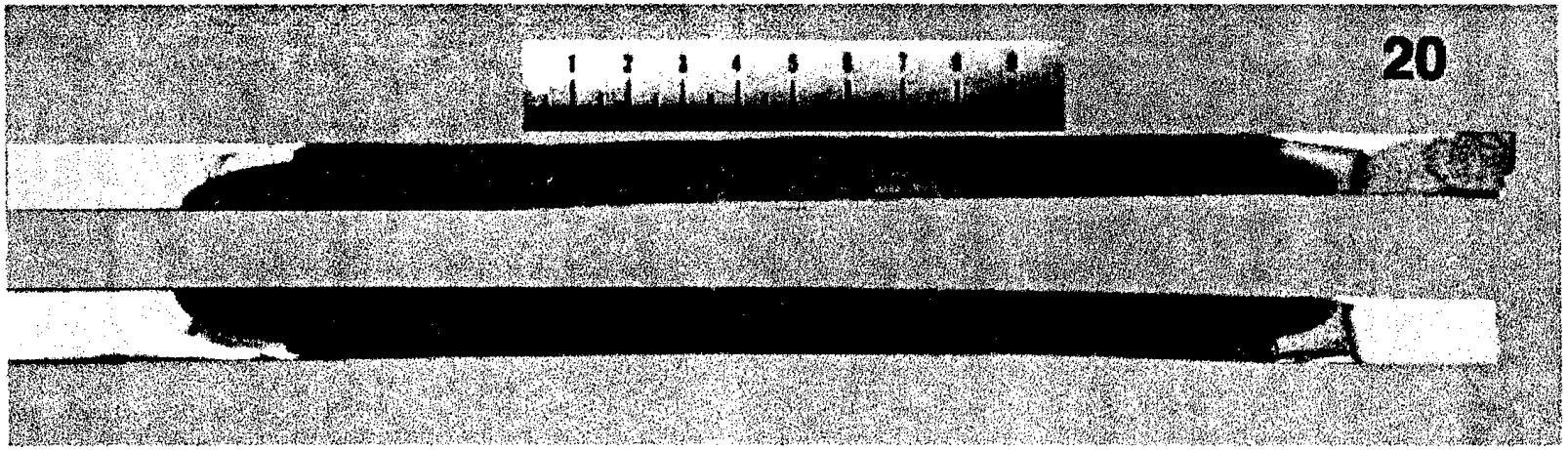


Figure 13. Two aircraft engine components that acquired cracks during service. (Component 39 had a part-through crack not visible in this photo, and Component 20 had a through-crack along the component axis.)



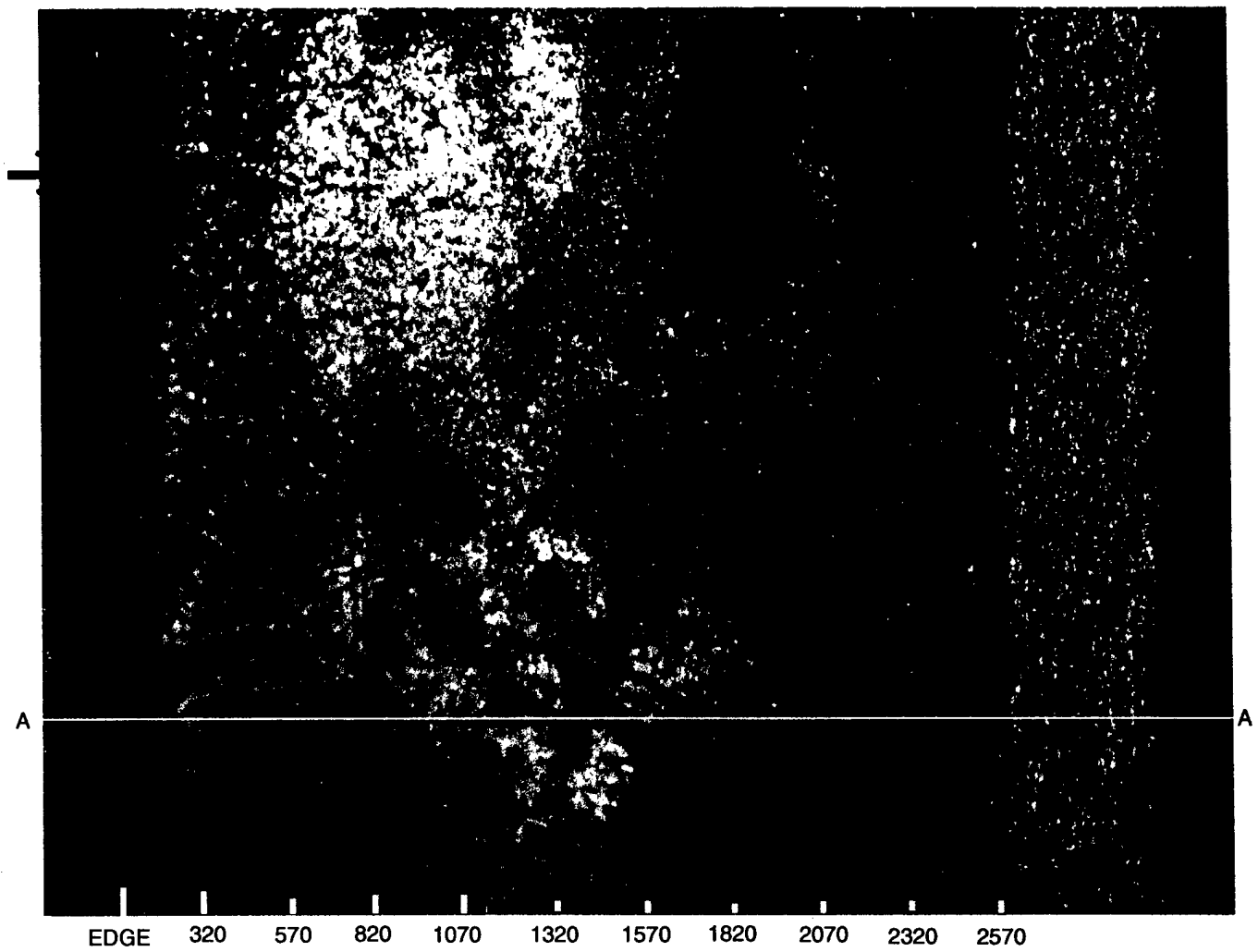
(a) Fracture surfaces of Component 39



(b) Fracture surfaces of Component 20

NM-2037-2

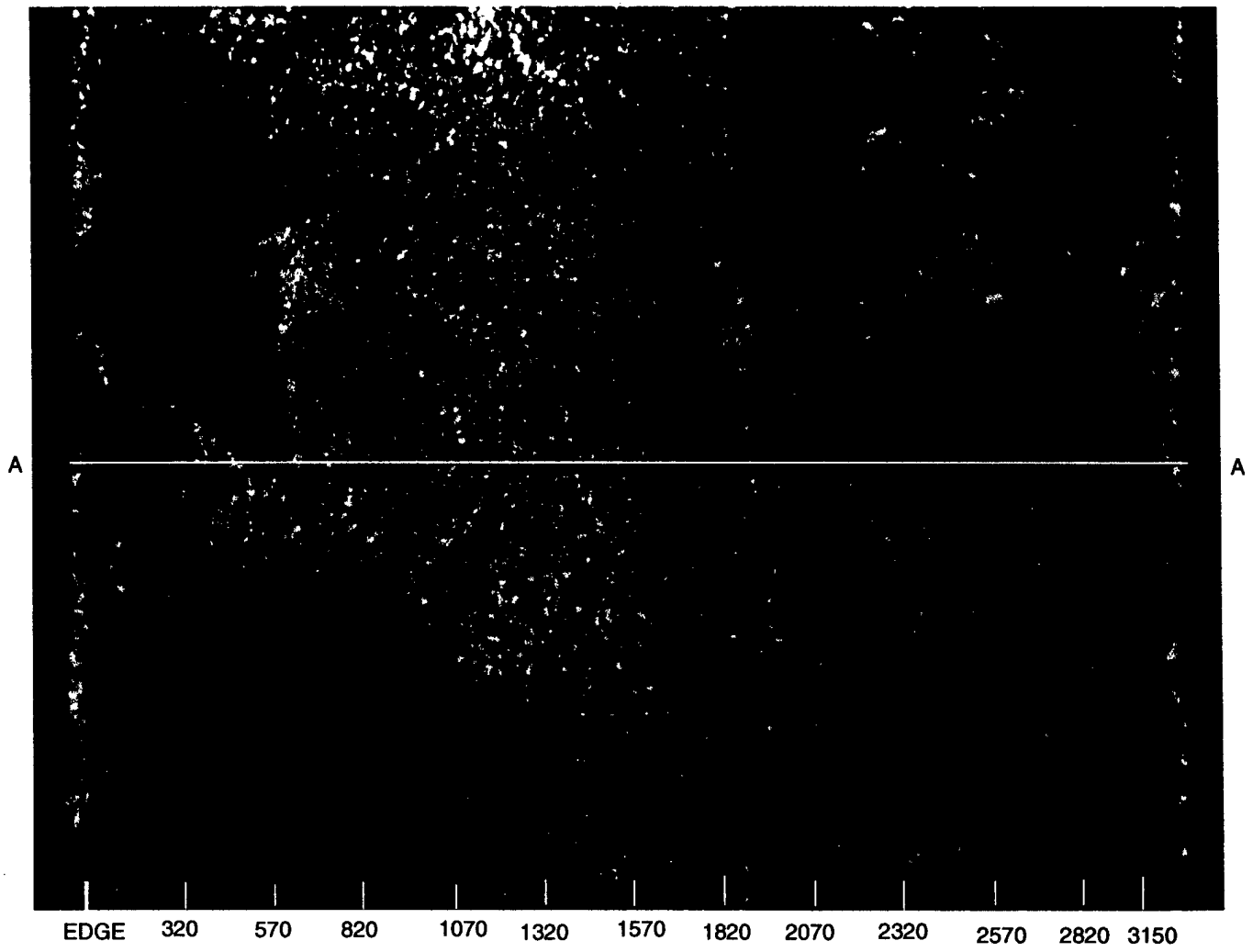
Figure 14. Plan view of fracture surfaces in Components 39 and 20.



(a) Fracture surface of Component 39
(Arrow on the left edge of the picture indicates the possible crack initiation site.)

NM-2037-003

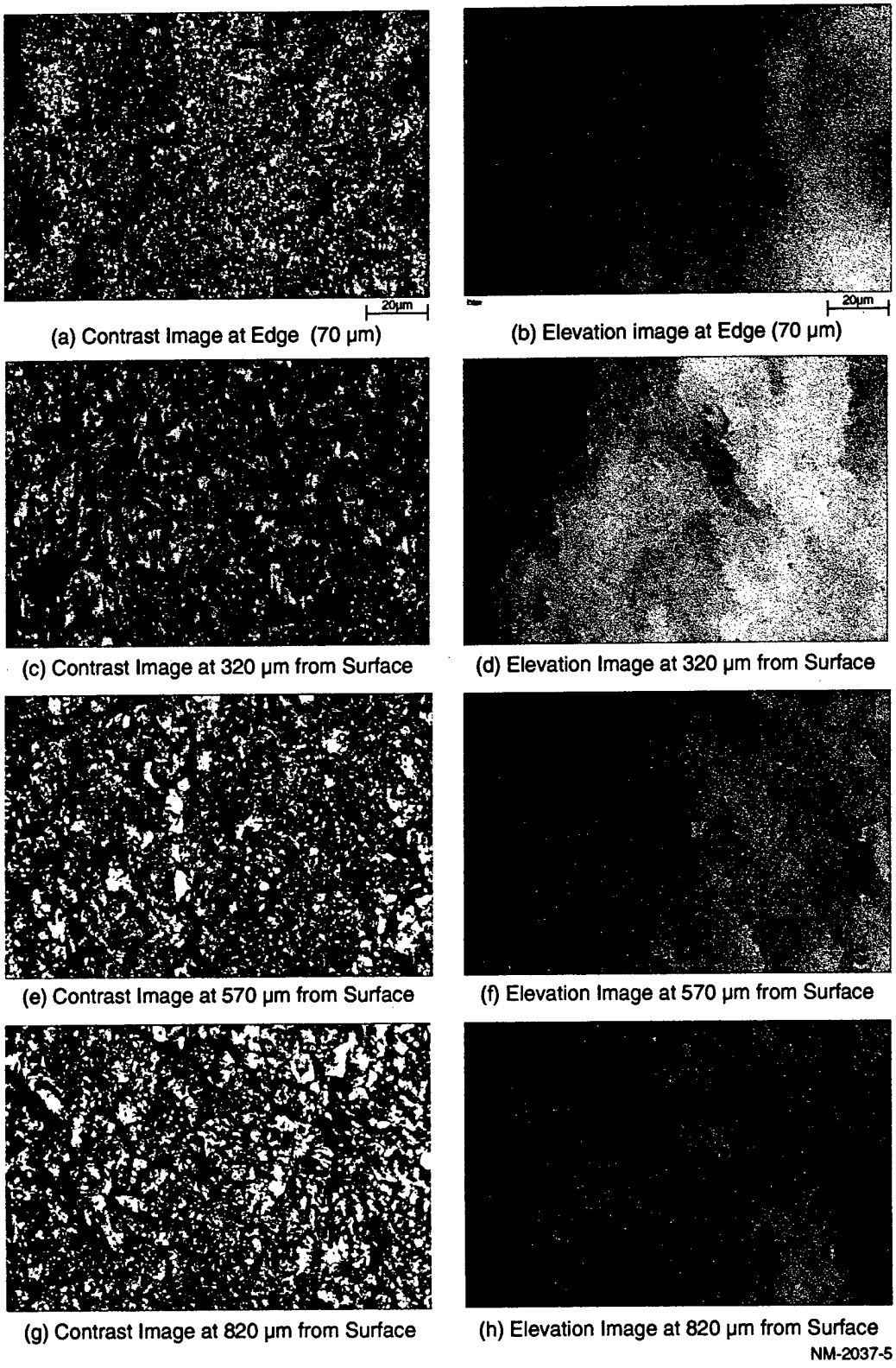
Figure 15. Fracture surface of Components 39 and 20 showing probable crack nucleation sites.
(The white Line A-A and tick marks indicate the location of the areas examined in detail.)



(b) Fracture surface of Component 20
(An oxide film is evident near both exterior surfaces.)

NM-2037-004

Figure 15. Fracture surface of Components 39 and 20 showing probable crack nucleation sites.
(The white Line A-A and tick marks indicate the location of the areas examined in detail.) (concluded)



(a) Contrast Image at Edge (70 µm)

(b) Elevation image at Edge (70 µm)

(c) Contrast Image at 320 µm from Surface

(d) Elevation Image at 320 µm from Surface

(e) Contrast Image at 570 µm from Surface

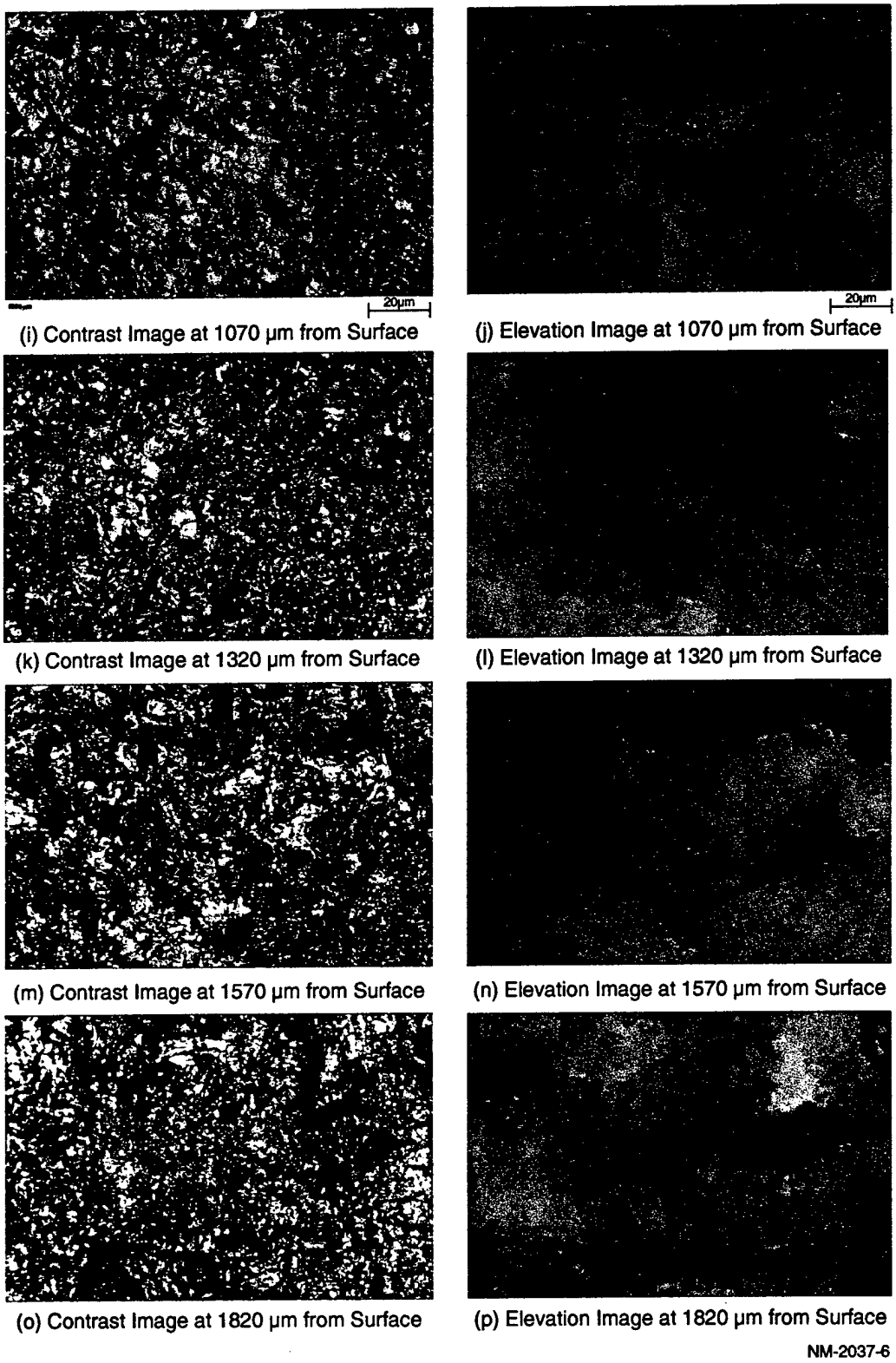
(f) Elevation Image at 570 µm from Surface

(g) Contrast Image at 820 µm from Surface

(h) Elevation Image at 820 µm from Surface

NM-2037-5

Figure 16. A series of contrast and gray-scale elevation images at various locations along Line A-A for Component 39.



(i) Contrast Image at 1070 µm from Surface

(j) Elevation Image at 1070 µm from Surface

(k) Contrast Image at 1320 µm from Surface

(l) Elevation Image at 1320 µm from Surface

(m) Contrast Image at 1570 µm from Surface

(n) Elevation Image at 1570 µm from Surface

(o) Contrast Image at 1820 µm from Surface

(p) Elevation Image at 1820 µm from Surface

NM-2037-6

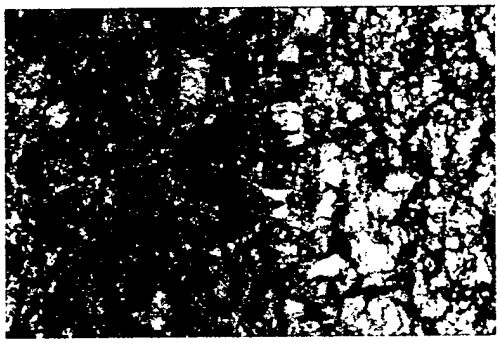
Figure 16. A series of contrast and gray-scale elevation images at various locations along Line A-A for Component 39. (continued)

(q) Contrast Image at 2070 μm from Surface

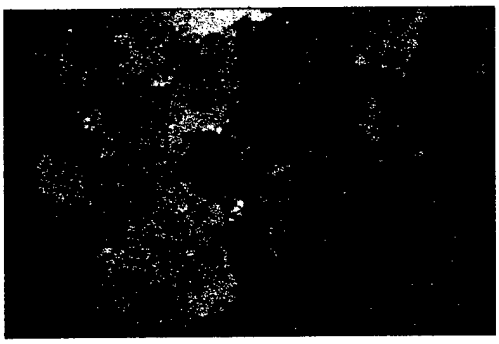
(r) Elevation Image at 2070 μm from Surface

(s) Contrast Image at 2320 μm from Surface

(t) Elevation Image at 2320 μm from Surface



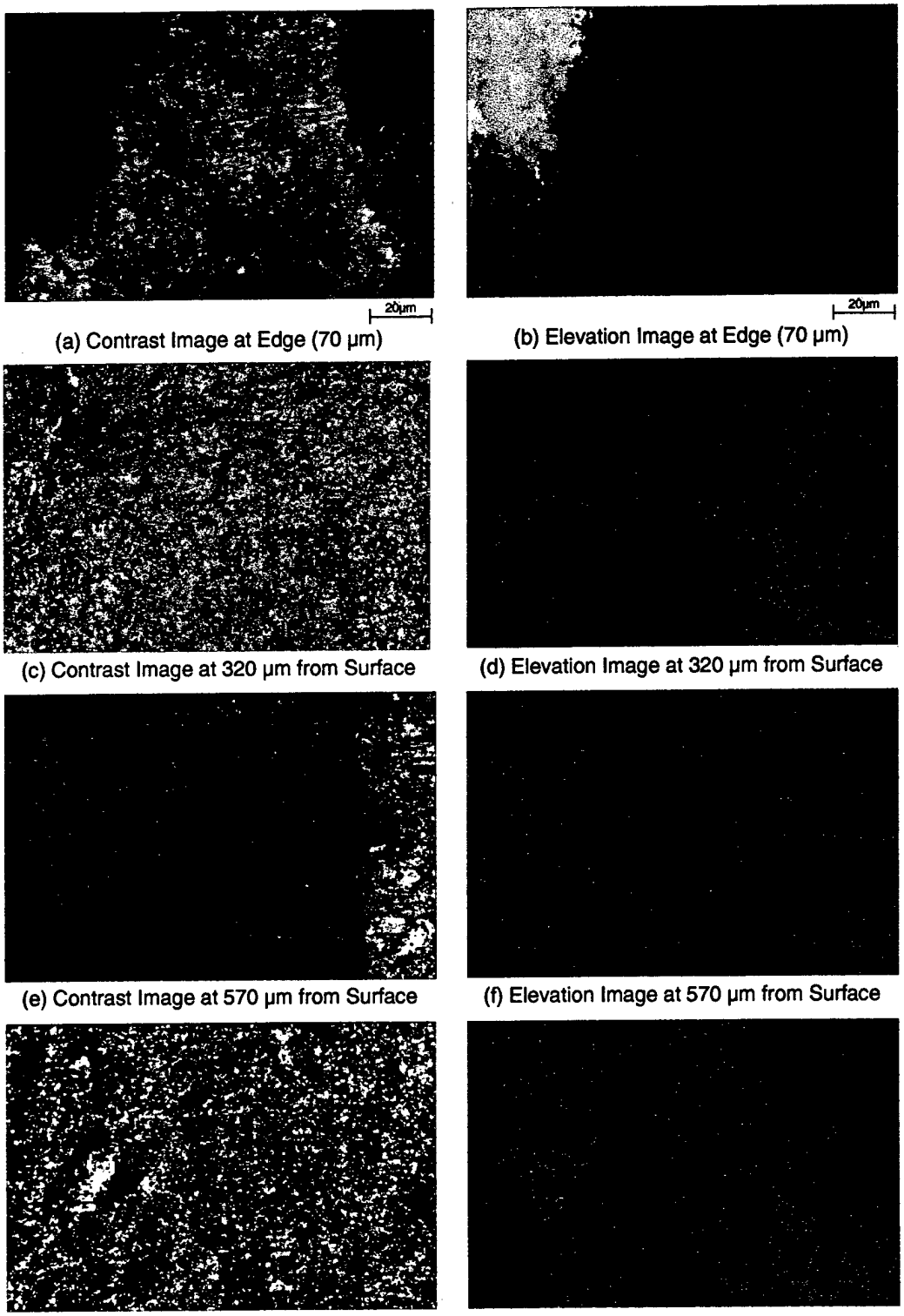
(u) Contrast Image at 2570 μm from Surface



(v) Elevation Image at 2570 μm from Surface

NM-2037-7

Figure 16. A series of contrast and gray-scale elevation images at various locations along Line A-A for Component 39. (concluded)



(a) Contrast Image at Edge (70 μm)

(b) Elevation Image at Edge (70 μm)

(c) Contrast Image at 320 μm from Surface

(d) Elevation Image at 320 μm from Surface

(e) Contrast Image at 570 μm from Surface

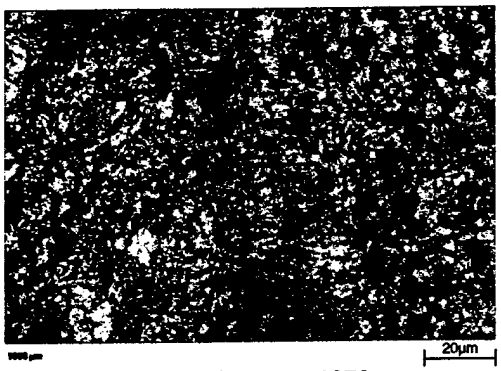
(f) Elevation Image at 570 μm from Surface

(g) Contrast Image at 820 μm from Surface

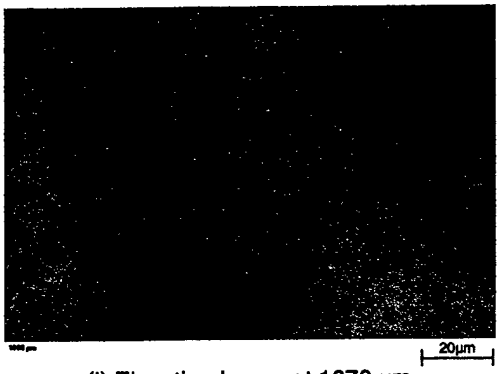
(h) Elevation Image at 820 μm from Surface

NM-2037-8

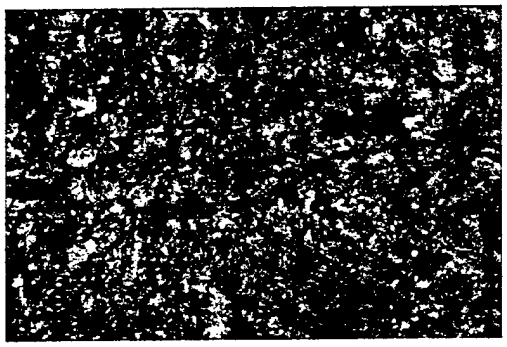
Figure 17. A series of contrast and gray-scale elevation images at various locations along Line A-A for Component 20.



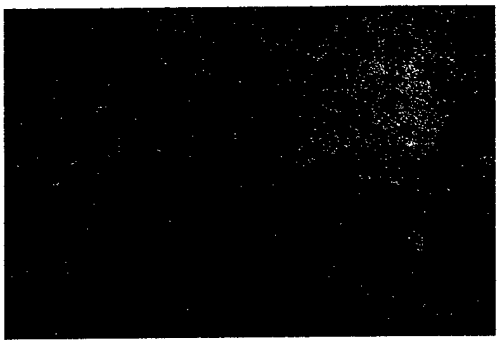
(i) Contrast Image at 1070 µm



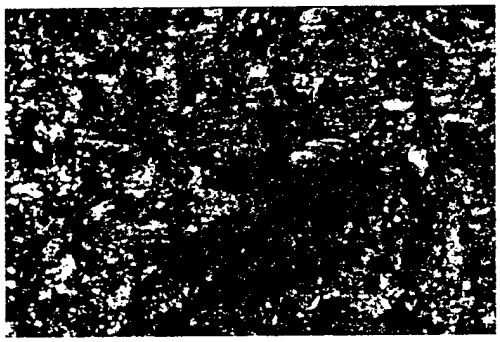
(j) Elevation Image at 1070 µm



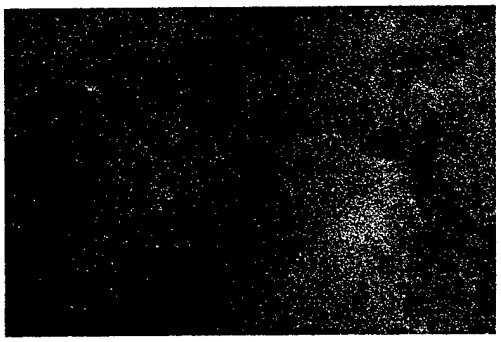
(k) Contrast Image at 1320 µm



(l) Elevation Image at 1320 µm



(m) Contrast Image at 1570 µm



(n) Elevation Image at 1570 µm

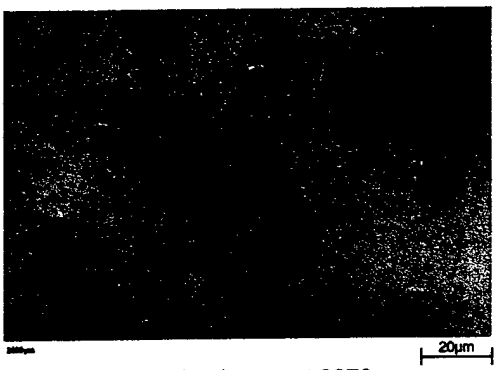
(o) Contrast Image at 1820 µm

(p) Elevation Image at 1820 µm

NM-2037-9

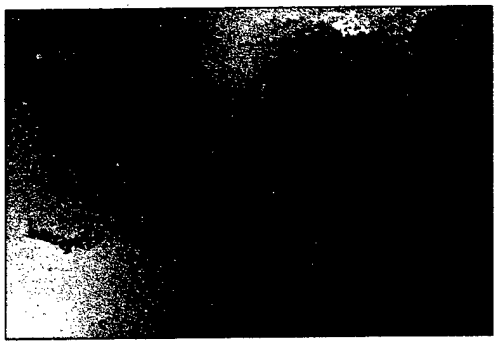
Figure 17. A series of contrast and gray-scale elevation images at various locations along Line A-A for Component 20. (continued)

(q) Contrast Image at 2070 µm

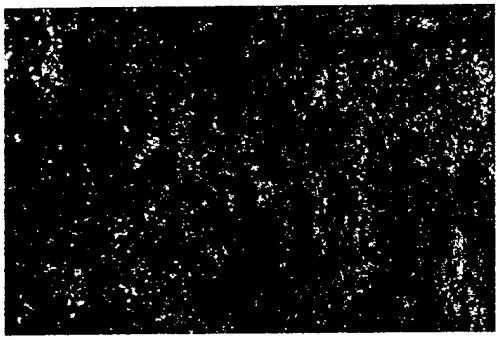


(r) Elevation Image at 2070 µm

(s) Contrast Image at 2320 µm

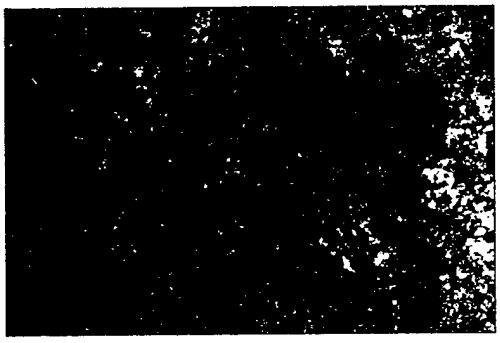


(t) Elevation Image at 2320 µm

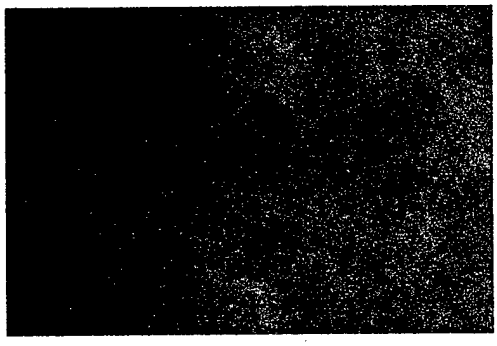


(u) Contrast Image at 2570 µm

(v) Elevation Image at 2570 µm



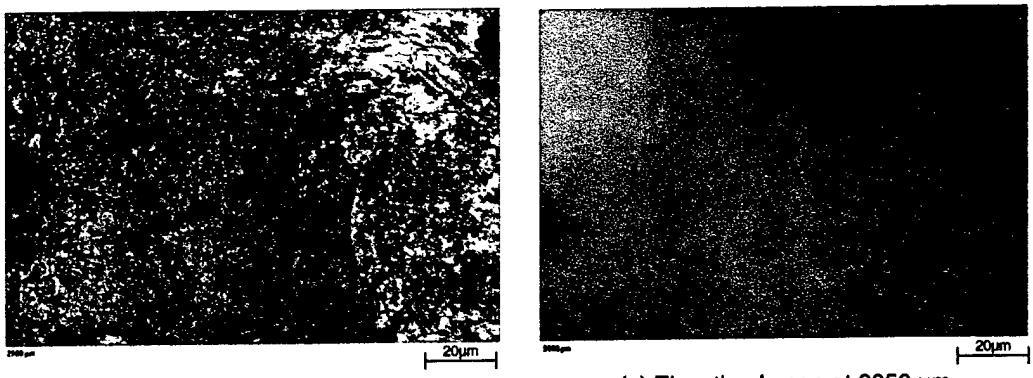
(w) Contrast Image at 2820 µm



(x) Elevation Image at 2820 µm

NM-2037-10

Figure 17. A series of contrast and gray-scale elevation images at various locations along Line A-A for Component 20. (continued)



(y) Contrast Image at 3050 µm

(z) Elevation Image at 3050 µm

NM-2037-11

Figure 17. A series of contrast and gray-scale elevation images at various locations along Line A-A for Component 20. (concluded)

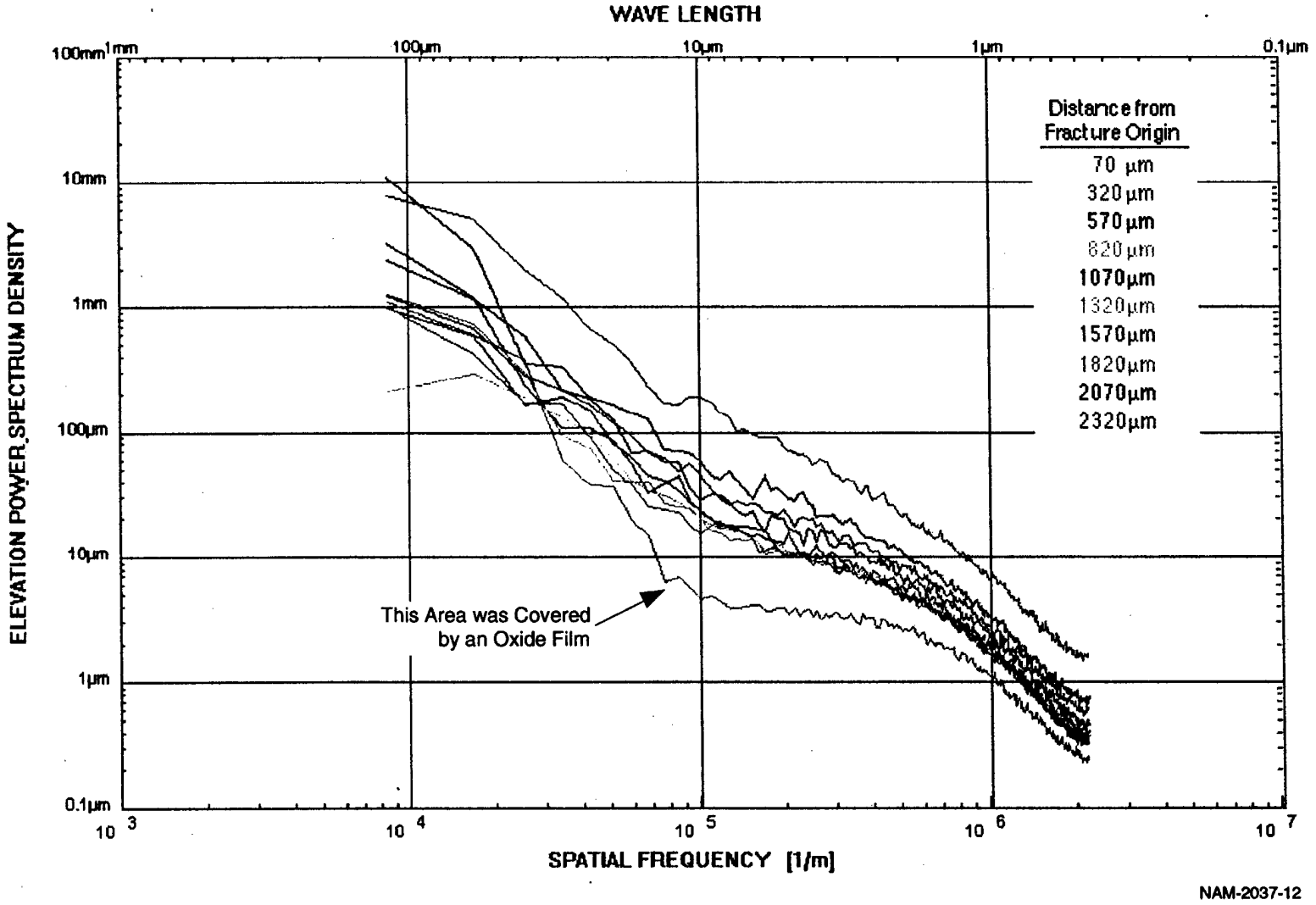


Figure 18. Elevation power spectrum density curves at various positions along the Line A-A for Component 39.

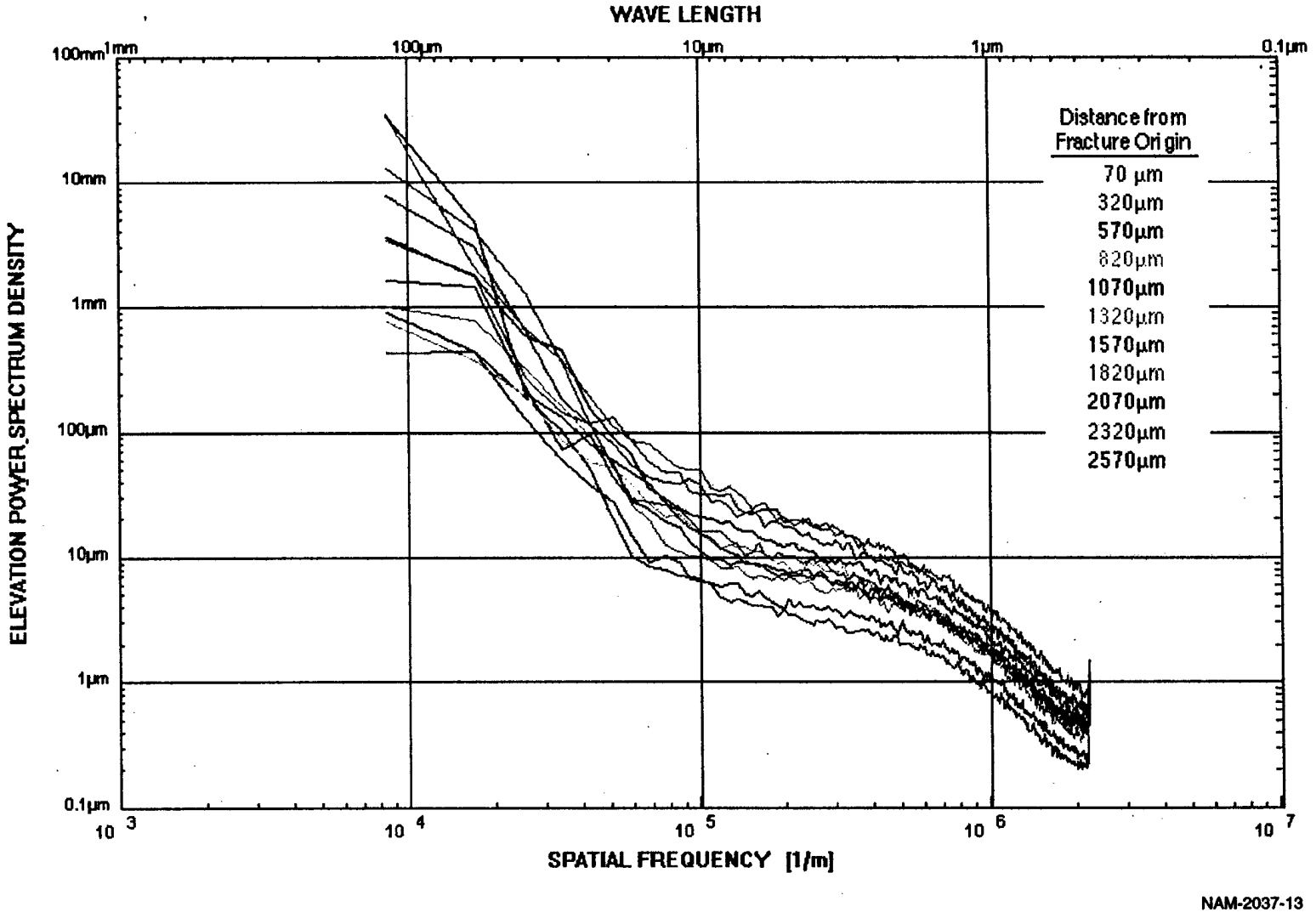


Figure 19. Elevation power spectrum density curves at various positions along the Line A-A for Component 20.

spatial frequency or roughness wavelength, rather than as a function of position (as in Figures 16 and 17).

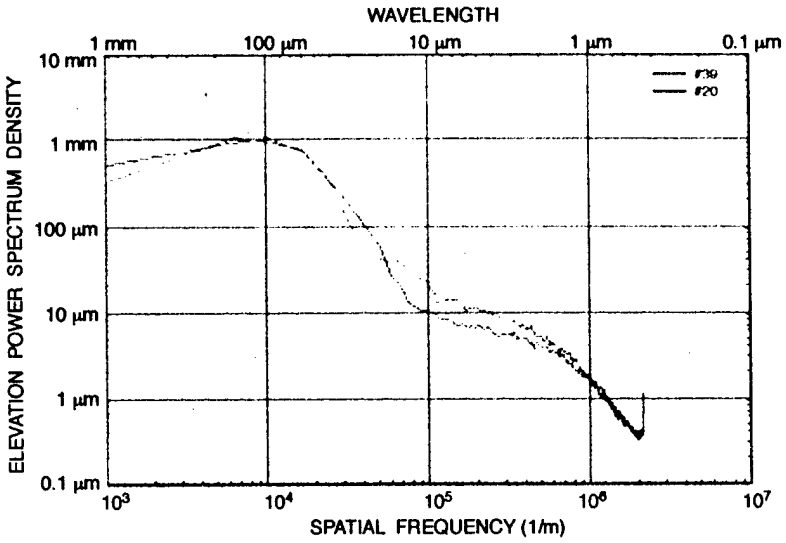
The averaged EPSD curves for the various fracture surface areas vary in height and shape and show a characteristic trend with crack length. The curves are roughly parallel to one another between wavelengths of 0.5 μm and 15 μm [with the exception of the EPSD curves from the areas closest to the fracture origins that were covered by the oxide film, see Figures 20(a) and 20(d)]. At distances 1.32 to 1.82 mm from the fracture origin, the curves for the two components are nearly identical; see Figures 20b and 20c, for example.

Figure 21 shows how the surface roughness varies with crack depth. We arbitrarily selected the EPSD at a wavelength of 5 μm and plotted the square root of this value versus crack depth. (We selected the square root function of EPSD because we observed a linear relationship between the square root of EPSD value and ΔK, during analysis of the laboratory specimens.) If we interpret this J-shaped roughness variation with crack length for Component 39 as the variation of ΔK, then crack initiation and early growth occurred at a relatively high ΔK, cracking in the component interior occurred at lower ΔK, and crack growth closer to the far end of the specimen occurred at increasing ΔK. This interpretation suggests that the component failed under some mode of bending, because bending produces stresses that are highest at the component exteriors and minimum at the central neutral axis.

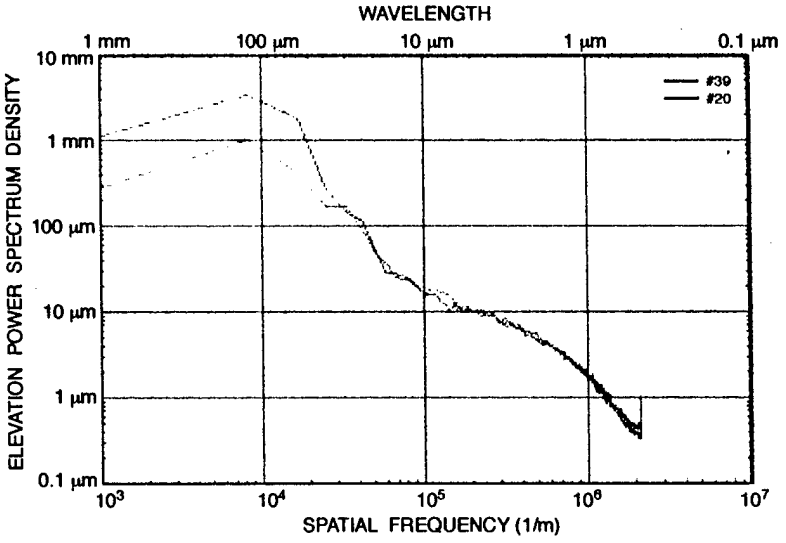
The curve for Component 20 lies close to that for Component 39, although fewer data points could be acquired because a larger portion of the fracture surface was covered by the oxide film. The effect of the oxide film is to lower the EPSD values, as seen in Figure 18 and Figures 20(a) and 20(d), but the results for the two components are consistent when no oxide exists on the surface.

Thus, the analysis so far has provided an answer to one aspect of the load question—the mode of loading. The next question is the magnitude of the load. With the lack of a value from analysis considerations, we turned to correlation with data from the laboratory specimens where the loading conditions were known.

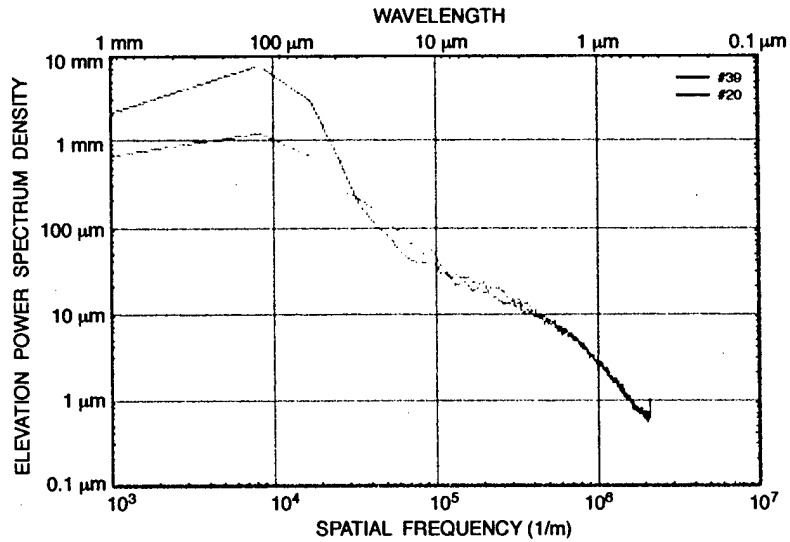
Comparisons were made from two viewpoints: (1) the shapes of the EPSD curves and (2) the heights of the EPSD curves. The shape of an EPSD curve seems to be sensitive to the stress-ratio, R. An EPSD curve from a higher R-value surface is straighter than that from a lower R-value surface. Figure 22 compares EPSD curves of Component 39 with curves from a



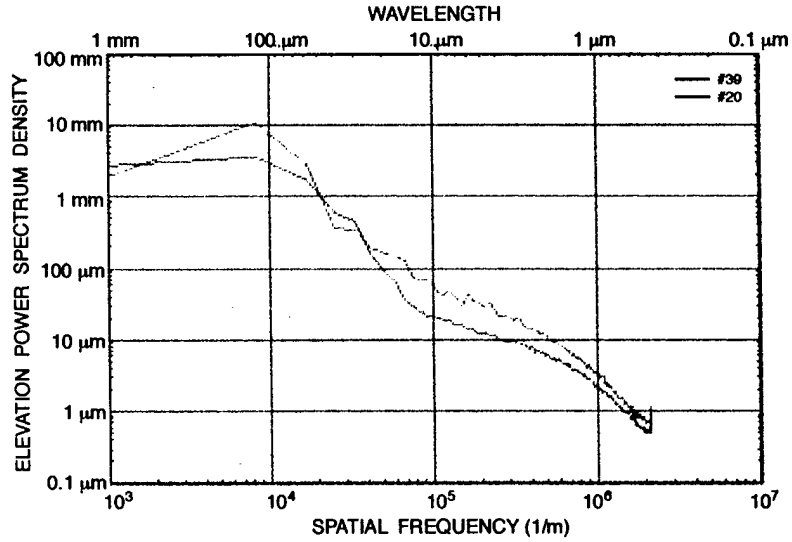
(a) Comparison of EPSD curve of Components 39 and 20 at a location 0.82 mm from the edge of component



(b) Comparison of EPSD curve of Components 39 and 20 at a location 1.57 mm from the edge of component



(c) Comparison of EPSD curve of Components 39 and 20 at a location 1.82 mm from the edge of component



(d) Comparison of EPSD curve of Components 39 and 20 at a location 2.07 mm from the edge of component

NAM-2037-14

Figure 20. EPSD curves of Components 39 and 20 at similar distances from the exterior surface.

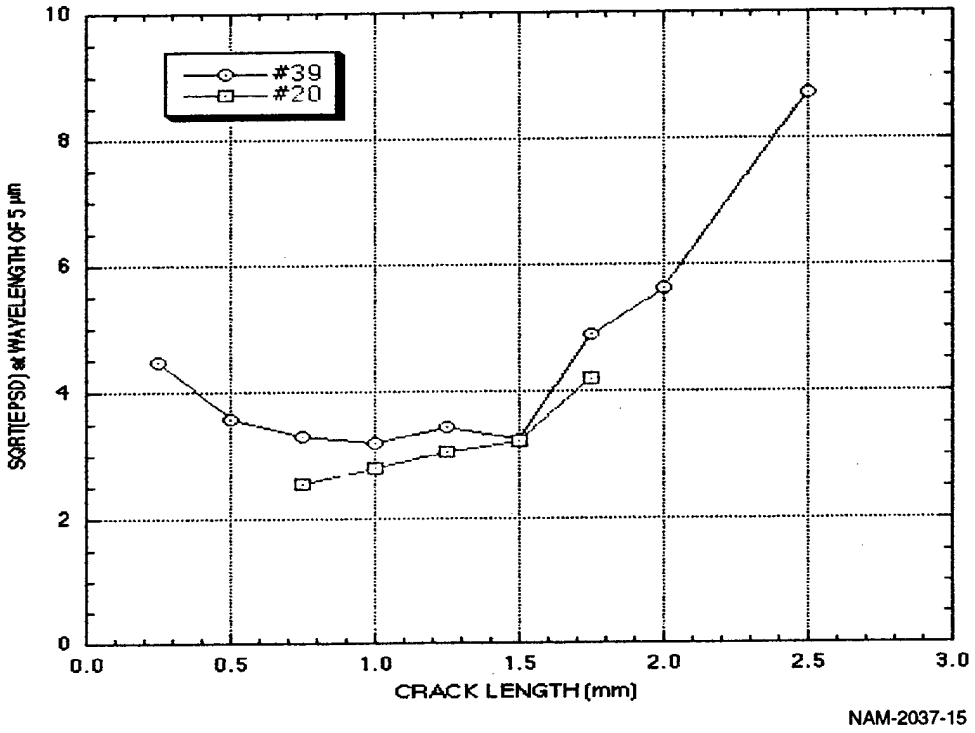


Figure 21. Square root of EPSD values at 5 µm wavelength as a function of crack position for Components 39 and 20.

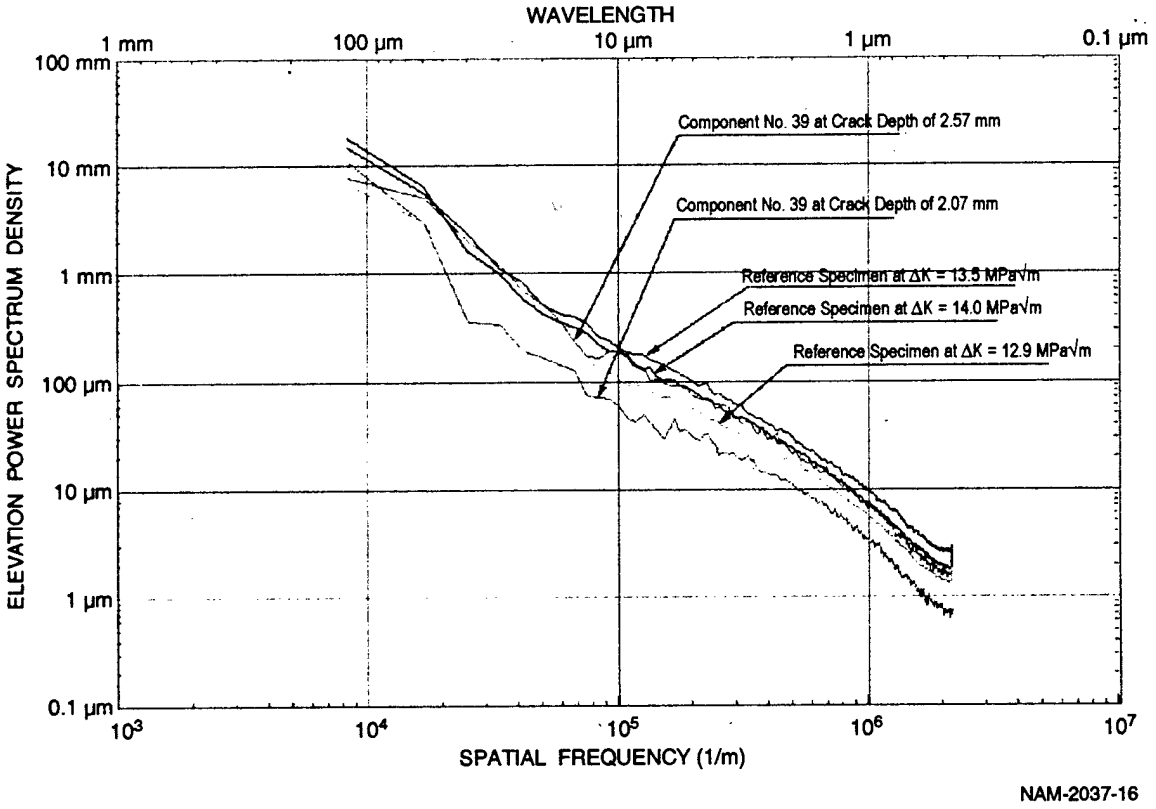


Figure 22. Comparison of EPSD curves of Component 39 with those of a laboratory reference specimen.

laboratory specimen in which the crack was propagated at a stress ratio of 0.8. The shapes of the curves match closely; therefore, chose to correlate the results of Components 39 and 20 with those from the specimen tested under a stress ratio of 0.8. Moreover, we speculate that the components experienced a high R value, perhaps about 0.8, because the EPSD curves from the laboratory specimen tested at R = 0.1 were significantly more convex.

We next attempted to correlate the heights of the curves to obtain ΔK values. When we correlated the ΔK with the square root of the EPSD at a wavelength of 5 μm for the laboratory specimen tested at R = 0.8, we obtained the following relationship

$$(EPSD)^{1/2} = -2.957 + 1.156 \times \Delta K \qquad (R = 0.8)$$

Because the ranges of the EPSD of both components were smaller than those from the laboratory specimen (due possibly to closure effects), we had to extrapolate to determine ΔK information. The converted results, shown in Figure 23, suggest that the cracks in the engine components grew to a length of 1.6 mm under rather constant ΔK conditions of about 5 $MPa\sqrt{m}$ after initiating at a higher ΔK. At this crack depth, the stress intensity range increased strongly and linearly to 10 $MPa\sqrt{m}$ in the next millimeter of crack advance. Such a variation in ΔK with crack depth is consistent with vibratory bending stresses, because bending deflections produce higher stresses near the component surface.

WAVELET ANALYSIS AND FRACTURE PARAMETER CORRELATION

To determine whether wavelet analysis could determine the load conditions that produced the fracture surfaces of Component 39, Professor Naoki Saito and Dr. Jean-Marie Aubry of the Mathematics Department of the University of California at Davis applied the Local Discriminate Basis algorithm and the Linear Discriminant Analysis algorithm to the elevation data. Training data were obtained from the laboratory specimens described above, which had areas of fracture surfaces produced at ΔK levels of 6.6, 8.8, 11.0, 13.2, and 16.5 $MPa\sqrt{m}$. Test data sets from Component 39 consisted of 80 randomly chosen square areas 64 pixels on a side (0.233 μm per pixel). A majority rule was then applied to get a classification for the whole image. When the classification rules obtained from the training data were applied to the test data, all areas of Component 39 were indicated to have been produced at a ΔK of 6.6 $MPa\sqrt{m}$ or lower. This agrees with the linear regression results of Figure 23, except for the results at crack lengths larger than 2.0 mm. Details of the wavelet analysis procedures are given in Appendix A.

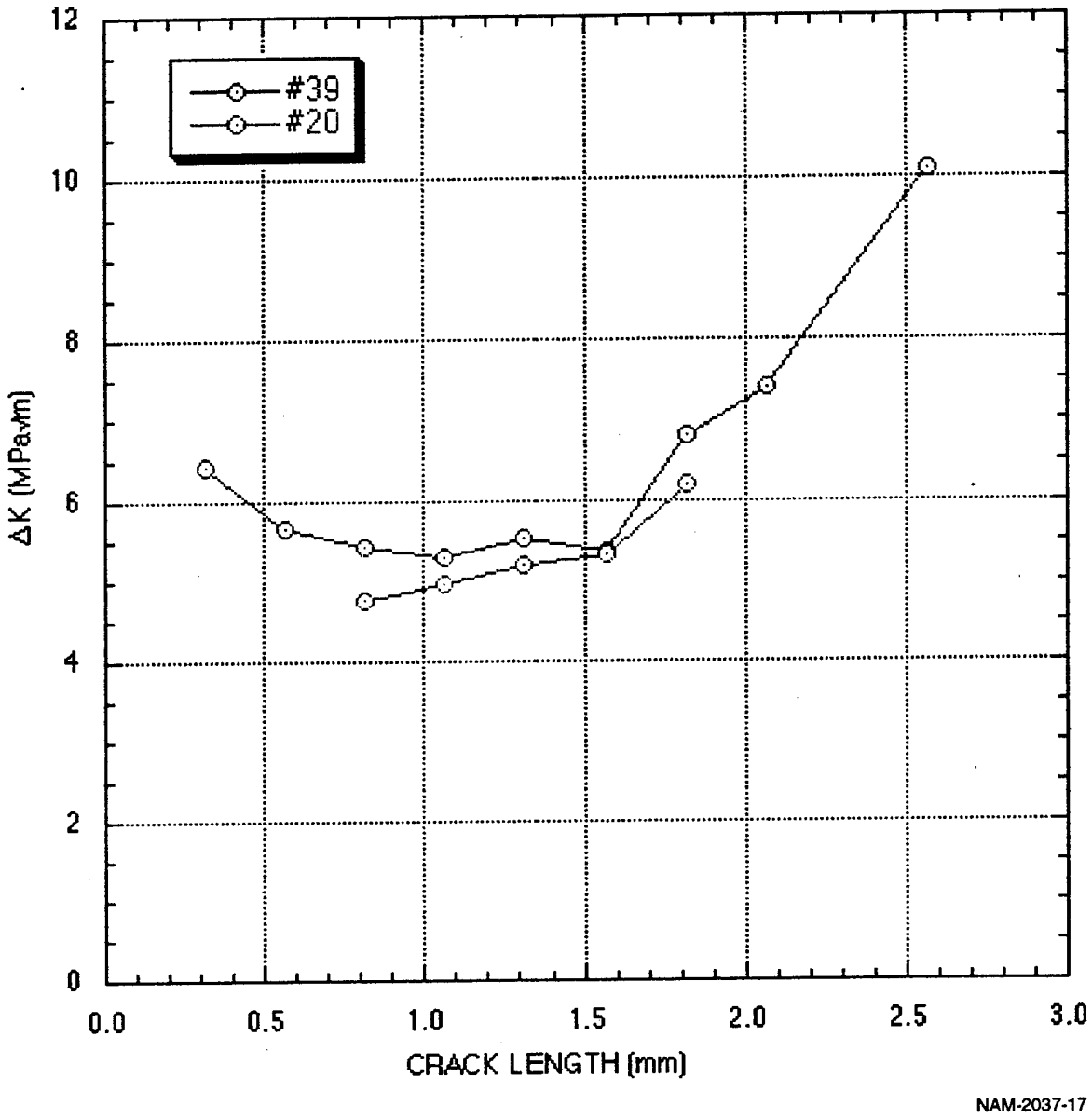


Figure 23. Estimated stress intensity range, ΔK, as a function of crack position in Components 39 and 20.

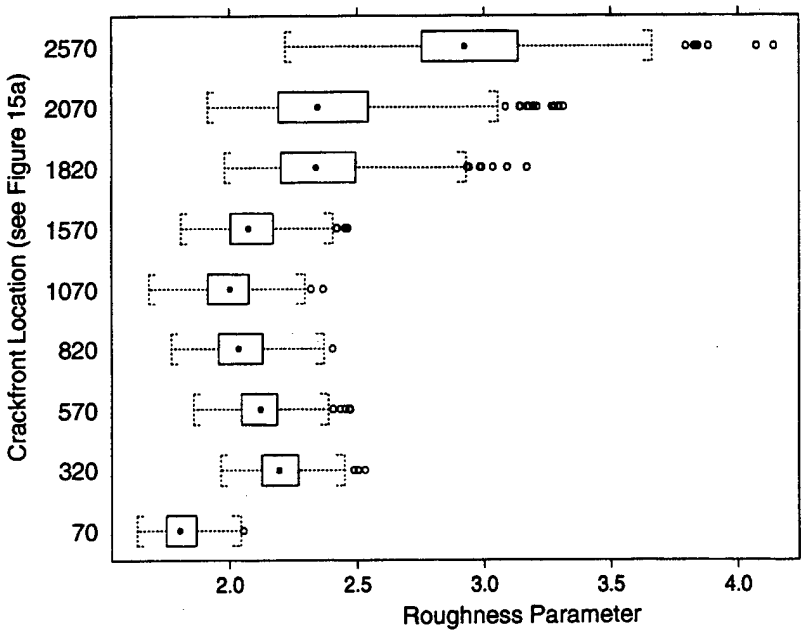
Saito and Aubry also attempted to classify the surfaces by using a roughness parameter defined as follows

$$R_x(y) = \int \sqrt{1 + f_x^2(x,y)}\, dx / \int dx$$

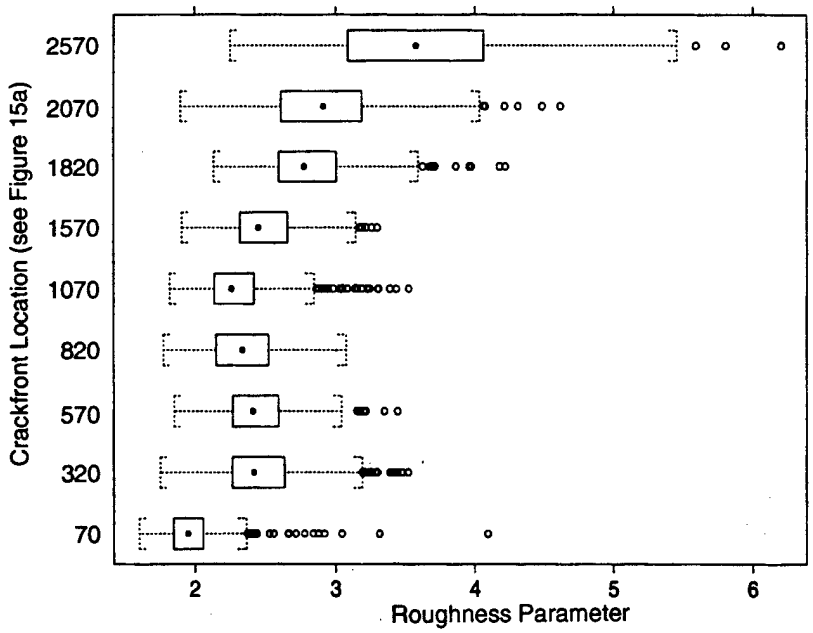
$$R_y(x) = \int \sqrt{1 + f_y^2(x,y)}\, dy / \int dy,$$

where $f_x = \partial z / \partial x$, $f_y = \partial z / \partial y$, and $z = f(x,y)$ is a fractograph representing the elevation data. $\int dx$ represents the length of the specimen in the horizontal direction, which in fact is parallel to the crack propagation for all the available datasets. $\int dy$ represents the length of the specimen in the vertical direction (perpendicular to the crack propagation). One can easily see that, if the profile is perfectly flat, then $R_x = 1$ (or $R_y = 1$).

The variation of this roughness index with crack depth (distance from the crack origin) for Component 39 is shown in Figure 24. Whether the parameter is measured in the direction of crack propagation or parallel to the crack front, the trend is J-shaped, in good agreement with the trend for EPSD or its square root shown in Figure 21. Furthermore, the roughness parameter was correlated with ΔK by applying it to the laboratory specimen and used to evaluate Component 39. Stress intensity range values of 11 $MPa\sqrt{m}$ and 6.6 $MPa\sqrt{m}$ were indicated at locations 2500 and 1000, respectively, in Figure 15(a). These values agree well with those from the Fourier analysis shown in Figure 23.



(a) Component 39 Roughness in Direction Parallel to the Crackfront



(b) Component 39 Roughness in Crack Propagation Direction

Figure 24. Variation of fracture surface roughness parameter with distance from crack origin in Component 39 parallel to the crack front and in crack propagation direction.

DISCUSSION

This seedling study has demonstrated a method for extracting quantitative estimates of load parameters from analysis of topographic features of failure surfaces. This is the first time such a prediction has been made. If the method can be validated and made user-friendly, it will represent a major advance in the state of practice of failure analysis.

The method consists of three parts: (1) characterization and mathematical analysis of fracture surfaces topography, (2) development of a reference database using well-characterized laboratory specimens, and (3) relating the reference data to field failure topography and determining the service load. The Fourier method demonstrated two important points: the shape of EPSD curves appears to depend on the stress-ratio, and the magnitude of the EPSD (or square root of EPSD) relates to DK. Further exploration of EPSD curve shape with wavelet techniques could allow other fatigue parameters such as R-value and Kmax to be determined.

The fast Fourier transform analysis of the data may not be optimal. The Fourier results presented here are highly averaged. Therefore, the EPSD curves are rather smooth and shift up or down within the 0.5 to 15 μm wavelength range. Individual lines of topographic data, however, show significant hash, because of the roughness contribution of the microstructure.

The topographic relief produced by the microstructure is several times greater than that produced by inelastic deformation and hence masks the effects produced by the loads, making the extraction of load parameters difficult. Therefore, we analyzed the averaged EPSD curves to suppress microstructural effects and accentuate the load-induced topography.

Although the EPSD curve showed that features having certain wavelengths exist on the surface, the locations on the fracture surface where these features exist were not indicated. This is a serious limitation of the fast Fourier analysis in interpreting fracture surface topography. Therefore, we applied wavelet analysis, which does not have this shortcoming, to the elevation data from the two aircraft engine components and used the results from the same laboratory specimens to evaluate the results.

To obtain the training data, the local discriminant basis method was applied to 1D and 2D Fourier treatments of the elevation data from the laboratory specimens. Classification rules were established for surfaces produced at DK values of 6.6, 8.8, 11.0, 13.2, and 16.5. When the rules

were applied to the regularly spaced locations on the fracture surface of Component 39, all areas were classified as DK = 6.6. This result suggests that nonstandard wavelet analysis techniques will be required to discriminate among fracture surfaces produced under different DK conditions.

We see three main reasons wavelet analysis did not produce more decisive results: (1) crack closure effects, (2) the size of the area (64 pixel by 64 pixel) chosen for LDB analysis, and (3) noise in the topography data file. We explain how these aspects influence the results below. We confirmed that crack closure significantly influences the fracture surface roughness and learned how the wavelet and FFT results and their relationship with DK are affected. Figure 8 shows that DK of 6.6, 8.8, and 11 are in the transition regions and a relationship between the surface roughness and DK in this region is not linear. Thus, when we analyze the fracture surfaces of these regions, the LDB results will be mixed. However, we believe that wavelet analysis can provide solutions.

Crack closure does influence the topography of fracture surfaces, and we believe that the degree of alteration of the surface is a function of the degree of closure, which in turn is a function of DK and the stress ratio. The crack closure effects result in local changes and thus are best characterized with wavelet techniques. We envision that wavelet can identify those local closure features on the surface that can be used to classify the surfaces in the transition region. One reason for the wide scatter in the LDB results was the small size of the area we chose to investigate relative to the material grain size. With a pixel spacing of 0.233 μm, the area investigated, 64 pixel by 64 pixel, was 15 μm square, which is approximately the area of an average grain (about 10-20 μm). Since the behavior of any individual grain at the macro crack tip will differ from other grains due to differences in orientation, properties, and surrounding grains, the topography of the examined area will not be representative of the material overall. Thus the local microstructure dominated the LDB analysis results and obscured the effects of loading condition. It is necessary to consider special way to assess the average effect of loading.

Another source of error was noise in the topography data, which arose from the inability of the confocal microscope to get unambiguous light reflection from high angle facets in the fracture surface. During scanning, the reflected light intensity was monitored at every pixel point as the distance between the fracture surface and the optical microscope changed. The light intensity of a specific pixel reaches its peak when the surface is in the focal plane of the microscope optical system and the microscope records the value of the peak light intensity and the elevation of the pixel. A problem arises, however, when the reflected light intensity is low, because the microscope cannot determine when the peak is reached at a specific pixel. Thus, the recorded peak

position could be arbitrary and its elevation is different from the surrounding area. This single pixel signal is anomalous and appears as high frequency noise. Since the wavelet results were influenced by this noise, a process is needed to eliminate noise. Paradoxically and fortunately, this is an ideal task for wavelet analysis.

Finally, the load shedding experiments used here are not the best type of experiment for generating topographic data. Rather, experiments in which the stress intensity increases as the crack propagates would likely provide better fracture surfaces for topographic analysis. The magnitude of the loads inferred for the cracked engine component, however, must be considered as approximate, because the data points from the laboratory tests did not overlap appreciably those from the service failure. Therefore, extrapolation was required to estimate the service loads. Moreover, correlation of the laboratory reference data with the data from the service component assumed that the component and the specimen experienced the same stress ratio. Both assumptions need to be examined and validated. However, if linear behavior can be assumed, the DKs at the different crack depths are as shown in Figure 23. When superimposed on the other stresses in the component, the bending stresses should result in a profile similar to that in Figure 23.

The next steps are to develop wavelet and Fourier techniques to identify and extract signals from topography data that relate to load and crack history parameters and to build a database of fracture surface topographs. A further step is to develop computational models for micro-deformation and failure behavior of the microstructure to provide the fundamental underpinnings for linking fracture surface topography features to load parameters. Future work is described in the section entitled Next Steps.

CONCLUSIONS

Failure-analysis-critical load parameter information is encrypted on the fracture surfaces of failed structural components.

Mathematical analysis of fracture surface topography can provide quantitative estimates of the load conditions that caused the failure.

A leapfrog advance in the state of practice of failure analysis is imminent.

A 3D fractography analysis technology would enhance air readiness and avoid hundreds of millions in annual maintenance costs.

NEXT STEPS

Much was learned during this seedling program. The load sensitivity of fracture surface topography was confirmed, a route for extracting load information was demonstrated, the limitations of Fourier analysis were shown, and the requirements for extending wavelet analysis were clarified. We realized the need to examine features at much higher resolution, the need for a microstructural failure model, the need to apply wavelet analysis to larger areas relative to the material grain size, and the need to handle noise in the elevation data.

REQUIRED SIZE SCALES

Both the Fourier and the wavelet analyses of the elevation data from the laboratory specimens discriminated fracture surfaces produced under different load conditions. This confirms previous findings suggesting that fracture surface features are sensitive to the loading conditions that caused the crack. The physical reason must lie in the deformation and microfailure behavior of the microstructure. Referring to Figure 1, the slip band activity and the grain boundary rupture activity are two processes that form the fracture surfaces in this nickel alloy, and hence these processes must be sensitive to the characteristics of the load history.

The confocal optics method used here to measure fracture surface topography has a resolution limit of 0.25 μm and hence is quite capable of capturing fracture surface features on the scale of the grains (about 10 μm in the titanium turbine blade alloy, Ti-6Al-4V). However, the steps on the grain surfaces produced by intersecting slip bands are typically below this resolution.

Thus, a higher resolution characterization instrument is needed. The atomic force microscope is capable of measurements as fine as several angstroms and hence capable of recording roughness features produced by slip bands. A problem then is combining features over a wide size range. Fortunately, much thought has been given to this problem, and several schemes exist for linking disparate size scales (see Reference 18, for example).

ANALYSIS OF ADDITIONAL SAMPLES OF DIFFERENT STRESS RATIOS

In this program, we examined the fracture surfaces of laboratory specimens produced under two stress ratios, R = 0.1 and R = 0.8, and found clear differences in the characteristics of the data. To establish the relationship between R and topographic features, we need to examine

surfaces produced under other stress ratios, say, R = 0.3 and R = 0.5. If the results show a consistent trend, then a unified empirical equation may predict the relationship among the SQRT(ESPSD), DK, and stress ratio. Such an equation can be used to extract the loading conditions from the fracture surface topography and to interpret crack closure phenomena at the crack tip.

ROBUST FEATURE EXTRACTION

From our classification experiments, we realized that it is critically important to extract the features that are

(1) Distinguishing the data belonging to the different classes

(2) Common and invariant within the data belonging to the same class.

It is item (2) that we need to address more for any tough classification problems such as the fractography data. The reason for emphasizing the common features among the same class is both for robustness of the features and for shedding light on the underlying physics; we need to extract the features insensitive to the variations of the data in the same class.

SPIKE REMOVAL BY CONTINUOUS WAVELETS

The fracture surface elevation data obtained with confocal optics microscopy contain undesirable spikes that may influence almost all frequency components in the Fourier analysis. Since the high roughness of the data contains useful information that must be preserved, conventional denoising techniques cannot be used. A continuous wavelet transform of these data and the properties of wavelet coefficients related to local smoothness of the data may be effective in detecting and removing the spikes.

LINKING LOCAL FEATURES TO PHYSICAL FRACTURE MODELS

The specific fracture features responsible for producing the "characteristic fingerprint" of a given load condition in a given microstructure may lead to a much more realistic fracture model. That is, if we understand how the microstructure is dissipating energy at specific load conditions, the fracture mechanicians and materials scientists may be able to generate a model that mimics those microstructural cracking phenomena. This would, in turn, generate a model that links the wavelet/Fourier analysis to physical reality. One thing to consider in this scenario would be the

effect of grain orientation, with respect to the crack propagation direction, on the fracture features produced. This effect could be used if or when we consider crack propagation within specific, nonisotropic grains.

It seems to us that the newer sets of dictionaries of bases, such as local Fourier, brushlets, edgelets, and curvelets, which are capable of detecting and characterizing oriented and striped patterns, may be better tools to extract features and link them to the physical models. This project requires close collaboration between mathematicians, fracture mechanicians, and materials scientists.

The empirical correlation between the stress intensity and the EPSD needs to be understood in terms of the micromechanisms that produce the fracture surface.

Needed at this point is a reference database of topographic maps and EPSD curves from fracture surfaces generated under conditions in which the important load parameters are varied independently and systematically over the entire range of interest for each alloy of interest. Such databases, known as fractography atlases, currently exist in a qualitative form as a collection of SEM micrographs and are used to reach subjective, approximate conclusions regarding the causes of a failure. An atlas of topographs and EPSD curves would allow objective and quantitative conclusions and represent a quantum improvement over current handbooks.

REFERENCES

1. ASM Handbook, Volume 12, *Fractography, 1987*, American Society of Materials International (1987).

2. U.S. Navy Workshop on Effect of Corrosion on Fatigue Resistance of Aircraft Materials, May 28-30, 1997, Hyannis, MA, organized by A. K. Vasudevan, Office of Naval Research, V. S. Agarwala, NAWCAD, and J. T. Cammett, NADEPCP.

3. P. C. Paris and F. Erdogan, "Critical Analysis of Crack Propagation Laws," Trans. Am. Soc. Mech. Engrs. **85D**, 528-534 (1963).

4. R.M.N. Pelloux, "Crack Extension by Alternating Shear," Engng. Fract. Mech. 1, 697-704 (1970).

5. M. Gehron and colleagues, Pratt & Whitney Aircraft, private communication, Program Review Meeting, West Palm Beach, FL, September 10-11, 1998.

6. Y. Murakami, N. Shiraishi, and K. Furukawa, "Estimation of Service Loading from the Width and Height of Fatigue Striations of 2017-T4 Al Alloy," Fatigue Fract. Engng. Mater. Struct. **14**, 897-906 (1991).

7. K. Furukawa, Y. Murakami, and S.-I. Nishida, "A Method for Predicting Service Load from the Width and Height of Striation," J. Soc. Maters. Sci, Japan **45**, 340-345 (March 1996).

8. T. Kobayashi, D. A. Shockey, C. G. Schmidt, R. W. Klopp, "Assessment of Fatigue Load Spectrum from Fracture Surface Topography," Int. J. Fatigue **19**, S237-S244 (1997).

9. C. G. Schmidt, T. Kobayashi, D. A. Shockey, and T. H. Flournoy, "Correlation of Fracture Surface Topography with Fatigue Load History," Proceedings of the International Committee on Aeronautical Fatigue 1997, Edinburgh, Scotland (June 16-20, 1997).

10. D. J. Morrison and J. C. Moosbrugger, "Effects of Grain Size on Cyclic Plasticity and Fatigue Crack Initiation in Nickel," Int. J. Fatigue, **19**, S51-S59 (1997).

11. ASTM Std. E647 – 95a, "Standard Test Method for Measurement of Fatigue Crack Growth Rates," published December 1995.

12. T. Kobayashi and D. A. Shockey, "FRASTA: A New Way to Analyze Fracture Surfaces, Part 1: Reconstructing Crack Histories," Advanced Materials & Processes, **140**, 28-34 (1991).

13. T. Kobayashi and D. A. Shockey, "Fracture Analysis Via FRASTA, Part 2: Determining Fracture Mechanisms and Parameters," Advanced Materials & Processes, **140**, 24-32 (1991).

14. J. C. Newman, Jr., "Analyses of Fatigue Crack Growth and Closure Near Threshold Conditions for Large-Crack Behavior," NASA/TM-1999-209133, Langley Research Center, Hampton, Virginia (April 1999).

15. N. Saito and R. R. Coifman, "Local Discriminant Bases and Their Applications," J. Mathematical Imaging and Vision, **5**, 337-358 (1995).

16. N. Saito and R. R. Coifman, "Improved Local Discriminant Bases Using Empirical Probability Density Estimation," Amer. Statist. Assoc. 1996 Proceedings of the Statistical Computing Section, pp. 312-321.'

17. T. Kobayashi and D. A. Shockey, "Deducing Fatigue Load Parameters from the Fracture Surfaces of a Cracked Aircraft Engine Component," SRI International Final Technical Report to University of Dayton Research Institute (SRI Project 2037, Contract No. RSC97006), Dayton, OH, November 1999.

18. A. M. Gokhale and S. Yang, "Application of Image Processing for Simulation of Mechanical Response of Multi-Length Scale Microstructures of Engineering Alloys," Met. and Mat. Trans. A, Volume 30A, September 1999, pp. 2369-2381.

APPENDIX A

CLASSIFICATION OF FATIGUE FRACTURE SURFACE TOPOGRAPHY USING MULTISCALE AND SPACE-FREQUENCY METHODS

Classification of fatigue fracture surface topography using multiscale and space-frequency methods

Naoki Saito and Jean-Marie Aubry

January 26, 2000

Abstract

We explored the application of local discriminant basis (LDB) methods, which efficiently extracts discriminant local features, to classify fractographs (fracture surface topography data) of titanium specimens according to the loading parameter ΔK, the stress intensity factor, under which the fracture propagated.

Our experiments are mainly grouped into two categories. 1) Application of 1D LDB algorithms along either row or column or circular directions with varying parameters (normalizations, filters, discriminant measures, number of features, space-domain data or frequency-domain data). 2) Application of 2D LDB algorithm to the same datasets viewed as images with varying parameters.

From the results on the above experiments, we found the following. 1) In general, the larger the difference between the ΔK values of these classes, the lower the misclassification rates. This suggests that fracture surface elevation data definitely contains the information sensitive to ΔK. 2) 2D classification on the Fourier domain is more stable than 1D classification, mainly because we can assume more statistical independence on the 2D image patches, and the Fourier magnitudes

are translation invariant. 3) It is still difficult, however, to single out the features only sensitive to the change of ΔK from the elevation data. The local variability of the data is so severe that even the two-class classification problems is tough (e.g., the misclassification rate ranges around 30% to 40%). We need to apply the majority rule on the classification results of many samples for correct classification.

Based on these findings, we recommend the following actions should be taken: 1) Develop *robust* feature extraction and classification algorithms, which are insensitive to the local variations in microstructures 2) Link local features to physical fracture models with the collaboration with the material science experts. 3) Build stochastic models that can simulate the similar topographic data from the available samples. 4) Integrate multiple sensors and resolutions for better and relevant feature extraction for load parameter estimation.

1 Introduction

1.1 High cycle fatigue and structural failure

High cycle fatigue is a cause of failure for materials, especially metals, occurring when a cyclic load is applied at some location. Usually a small initial defect already exists, around which the stress concentrates. When the cyclic load is applied, the stress amount oscillates between values K_{min} and K_{max} and a fatigue fracture propagates through the material.

If this happens to a critical part in an aircraft, such as a turbine blade, and the undetected fractures reach the point of final rupture, consequences can be dramatic. Therefore, a better understanding of this phenomenon is very important. After the failure occurred, one would like to diagnose the causes of the fracture propagation, for instance, the stress intensity factors K_{min} and K_{max} or their difference ΔK. This information is to be used later, for example, to determine during which part of the flight most fracture

propagation occurs, and to improve the fracture-resistance of the material.

1.2 Existing approaches

The method currently used by aircraft engine manufacturers to analyze fractographs is by trained human eyes. These people usually work on SEM images, not elevation data as we do.

First studies at SRI [5, 6, 7, 8, 9, 19] proved some feasibility of using the (mean) power spectrum of adjacent rows of elevation data as a tool to discriminate and estimate the ΔK. Some aspects of our work are an extension of this idea (§ 5.2.3 and 5.3.2).

1.3 Fundamental hypotheses

This study is based on the first fundamental assumption that the history of the fatigue parameters is encrypted on the fracture surface. A second hypothesis is that, for a given material, only the ΔK parameter is relevant and governs the fracture surface topography; all other variations due, for instance, to the fine-scale material structure, are considered as random but constant in a probabilistic sense. In other words, we suppose that the law of the stochastic process creating the fracture depends only on ΔK.

Note that this is a very strong hypothesis. One aspect of this work is to try to validate it.

2 Mathematical formulation

2.1 Probabilistic setting

In all generality, we can pose the problem as follows. Suppose that the material is fixed, and let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space. According to the fundamental hypotheses, there exists, for each value of the parameter ΔK

in a certain range, a random process $\omega \in \Omega \mapsto f_{\Delta K}(\omega)$ whose realizations represent the fracture surface topography of a piece of material that was broken under high-cycle fatigue with the parameter ΔK. Each realization $f_{\Delta K}(\omega)$ is thus a function from $S \to \mathbb{R}$, where $S \subset \mathbb{R}^2$ is the area where the fracture surface is analyzed. For a given location $x \in S$, $f_{\Delta K}(\omega)(x)$ is the actual elevation of this particular sample at coordinate x (note that in practical applications, S is a finite set of discrete points).

The problem we want to solve is now: given a function $f_X : S \to \mathbb{R}$ representing the fracture topography of a sample broken under some unknown conditions (X stands here for experiments), determine or estimate ΔK such that f_X can be viewed as a realization of the process $f_{\Delta K}$.

This is not a good question in our setting, because the answer may not be unique. Rather, we formulate it in a Bayesian sense: considering all possible values of ΔK, what is the probability density function (pdf) $\Delta K \mapsto p_{f_X}(\Delta K)$ that f_X is indeed a realization of $f_{\Delta K}$? Once this is answered, we can chose "the" ΔK to be the point of maximum likelihood in p_{f_X}, or its expectation $\int \Delta K p_{f_X}(\Delta K)$, or whatever suits best to the application.

The calculation of p_{f_X}, however, requires two types of information: a prior probability p_0 on ΔK, and the law of the process $\omega \mapsto f_{\Delta K}(\omega)$. The first one may be obtained from technical considerations: for instance in the case of an airplane turbine blade, we may know what sort of loads to expect in normal service, and thus have a prior probability distribution for ΔK (we can use uniform distribution, if this knowledge is not within the range of aeronautical engineers' predictions). The second one is not available, as we discussed in Introduction.

As is often the case in statistics, we do not have the law of the stochastic process of interest and we have to estimate it from a few sample realizations, which are the laboratory samples fractured under known conditions in our case.

Here we must leave the world of the rigorous Bayesian theory for more

practical matters.

2.2 Classification problem

In our case, the number of available realizations of the process is very small (less than five per ΔK), and moreover we only have them for a limited set of ΔK values (see § 3). Therefore we must (temporarily) give up the estimation problem and concentrate on the simpler classification problem.

Let us give the classical formulation of a signal classification problem. We first define appropriate spaces of input signals (or patterns), extracted features, outputs (or responses), and mapping functions among them. Let $\mathfrak{X} \subset \mathbb{R}^n$ denote a *signal space* (or a *pattern space*) which is a subset of the standard n-dimensional vector space and which contains all signals (or samples/patterns) under consideration. In this case, the *dimensionality* of the signal space or equivalently the length of each signal is n. Let $\mathfrak{Y} = \{1, 2, \ldots, K\}$ be a set of the class or category names to which the input signals belong. We call this space a *response space*.

Signal classification can be considered as a mapping function (usually many-to-one) $d : \mathfrak{X} \to \mathfrak{Y}$ between these two spaces. Direct manipulation of signals in the signal space for classification is prohibitive because: 1) the signal space normally has very high dimensionality (e.g., for a typical segment of fractograph image we use for our experiments, $n = 64 \times 64 = 4096$), and 2) the existence of noise or undesired components (whether random or not) in signals makes classification difficult. On the other hand, the signal space is overly redundant compared to the response space. Therefore, it is extremely important to reduce the dimensionality of the problem, i.e., extract only relevant features for the problem at hand and discard all irrelevant information. If we succeed in doing this, we can greatly improve classification performance both in its accuracy and efficiency.

For this purpose, we set up a *feature space* $\mathfrak{F} \subset \mathbb{R}^k$ where $k \leq n$ between the signal space and the response space. A *feature extractor* is defined as a

map $f : \mathfrak{X} \to \mathfrak{F}$, and a *classifier* (or *predictor*) as a map $g : \mathfrak{F} \to \mathfrak{Y}$. Let $\mathfrak{T} = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^N \subset \mathfrak{X} \times \mathfrak{Y}$ be a *training* (or *learning*) dataset with N pairs of signals $\boldsymbol{x}_i$ and responses (class names) y_i This is the dataset to be used to construct a feature extractor f. Let N_c be the number of signals belonging to class c so that we have $N = N_1 + \cdots + N_K$. Also, let us denote a set of class c signals by $\{\boldsymbol{x}_i^{(c)}\}_{i=1}^{N_c} = \{\boldsymbol{x}_i\}_{i \in I_c}$ where $I_c \subset \{1, \ldots, N\}$ is a set of indices for class c signals in the training dataset with $|I_c| = N_c$.

Preferably, the performance of the whole process should be measured by the misclassification rate using a *test* dataset $\mathfrak{T}' = \{(y_i', \boldsymbol{x}_i')\}_{i=1}^{N'}$ (which has not been used to construct the feature extractors and classifiers) as $(1/N') \sum_{i=1}^{N'} \delta(y_i' - d(\boldsymbol{x}_i'))$, where $\delta(r \neq 0) = 1$ and $\delta(0) = 0$. If we use the *resubstitution* (or *apparent*) error rates (i.e., the misclassification rates computed on the training dataset), we obviously have overly optimistic figures.

For this application, we focus on the feature extractors of the form

$$f = \Theta^{(k)} \circ \Psi,$$

where $\Theta^{(k)} : \mathfrak{X} \to \mathfrak{F}$ represents the selection rule (e.g., picking most important k coordinates from n coordinates), and $\Psi \in \mathrm{O}(n)$, i.e., an n-dimensional orthogonal matrix. In particular, we consider matrices representing the orthonormal bases in the basis library (consisting of wavelet packets or local trigonometric bases) as candidates for Ψ. As a classifier g, we adopt Linear Discriminant Analysis (LDA) of R. A. Fisher [2].

2.3 Feature extraction using LDB

We briefly review the principles of feature extraction using the Local Discriminant Basis algorithm of Coifman and Saito [18, 17].

Let $\mathbb{R}^n$ be the space of the signal to analyze. We assume that n is a dyadic number of the form 2^{J_0}, and we are given a hierarchical system of orthonormal bases for $\mathbb{R}^n$.

Such a system is composed of: first an orthonormal basis B_0^0 which is (usually) simply the standard basis of $\mathbb{R}^n$. Then, $\mathbb{R}^n$ is decomposed into two orthogonal subspaces of dimension $n/2$, the first one having an orthonormal basis B_0^1 and the second one an orthonormal basis B_1^1. The same decomposition is applied to each of the subspaces, recursively up to some maximum decomposition level $j_1 \leq j_0$ (in practice $j_1 = j_0 - 1$ or $j_0 - 2$). At level j, $\mathbb{R}^n$ is thus decomposed as the orthogonal sum of 2^j subspaces of dimension 2^{j_0-j}, each of them with its orthonormal basis B_k^j ($k = 0, \ldots, 2^j - 1$). The number N_{j_1} of such bases for $\mathbb{R}^n$ that can be built by concatenating complementary sub-bases (with maximum decomposition level j_1) is given by the recurrence: $N_j = N_{j-1}^2 + 1$. These bases form a redundant *dictionary* for $\mathbb{R}^n$.

For instance, with $j_1 = 2$, the decomposition is symbolized on Figure 1, and we have five possible bases: $\{B_0^0\}$, $\{B_0^1, B_1^1\}$, $\{B_0^1, B_2^2, B_3^2\}$, $\{B_0^2, B_1^2, B_1^1\}$, $\{B_0^2, B_1^2, B_2^2, B_3^2\}$.

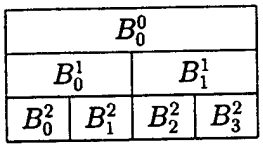


Figure 1: Hierarchy of bases

Such a hierarchical system of orthonormal bases can be obtained using the discrete wavelet transform (DWT) or local trigonometric transform (Local Cosine, Local Sine, Local Fourier). See [11] for more details.

LDB achieves the construction of a feature extractor of the form $f = \Theta^{(k)} \circ \Psi$ in two steps. First, Ψ is selected as the rotation to the "best" basis among the N_{j_1} available. Then $\Theta^{(k)}$ selects the k "best" vectors in this basis. Here "best" is to be understood as most suitable for the discrimination problem at hand.

The choice of the "best" basis and vectors is the central problem of LDB. For this we use some criterion δ, saying that a basis B is better than a basis

B' if $\delta(B) > \delta(B')$.

However, the number N_{j_1} can be huge, and it would be very inefficient to check each possible basis one by one. So we need the criterion δ to be in additive form, i.e. for each vector v we have $\delta(v)$ and $\delta(B) = \sum_{v \in B} \delta(v)$. This allows the selection of the basis to be done hierarchically, by the following procedure. Let B_k^j be some sub-basis in our dictionary, itself being decomposed in two sub-bases B_{2k}^{j+1} and B_{2k+1}^{j+1}. If $\delta(B_k^j) > \delta(B_{2k}^{j+1}) + \delta(B_{2k+1}^{j+1})$, then we keep $\{B_k^j\}$ in the decomposition of $\mathbb{R}^n$; else we replace it by the pair of bases $\{B_{2k}^{j+1}, B_{2k+1}^{j+1}\}$.

Starting from the bottom level $\{B_k^{j_1}, k = 0, \ldots, 2^j - 1\}$, this recursion yields the basis in the dictionary that maximizes the criterion δ. Moreover, once the basis is selected, the same criterion (already computed) can be used to select the k best vectors.

We have a few important remarks. First, imposing additivity is a huge restriction. Actually, the true criterion to use would be to minimize the misclassification rate when the k best vectors in the selected basis are supplied to the classification module g. But this criterion is clearly not additive, and this problem creates combinatorial explosion. The choice of an additive criterion will be discussed in § 2.5. Another way to understand this remark is to remember that in general, "the best k vectors are not the k best".

Secondly, once we resign ourselves to using an additive criterion, it is no longer necessary to chose the discriminating vectors in a basis. We can simply take the k vectors in the whole dictionary that maximize δ, regardless of their linear dependence. However, experimentally, lifting this restriction does not improve the results, but makes the computation slightly more demanding (in memory).

2.4 Linear Discriminant Analysis

Fisher's LDA first tries to do its own feature extraction by a linear map $A^T : \mathfrak{X} \to \mathfrak{F}$ (in this case not necessarily orthogonal matrix). This map

A simultaneously minimizes the scatter of sample vectors (signals) within each class and maximizes the scatter of mean vectors of classes around the total mean vector. To be more precise, let $\boldsymbol{m}_c := (1/N_c)\sum_{i=1}^{N_c} \boldsymbol{x}_i$ be a mean vector of class c signals.[1] Then the total mean vector $\boldsymbol{m}$ can be defined as $\boldsymbol{m} := \sum_{c=1}^{K} \pi_c \boldsymbol{m}_c$, where π_c is the prior probability of class c (which can be set to N_c/N without the knowledge on the true prior probability). The scatter of samples within each class can be measured by the *within-class covariance matrix* $\Sigma_w := \sum_{c=1}^{K} \pi_c \Sigma_c$, where Σ_c is the *sample covariance matrix* of class c: $\Sigma_c := (1/N_c)\sum_{i=1}^{N_c} (\boldsymbol{x}_i^{(c)} - \boldsymbol{m}_c)(\boldsymbol{x}_i^{(c)} - \boldsymbol{m}_c)^T$. The scatter of mean vectors around the total mean can be measured by the *between-class covariance matrix* $\Sigma_b := \sum_{c=1}^{K} \pi_c (\boldsymbol{m}_c - \boldsymbol{m})(\boldsymbol{m}_c - \boldsymbol{m})^T$. Then, LDA maximizes a class separability index $J(A) := \mathrm{tr}[(A^T \Sigma_b A)^{-1}(A^T \Sigma_w A)]$, which measures how much these classes are separated in the feature space.

This requires solving the so-called generalized (or pencil-type) eigenvalue problem $\Sigma_b A = \Sigma_w A \Lambda$, where Λ is a diagonal matrix containing the eigenvalues. Once the map A is obtained (normally $k = K - 1$), then the feature vector $A^T \boldsymbol{x}_i$ is computed for each i, and finally it is assigned to the class which has the mean vector closest to this feature vector in the Euclidean distance in the feature space. This is equivalent to bisecting the feature space $\mathfrak{F}$ by hyperplanes. In this application we regard LDA as a classifier although, as explained, it also includes its own feature extractor A^T. LDA is the optimal strategy if all classes of signals obey multivariate normal distributions with different mean vectors and an equal covariance matrix [3], [12]. In reality, however, it is hard to assume this condition. Moreover, since it relies on solving the eigensystem, LDA can only extract global features (or squeezes all discriminant information into a few $[K-1]$ basis vectors) so that the interpretation of the extracted features becomes difficult, it is sensitive

[1] The sample mean operation $(1/N_c)\sum_{i=1}^{N_c}$ in this subsection can be replaced by expectation E_c for general cases; however, in this paper, we focus our attention on the cases of a finite number of samples, so we stay with the sample mean operations.

to outliers and noise, and it requires $O(n^3)$ calculations.

2.5 Discriminant criterion

Let us first concentrate on the two-class classification problem. Suppose also that the probability laws for each of these classes are known: probability p_0 for class 0 and p_1 for class 1. These distributions are the pdf's for $\mathbb{R}^n$-valued random variables C_0 and C_1, respectively.

Let v be a unit vector in $\mathbb{R}^n$. We consider the projected probabilities $p_{0|v}$ and $p_{1|v}$, that is, the pdf (on $\mathbb{R}$) of $\langle C_0, v\rangle$ and $\langle C_1, v\rangle$, respectively.

If $p_{0|v} = p_{1|v}$ then obviously v is not a discriminating feature: viewed from this direction, C_0 and C_1 look exactly the same. On the contrary, if $p_{0|v}$ is very different from $p_{1|v}$, then v is a good discriminating feature. Our criterion δ should measure this difference.

In information theory, the difference between two pdf's can be measured by the relative entropy, or Kullback-Leibler Information, defined by:

$$J(p_{0|v}, p_{1|v}) := \int p_{0|v} \log\left(\frac{p_{0|v}}{p_{1|v}}\right).$$

This formula is not symmetric, however, so we may prefer its symmetrized version:

$$S(p_{0|v}, p_{1|v}) := J(p_{0|v}, p_{1|v}) + J(p_{1|v}, p_{0|v}).$$

The symmetric relative entropy can easily be generalized to any number of classes, taking the sum of (asymmetric) relative entropies for all pairs of classes.

Certainly, this is the most straightforward choice for $\delta(v)$. However, in practice, we stumble on the same ubiquitous issue: we still do not know the probabilities p_0 and p_1. What we have are only a finite number of realizations of each class, constituting the training set.

Two workarounds are available. First, one can try to evaluate the probabilities $p_{0|v}$ and $p_{1|v}$. Since these are one-dimensional distributions, it is

possible to approximate them by the histogram of the values in the training set, or its smoothed version by convolution with some kernel or averaged with shifted versions of itself. Averaged Shifted Histograms (ASH) are, by the way, a particular case of kernel convolution.

Another option is to give up the pdf-based criteria, and use the sum of energies instead, simply taking

$$J_E(0|v, 1|v) := \sum_{c_0,c_1} |\langle c_0, v\rangle|^2 \log\left(\frac{|\langle c_0, v\rangle|^2}{|\langle c_1, v\rangle|^2}\right)$$

where c_0 and c_1 are taken in the training set for class 0 and 1, respectively.

And its symmetrized version:

$$S_E(0|v, 1|v) := J_E(0|v, 1|v) + J_E(1|v, 0|v)$$

with the straightforward generalization to any number of classes.

As before, this measures the difference between the projection of each class on the feature v.

2.6 Modified Relative Entropy

At some point in this study we tried to improve the criterion used. In our case, our main problem is that we only have a very limited set of training data. We cannot avoid the fact that they all come from the same laboratory sample, for instance. As for the selected features, we want them to be truly discriminating between elements of different ΔK classes for any given elevation data, as long as the specimen is the same metal (e.g. Titanium). In some sense, we would like to reject features that are apparently too specific to the actual piece of metal used for the training.

Let H denotes the classical (Shannon) entropy of a probability distribution function. We recall that $H(q) := -\int q \log(q)$ is small when q is a well concentrated distribution. Suppose that $S(p_{0|v}, p_{1|v})$ is large (i.e., very discriminant). Then there are several cases to which we need to pay attention.

(1) If both $H(p_{0|v})$ and $H(p_{1|v})$ are small (i.e., well concentrated), then this vector is an important feature extractor for both classes and it is very unlikely that "chance" (i.e., something too specific in this training data) would make this vector appear important for both classes. So we should keep this vector.

(2) If, for instance, $H(p_{0|v})$ is small and $H(p_{1|v})$ is large, then this is precisely what we want to avoid, because it indicates that this vector is good only for this class 1 training samples and may not carry the generalization ability for the classification problems. LDB might retain such vectors.

(3) If both $H(p_{0|v})$ and $H(p_{1|v})$ are large, then this vector is specific to neither class 0 nor class 1. However, there is no reason to discard this vector since it still carries good discriminant information. So, we keep this one too.

From the above consideration, we propose a Modified Relative Entropy based on pdf's:

$$M(p_{0|v}, p_{1|v}) = S(p_{0|v}, p_{1|v}) - A \max(H(p_{0|v}), H(p_{1|v})) + B \min(H(p_{0|v}), H(p_{1|v}))$$

where A and B are positive constants balancing the relative importance one may want to impose on effects. We then search the basis that maximizes the sum of M (or its obvious modification in the case of more than two classes) from the dictionaries of bases.

3 Data sets

The following elevation matrices were provided by SRI. We mention the material (a Titanium alloy), a reference identification, the ΔK (when known), the dimension in pixels, the horizontal and vertical scale of the originals (matrices used for classification were rescaled at the same scale).

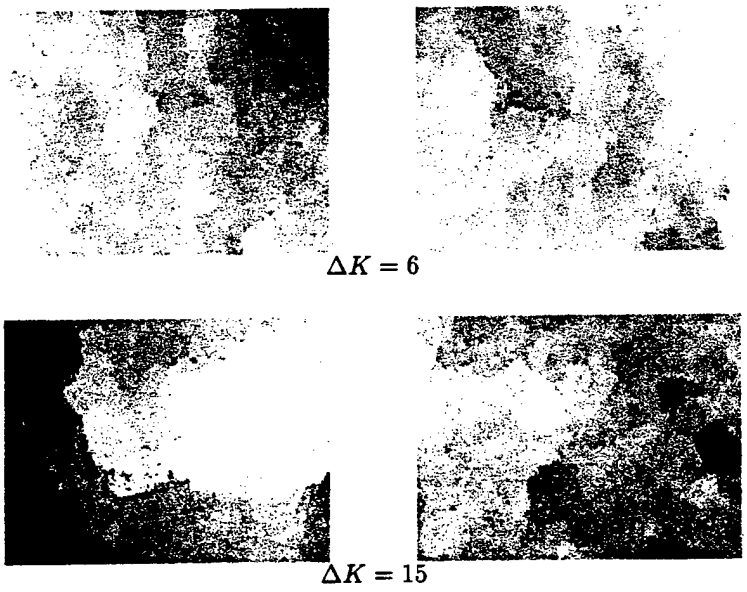


Figure 2: Example of elevation data

3.1 Titanium sample #8354

This is a laboratory sample that was submitted to a stress test with constant ratio $R = K_{\text{max}}/K_{\text{min}} = 0.8$. Parameter $\Delta K = K_{\text{max}} - K_{\text{min}}$ varies from 2.2 to 14.6 (units: ksi$\sqrt{\text{in.}}$).

The stress test is composed of a series of *runs*. The first run starts with a value of ΔK high enough for the fracture, initiated by an initial dent, to propagate. The value of ΔK is then gradually reduced, until it is under the threshold below which the fracture does not propagate anymore. Another run is then started. The last run receives increasing values of ΔK until the final failure.

3.1.1 First data

This batch of data is the first dataset we received. The complete runs of 8353 and 8354 came later.

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8354/Ti6_0	6	718 × 478	0.36	78.25/256
8354/Ti6_1	6	479 × 638	0.29	62.66/256
8354/Ti6_2	6	479 × 638	0.29	62.66/256
8354/Ti6_3	6	479 × 638	0.29	62.66/256
8354/Ti8_0	8	718 × 478	0.36	78.25/256
8354/Ti8_1	8	479 × 638	0.29	62.66/256
8354/Ti8_2	8	479 × 638	0.29	62.66/256
8354/Ti8_3	8	479 × 638	0.29	62.66/256
8354/Ti10_0	10.1	718 × 478	0.36	78.25/256
8354/Ti10_1	10.1	479 × 638	0.29	80.42/256
8354/Ti10_2	10.1	479 × 638	0.29	80.42/256
8354/Ti10_3	10.1	479 × 638	0.29	80.42/256
8354/Ti12_0	12	718 × 478	0.36	116.59/256
8354/Ti12_1	12	479 × 638	0.29	88.11/256
8354/Ti12_2	12	479 × 638	0.29	105.80/256
8354/Ti12_3	12	479 × 638	0.29	137.08/256
8354/Ti15_0	15	718 × 478	0.36	142.37/256
8354/Ti15_1	15	479 × 638	0.29	128.60/256
8354/Ti15_2	15	479 × 638	0.29	84.09/256
8354/Ti15_3	15	479 × 638	0.29	128.60/256

3.1.2 Run 1

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8354/Ti10927	4.52	600 × 2000	0.23	48.01/256
8354/Ti11554	4.01	600 × 2000	0.23	53.95/256
8354/Ti12266	3.49	600 × 2000	0.23	58.42/256
8354/Ti13063	2.99	600 × 2000	0.23	58.42/256
8354/Ti14008	2.50	600 × 2000	0.23	58.42/256
8354/Ti14661	2.20	600 × 2000	0.23	58.42/256

3.1.3 Run 2

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8354/Ti15400	4.05	600 × 2000	0.23	65.07/256
8354/Ti15705	3.49	600 × 2000	0.23	65.07/256
8354/T16017i	3.01	600 × 2000	0.23	65.07/256
8354/Ti16391	2.53	600 × 2000	0.23	72.39/256

3.1.4 Run 3

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8354/Ti17140	3.97	600 × 2000	0.23	72.39/256
8354/Ti17315	3.53	600 × 2000	0.23	72.39/256
8354/Ti17562	3.01	600 × 2000	0.23	72.39/256
8354/Ti17866	2.47	600 × 2000	0.23	72.39/256

3.1.5 Run 5

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8354/Ti18519	3.45	600 × 2000	0.23	72.39/256
8354/Ti18908	4.05	600 × 2000	0.23	72.39/256
8354/Ti19195	4.57	600 × 2000	0.23	72.39/256
8354/Ti19393	4.95	600 × 2000	0.23	72.39/256
8354/Ti19682	5.58	600 × 2000	0.23	72.39/256
8354/Ti19870	6.07	600 × 2000	0.23	72.39/256
8354/Ti20076	6.58	600 × 2000	0.23	72.39/256

3.1.6 Run 7

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8354/Ti21580	4.47	600 × 2000	0.23	72.39/256
8354/Ti22075	4.03	600 × 2000	0.23	72.39/256
8354/Ti22748	3.51	600 × 2000	0.23	72.39/256
8354/Ti23406	3.05	600 × 2000	0.23	72.39/256
8354/Ti24323	2.52	600 × 2000	0.23	72.39/256
8354/Ti24930	2.22	600 × 2000	0.23	72.39/256

3.1.7 Run 8

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8354/Ti25756	4.98	600 × 2000	0.23	72.39/256
8354/Ti26299	5.52	600 × 2000	0.23	72.39/256
8354/Ti26794	6.02	600 × 2000	0.23	72.39/256
8354/Ti27181	6.47	600 × 2000	0.23	72.39/256
8354/Ti27645	7.05	600 × 2000	0.23	84.00/256
8354/Ti27950	7.46	600 × 2000	0.23	96.38/256
8354/Ti28329	8.00	600 × 2000	0.23	96.38/256
8354/Ti28628	8.45	600 × 2000	0.23	109.07/256
8354/Ti29002	9.06	600 × 2000	0.23	109.07/256
8354/Ti29294	9.56	600 × 2000	0.23	109.07/256
8354/Ti29583	10.10	600 × 2000	0.23	112.71/256
8354/Ti29804	10.50	600 × 2000	0.23	112.71/256
8354/Ti30099	11.09	600 × 2000	0.23	112.71/256
8354/Ti30307	11.54	600 × 2000	0.23	112.71/256
8354/Ti30526	11.98	600 × 2000	0.23	139.67/256
8354/Ti30736	12.45	600 × 2000	0.23	139.67/256
8354/Ti30963	12.93	600 × 2000	0.23	139.67/256
8354/Ti31166	13.58	600 × 2000	0.23	139.67/256
8354/Ti31440	13.99	600 × 2000	0.23	151.97/256
8354/Ti31722	14.60	600 × 2000	0.23	162.30/256

3.2 Titanium sample #8353

This is a laboratory sample that was submitted to a stress test with constant ratio $R = K_{\text{max}}/K_{\text{min}} = 0.1$. Parameter $\Delta K = K_{\text{max}} - K_{\text{min}}$ varies from 4.15 to 44.68.

3.2.1 Run 1

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8353/Ti09924	8.52	600 × 2000	0.23	74.93/256
8353/Ti10320	7.99	600 × 2000	0.23	74.93/256
8353/Ti10739	7.45	600 × 2000	0.23	93.60/256
8353/Ti11148	6.96	600 × 2000	0.23	93.60/256
8353/Ti11557	6.53	600 × 2000	0.23	120.22/256
8353/Ti12128	5.96	600 × 2000	0.23	120.22/256
8353/Ti12670	5.46	600 × 2000	0.23	120.22/256
8353/Ti13236	5.00	600 × 2000	0.23	120.22/256
8353/Ti13927	4.48	600 × 2000	0.23	120.22/256

3.2.2 Run 2

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8353/Ti15222	6.94	600 × 2000	0.23	120.22/256
8353/Ti15448	6.14	600 × 2000	0.23	120.22/256
8353/Ti15659	5.46	600 × 2000	0.23	120.22/256
8353/Ti15878	4.86	600 × 2000	0.23	120.22/256
8353/Ti16050	4.47	600 × 2000	0.23	120.22/256
8353/Ti16165	4.15	600 × 2000	0.23	120.22/256

3.2.1 Run 1

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8353/Ti09924	8.52	600 × 2000	0.23	74.93/256
8353/Ti10320	7.99	600 × 2000	0.23	74.93/256
8353/Ti10739	7.45	600 × 2000	0.23	93.60/256
8353/Ti11148	6.96	600 × 2000	0.23	93.60/256
8353/Ti11557	6.53	600 × 2000	0.23	120.22/256
8353/Ti12128	5.96	600 × 2000	0.23	120.22/256
8353/Ti12670	5.46	600 × 2000	0.23	120.22/256
8353/Ti13236	5.00	600 × 2000	0.23	120.22/256
8353/Ti13927	4.48	600 × 2000	0.23	120.22/256

3.2.2 Run 2

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8353/Ti15222	6.94	600 × 2000	0.23	120.22/256
8353/Ti15448	6.14	600 × 2000	0.23	120.22/256
8353/Ti15659	5.46	600 × 2000	0.23	120.22/256
8353/Ti15878	4.86	600 × 2000	0.23	120.22/256
8353/Ti16050	4.47	600 × 2000	0.23	120.22/256
8353/Ti16165	4.15	600 × 2000	0.23	120.22/256

3.2.3 Run 3

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8353/Ti16685	6.50	600 × 2000	0.23	120.22/256
8353/Ti16800	5.98	600 × 2000	0.23	120.22/256
8353/Ti16919	5.53	600 × 2000	0.23	120.22/256
8353/Ti17064	5.01	600 × 2000	0.23	120.22/256
8353/Ti17203	4.54	600 × 2000	0.23	120.22/256

3.2.4 Run 4

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8353/Ti17633	6.53	600 × 2000	0.23	120.22/256
8353/Ti17696	6.01	600 × 2000	0.23	120.22/256
8353/Ti17770	5.50	600 × 2000	0.23	120.22/256
8353/Ti17849	5.01	600 × 2000	0.23	120.22/256
8353/Ti17932	4.51	600 × 2000	0.23	120.22/256

3.2.5 Run 6

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8353/Ti19228	7.46	600 × 2000	0.23	120.22/256
8353/Ti19489	7.03	600 × 2000	0.23	120.22/256
8353/Ti19789	6.54	600 × 2000	0.23	120.22/256
8353/Ti20147	6.01	600 × 2000	0.23	120.22/256
8353/Ti20508	5.53	600 × 2000	0.23	120.22/256
8353/Ti20858	5.08	600 × 2000	0.23	120.22/256
8353/Ti21364	4.51	600 × 2000	0.23	120.22/256

3.2.6 Run 7

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8353/Ti22545	7.50	600 × 2000	0.23	120.22/256
8353/Ti22784	6.96	600 × 2000	0.23	120.22/256
8353/Ti23256	5.94	600 × 2000	0.23	120.22/256
8353/Ti23487	5.53	600 × 2000	0.23	120.22/256
8353/Ti23815	4.97	600 × 2000	0.23	120.22/256

3.2.7 Run 8

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
8353/Ti24326	9.11	600 × 2000	0.23	120.22/256
8353/Ti25065	9.52	600 × 2000	0.23	120.22/256
8353/Ti25880	10.00	600 × 2000	0.23	120.22/256
8353/Ti26695	10.53	600 × 2000	0.23	120.22/256
8353/Ti27404	11.02	600 × 2000	0.23	120.22/256
8353/Ti28026	11.49	600 × 2000	0.23	120.22/256
8353/Ti28628	11.97	600 × 2000	0.23	120.22/256
8353/Ti29246	12.51	600 × 2000	0.23	120.22/256
8353/Ti29746	12.97	600 × 2000	0.23	120.22/256
8353/Ti30259	13.49	600 × 2000	0.23	120.22/256
8353/Ti30770	14.03	600 × 2000	0.23	120.22/256
8353/Ti31173	14.48	600 × 2000	0.23	120.22/256
8353/Ti31684	15.10	600 × 2000	0.23	120.22/256
8353/Ti31989	15.49	600 × 2000	0.23	120.22/256
8353/Ti32390	16.03	600 × 2000	0.23	120.22/256

3.3 Component #39 (titanium)

Component #39 is one of the aircraft engine components that actually failed in service. The ΔK parameter is unknown in this case. Fractographs were taken at nine regularly spaced locations, starting from the edge of the component where the crack initiated.

Name	ΔK	Dimension	xy scale (μm/pixel)	z scale (μm/gray level)
B39/TiEDG	?	719 × 478	0.23	29.32/256
B39/Ti0250	?	719 × 478	0.23	29.32/256
B39/Ti0500	?	719 × 478	0.23	29.32/256
B39/Ti0750	?	719 × 478	0.23	29.32/256
B39/Ti1000	?	719 × 478	0.23	29.32/256
B39/Ti1500	?	719 × 478	0.23	29.32/256
B39/Ti1750	?	719 × 478	0.23	29.32/256
B39/Ti2000	?	719 × 478	0.23	29.32/256
B39/Ti2500	?	719 × 478	0.23	29.32/256

4 Data preparation

As one can see, it is very difficult to conduct statistical analysis on these datasets, for we have a very small number of samples (maximum five images in each ΔK class) of high dimension (up to 600 × 2000). It would make no sense to try to extract features in this space. Instead, we simulate a large number of samples in a smaller space by sub-sampling the images. This artificially creates a nice statistical set, but one must keep in mind that it is basically biased by the fact that the samples are not at all independent: they come from the same piece of metal, sometimes they even overlap. One must be very careful on this point, and try to correct the bias as much as possible.

Also, one must check that the horizontal and vertical scales of the files we want to compare are the same, and rescale them if needed.

4.1 One-dimensional sampling

We can sample rows or columns of the matrix (the "row" direction corresponding to the direction of propagation of the fracture), or along other shapes such as circles. Optionally, we could take their Fourier transform and study them in the frequency domain.

One-dimensional computation is easier and faster, and it was the first way we tried to validate the method. However, we should keep a critical eye on the results, because the dependence between the samples in this case is very strong (by continuity, two neighboring rows have about the same profile).

Moreover, at the very beginning of this study, we were given only five images, each one in its ΔK class (8354/Ti6_0, 8354/Ti8_0, 8354/Ti10_0, 8354/Ti12_0, 8354/Ti15_0). So we had to split them to make the training and test data sets: first half of the matrix in the training set, the second one in the test set, eventually with some randomization. In this case, the test set is very dependent from the training set, and the significance of the results may be very weak. For instance, the classification could be (and actually was) based on the mean elevation or slope of the sample, which depend only on the data acquisition procedure and not on the ΔK. This is not what we want: we want to select features that are pertinent only to the parameter ΔK. We tried to correct this baleful effect by normalizing the mean and the first order moment to zero, so that each sample has the same overall elevation and slope.

Later, other locations (but still from the same piece of metal) for each class were added to our collection (8354/TiΔK_n, $n = 1, 2, 3$). Since, in each class, locations 1 and 2 do not overlap, we could use one of them as training set, the other one as test set, for a more realistic trial.

4.2 Two-dimensional sampling

We also prepared and examined two-dimensional samples. We took small square "patches" of size 32×32 or 64×64 (powers of two are more suitable to multiresolution analysis). Usually they were taken at random, but sometimes with a fixed pattern such as checker-board tiling. Optionally, we could take their Fourier transform and study them in the frequency domain. A typical number of patches is 128 on each image.

In the case of two-dimensional sampling, the problem of the dependence between neighboring samples seems less present. We do not have the means to give a precise formulation of this statement, because this requires a probabilistic model of the surface, which we do not have, but a coarse justification is given by the remark that, when using a regular tiling, a square of side c only has $4c$ points that are close to points of its neighbors, while its total number of points c^2.

4.3 Computational aspects

The computations were done on an Intel Pentium II machine (450 MHz CPU, 128 MB RAM) running Linux as an operating system.

The data preparation and feature extraction routines were implemented in C++. The files were converted and manipulated as Scientific Data Sets in Hierarchical Data Format (a file format developed by the National Center for Supercomputing Applications, freely available at http://hdf.ncsa.uiuc.edu).

For one-dimensional analysis, the Local Discriminant Basis code (`ldb`) and its version using Averaged Shifted Histograms (`ldbkash`) had already been developed by one of us (Saito).

For two-dimensional analysis, the program to compute the energy distribution on the basis dictionary (`tfenergy2d`), and the program using the output of the latter to compute the LDB (`ldb2d`) had also already been written. The code computing the Averaged Shifted Histograms of the coef-

ficients in the basis dictionary (`tfkash2d`), as well as the code using these histograms to compute the probability distributions-based LDB (`ldbkash2d`) were developed as a part of this project.

The actual classification using LDA was done in S-plus using a code from Venables and Ripley [20].

5 Overview of the results

5.1 Note on the interpretation

The results are given in the form of misclassification rates, that is, the percentage of total errors made by the classifier over the total number of samples. The classifier was first tested on the training data (resubstitution error rate), then on the test data. We often give both in the following tables, but as explained in Introduction, the second number is the one that has more significance.

For two-class classification problems, the worst misclassification rate should be 50%, in which case the classifier operates "at random", and is of no use. Misclassification rates $> 50\%$, if they occur, indicate a potential flaw in our fundamental hypotheses.

Misclassification rates $< 50\%$ are positive results. Though an error rate of, say, 35%, may not seem very encouraging, we should remember that we are trying to classify the whole images, not only the small samples that were created for doing the statistics.

So we can apply a rule deciding that the class of the whole image is the class of the majority of the samples. Assuming that we have N independent sample on the image, and that the probability of making an error on one sample is p, the probability p_M of having a majority of errors is

$$p_M = \sum_{k=0}^{\left\lceil \frac{N+1}{2} \right\rceil} \binom{N}{k} (1-p)^k p^{N-k}.$$

Using the Central Limit Theorem, we can approximate this probability for large N by

$$p_M \simeq \mathrm{erfc}\left(\frac{1/2-p}{\sqrt{p-p^2}}\sqrt{N}\right)$$

where erfc is the normal tail error function defined by

$$\mathrm{erfc}(x) := \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt.$$

As long as $p < 1/2$, the error rate for the whole image decreases exponentially with N.

5.2 One-dimensional classification

5.2.1 Data preparation: one file per class

As we saw it before, constructing a statistical classifier requires a large number of training samples. In our case, the number of available images (elevation matrices) is very limited; actually, when we started this study, we only had one image for each of five values of ΔK: 8354/Ti6_0, 8354/Ti8_0, 8354/Ti10_0, 8354/Ti12_0, and 8354/Ti15_0.

Since a preliminary study had shown some results using one-dimensional Fourier transform on the rows of these images, we began our exploration in the same direction. For the simplest problem of two-class classification (for instance between 8354/Ti6_0 and 8354/Ti15_0), we would take half of the rows of each image for training, and the second half for testing the classifier.

Other options include : taking columns, randomized rows or columns, random circles... We also tried to take off the mean and/or the first order moment of each sample, in order to suppress the artifact caused by the overall elevation of the samples. Here is an example of misclassification rates between 8354/Ti6_0 and 8354/Ti15_0, obtained using ldb with Daubechies 10 filter to extract the 50 most discriminant (in the sense of symmetric relative entropy) features.

The first number represents the misclassification rate when the classifier is applied to the training data itself; the second number is for the test data.

	raw	mean 0	first moment 0	both 0
rows	0% 30%	0% 1%	0% 31%	1.5% 23%
columns	0% 0%	0% 0%	0% 1.8%	0% 1.1%
random rows	0% 0%	0% 0%	0.42% 0.84%	2.1% 3.3%
random col.	0% 0%	0% 0%	0.14% 0%	0.42% 0.7%
random circles	3.1% 6.1%	0.63% 4.2%	12% 14%	35% 44%

Unfortunately, these figures are not statistically significant, for the simple reason that training and test samples are not statistically independent: they come from the same image! This bias is even worse when they are randomized. Moreover, we see that when the mean and the first order moment are removed, the performance drops. This suggests that classification is actually done on the overall elevation and slope of each image, rather than on its intrinsic surface characteristics.

Nevertheless, this kind of very primitive experiment can yield some interesting conclusion. We explored further the two-class one-dimensional classification on all possible pairs of the five classes we had at that point. Here are the results obtained using `ldb` with Daubechies 10 filter, 50 most discriminant features, random columns, mean and first order moment set to zero. Results above the diagonal use the first half of randomized columns for training, second half for testing; results below the diagonal use the second half of randomized columns for training, first half for testing.

	$\Delta K = 6$	$\Delta K = 8$	$\Delta K = 10$	$\Delta K = 12$	$\Delta K = 15$
$\Delta K = 6$		8.8% 13%	3.1% 6.7%	7.4% 8.5%	0.42% 0.7%
$\Delta K = 8$	6.8% 11%		12% 17%	2.8% 5.7%	0.7% 0.42%
$\Delta K = 10$	4.3% 5.8%	8.8% 12%		4.2% 6.4%	3.2% 3.8%
$\Delta K = 12$	7.2% 7.2%	4% 4.5%	4.5% 6.3%		0.14% 0.42%
$\Delta K = 15$	0.14% 0.84%	0.56% 0.97%	1.8% 2.5%	0% 0.56%	

As previously noticed, because of the statistical dependence, these results are completely biased (in regard of what would be the performance of the classifier in a real situation). But a general tendency can be observed: *the more different the ΔKs, the lower the misclassification rate.* This is an important confirmation of our fundamental hypothesis.

5.2.2 Two files per class

A more realistic test of our methods is of course the case when the samples used to test the classifier are independent from the ones used to train it. In order to keep all other parameters (the physical conditions of the experiments) as neutral as possible for the classification problem, training and test samples were taken from the same piece of metal at different locations (ΔK being the same within each class). These samples are noted 8354/TiΔK_n, with ΔK=6, 8, 10, 12, 15 and n=1, 2, 3 corresponding to three different locations. Locations 1 and 2 do not overlap at all, and thus we can use them for the test where training and test samples are independent.

Here we show the results for Ti6/Ti15 classification, with various sampling schemes, using Daubechies 10 wavelets for ldb or ldbkash and extracting 50 basis vectors. M1 means that the mean and the first moment are set to zero. Here the first number is the misclassification rate obtained with files #1 as training set, files #2 as test set; the second number is obtained with files #2 as training set, files #1 as test set.

	ldb	M1 + ldb	ldbkash
rows	100% 50%	30% 50%	100% 50%
columns	9.5% 41%	8.4% 40%	19% 37%
rows, randomly shifted	8.1% 50%	48% 47%	7.7% 42%
columns, randomly shifted	23% 43%	49% 46%	24% 42%
random circles	75% 50%	47% 45%	96% 50%

Bad results and extreme variation suggests that ldb(kash) is much too sensitive to the particular feature of the samples. The asymmetry between each

pair of results is quite surprising : in almost all cases the pair Ti6.2/Ti15.2 seems better for training than Ti6.1/Ti15.1. This could result from some special property of these files, but this is obviously not what we are looking for.

Figure 3 shows the 10 most discriminant LDB vectors computed for the Ti6.1/Ti15.1 classification problem. Some of the vectors suggests that the discriminant information may be in the Fourier (i.e., wavenumber or spatial frequency) domain. This observation lead us to the next set of experiments.

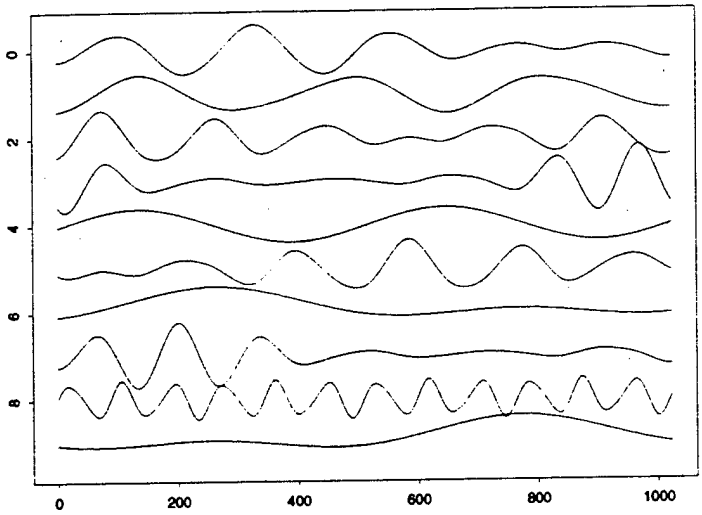


Figure 3: Example of top 10 seleted basis vectors

5.2.3 Fourier domain

We then conducted the classification experiments in the Fourier domain, and use a basis of Local Cosines to analyze it. This is a refinement of the original idea of looking at the power spectrum of each sample: the effect of using the

LCT-based LDB is to (automatically) find a discriminating portion of the Fourier domain.

We FFT-ed each row, column and circle, and computed the energy distributions in the Fourier domain. These data were supplied to the LDB algorithm. Note that when no filter was employed, which means that LDA classification was directly applied on the whole energy distributions, the computation time severely increased because we used all (or half) of the 512 coordinates.

	d10+s	lct	lda all	lda 2nd half
rows	24% 41%	18% 35%	20% 45%	51% 51%
columns	14% 39%	10% 41%	5.9% 36%	0% 50%
random circles	16% 50%	67% 50%	65% 48%	0% 50%

The last column shows the result obtained when we tried the classification based on the second half of the Fourier domain (high frequencies). This attempt was suggested by other evidence from SRI that some useful information may exist in the highest frequency region.

However, the results are still not very conclusive (slightly better than in the spatial domain, but not much).

5.2.4 Summary of results for one-dimensional classification

- In general, the more different the stress intensity factors (ΔK), the lower the misclassification rates.
- The misclassification rate varies with the number of LDB features used (violating independence of each sample).
- If both the training and test datasets are selected from the same region randomly, then the misclassification rates are very small (0.4 % to 17 %).

- If the training and test datasets are selected from the two different regions without randomization, then the misclassification rates are much higher (8.4 % to 41 %).

5.3 Two-dimensional classification

We thus decided to move on to two-dimensional classification. Although theoretically the problems are very similar, in practice a number of technical difficulties appear. The main one is that the dimensionality of the signal space. In the one-dimensional problems, the dimension of the signal space was the length of the rows or columns, at most 700 points. In two-dimensional problems, if we want to take patches of reasonable size as our samples, say 64×64, the dimension jumps to 4096. Here the need for extracting the significant features before doing the classification is even bigger, because the classifiers such as LDA perform very badly (and need a huge computation time) in high dimensions.

We work here on the same files as before: 8354/TiΔK_n, with $\Delta K = 6, 8, 10, 12, 15$ and $n = 1, 2, 3$ corresponding to three different locations. In the following we used location 1 as training samples and location 2 as test samples, and vice-versa.

On each image, we take a certain number (128) of small square samples (64×64) as our training or test datasets. Here the patches are randomly sampled. Note that in these two-dimensional experiments, we did not remove the mean and the first moment from each sample.

5.3.1 Spatial domain

The following results were obtained using `ldb2d` with Daubechies 10 filter, extracting the 50 most discriminant features. First number of each pair is the training misclassification rate, the second one is the test misclassification rate. Results above the diagonal use file #1 for training, file #2 for testing;

Results below the diagonal use file #2 for training, file #1 for testing.

	$\Delta K = 6$	$\Delta K = 8$	$\Delta K = 10$	$\Delta K = 12$	$\Delta K = 15$
$\Delta K = 6$		18% 29%	21% 45%	21% 44%	6% 46%
$\Delta K = 8$	23% 30%		29% 59%	23% 55%	7% 44%
$\Delta K = 10$	19% 47%	26% 57%		35% 52%	8% 57%
$\Delta K = 12$	17% 46%	24% 55%	28% 54%		6% 63%
$\Delta K = 15$	7% 48%	8% 39%	6% 55%	6% 64%	

5.3.2 Fourier domain

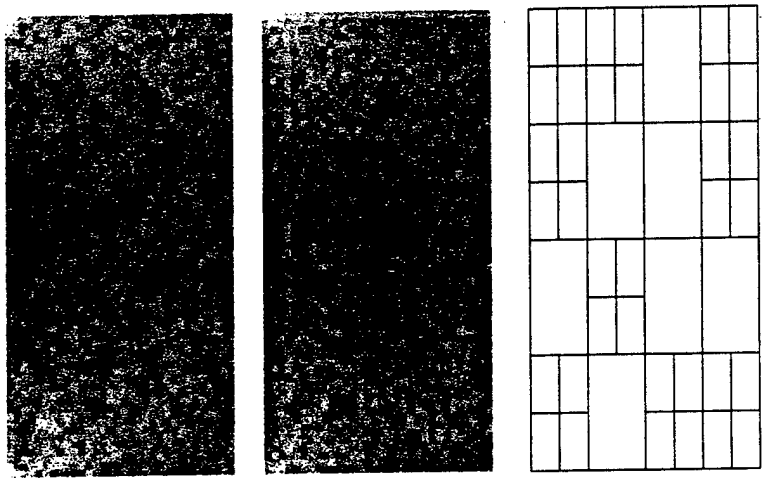
We obtained the misclassification rates using the 50 most discriminant features computed by `ldb2d` with LCT on the energy distributions of the image patches in the Fourier domain. First number of each pair is the training misclassification rate, the second one is the test misclassification rate. Results above the diagonal use file #1 for training, file #2 for testing; Results below the diagonal use file #2 for training, file #1 for testing.

	$\Delta K = 6$	$\Delta K = 8$	$\Delta K = 10$	$\Delta K = 12$	$\Delta K = 15$
$\Delta K = 6$		16% 48%	15% 48%	8% 29%	2% 36%
$\Delta K = 8$	27% 49%		16% 43%	12% 30%	4% 31%
$\Delta K = 10$	17% 50%	18% 39%		23% 44%	5% 46%
$\Delta K = 12$	11% 34%	12% 28%	23% 47%		13% 52%
$\Delta K = 15$	2% 33%	2% 30%	8% 47%	14% 55%	

The results appear to be slightly better in the Fourier domain. Figure 4 shows the typical Fourier domain energy distributions on these two classes of surfaces, and the partition patterns of the Fourier domain computed by the LDB with LCT.

5.4 Two-dimensional classification using new data

After doing the experiments described in § 5.3, we received the complete elevation data for samples 8353 and 8354. Among the different runs, ΔK



$\Delta K = 6$ (Fourier) $\Delta K = 15$ (Fourier) Localization (LCT)

Figure 4: Fourier-transformed samples (energy) and the selected LCT localization

takes a large range of values, that can be used later for the regression problem of estimating ΔK from samples. However, we still concentrate on the simpler two-class classification problem here, the classes being chosen among $\Delta K = 6, 8, 10, 12, 15$, for the purpose of comparison with the previous experiments. In each run, we take the closest values to these ΔK's:

- in the $\Delta K = 6$ class we have: 8354/Ti19870 and 8354/Ti26794 for sample 8354, and 8353/Ti12128, 8353/Ti16800, 8353/Ti17696, 8353/Ti20147, 8353/Ti23256 for sample 8353;
- in the $\Delta K = 8$ class we have: 8354/Ti28329 for sample 8354, and 8353/Ti10320 for sample 8353;
- in the $\Delta K = 10$ class we have: 8354/Ti29583 for sample 8354, and 8353/Ti25880 for sample 8353;

- in the $\Delta K = 12$ class we have: 8354/Ti30526 for sample 8354, and 8353/Ti26628 for sample 8353;
- in the $\Delta K = 15$ class we have: 8354/Ti31722 for sample 8354, and 8353/Ti31684 for sample 8353.

This time we can use the two different pieces as training data and test data, which are now completely independent. For different results, we use here `ldbkash2d` with the modified relative entropy presented in § 2.6. The test is run in the Fourier domain using LCT, and the number of features selected is 5. Results above the diagonal use 8353 for training, 8354 for testing; Results below the diagonal use 8354 for training, 8353 for testing.

	$\Delta K = 6$	$\Delta K = 8$	$\Delta K = 10$	$\Delta K = 12$	$\Delta K = 15$
$\Delta K = 6$		16% 33%	16% 33%	16% 33%	15% 41%
$\Delta K = 8$	33% 18%		44% 44%	43% 56%	42% 50%
$\Delta K = 10$	33% 16%	40% 53%		47% 54%	45% 53%
$\Delta K = 12$	32% 19%	41% 54%	44% 47%		45% 49%
$\Delta K = 15$	32% 20%	30% 58%	42% 53%	47% 49%	

With a few exceptions, the results are clearly negative (the same experiments with the classical relative entropy are neither better, nor worse). Several reasons can be imagined, but we certainly need to review our fundamental hypotheses. Most likely, the parameter ΔK is not the only one to influence the surface topography in this case, and we must take into account other factors, such as the ratio $R = K_{\text{max}}/K_{\text{min}}$, which is different on the samples 8353 ($R = 0.1$) and 8354 ($R = 0.8$). Some other evidence from SRI, using the mean Fourier power spectrum of rows in these samples, corroborate this hypothesis.

5.5 Classification of Component #39 data

We finally conducted a series of experiments on the fractographs taken on Component #39 (see page 3.3), one of the aircraft engine components that

actually failed in service. The ΔK parameter in this case is truly unknown. Since we did not develop any regression code, the best thing we could do was to apply to each fractograph the classifier constructed, as explained above, to discriminate between two known ΔK values.

Training samples were taken from the 8353 data set in a first experiment, and from both 8353 and 8354 data sets in a second experiment. We used all files corresponding to the values of ΔK previously analyzed; 6, 8, 10, 12, 15. The same parameters as in §5.4 were used for the LDB/LDA algorithm. As the test datasets, 80 samples of size 64×64 were randomly chosen in each C39 fractographs, and a majority rule was then applied in order to get a classification for the whole image.

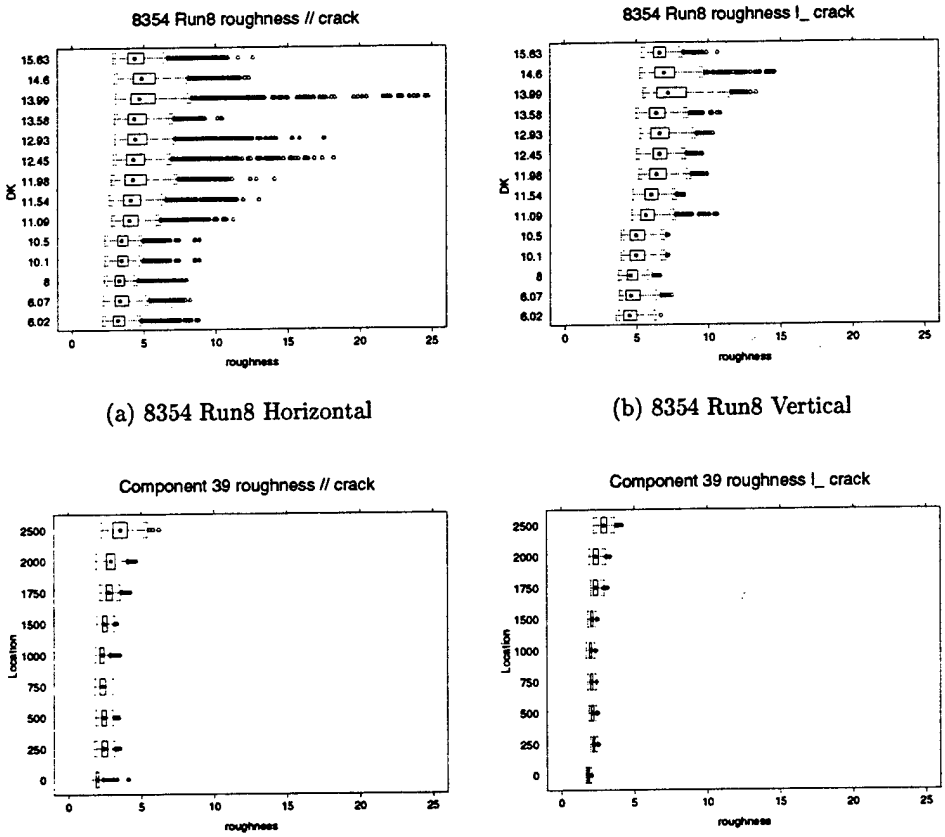
Using the classification rules for Ti6/TiX where X=8,10,12,15, all the C39 fractographs were classified as Ti6, with an overwhelming majority. In all other classification experiments, such as Ti8/Ti10, the classification results appeared random, and each case the majority vote was close to 50 %.

The results seem to indicate that the C39 fractographs were fractured under the stress conditions lower than $DK < 6$. This observation is also supported by the following experiments using the *roughness parameters*. Let $z = f(x, y)$ be a fractograph representing the elevation data. We quantified the roughness of a profile of a fractograph by

$$\begin{aligned} R_x(y) &= \int \sqrt{1 + f_x^2(x,y)}\, \mathrm{d}x \,/ \int \mathrm{d}x \\ R_y(x) &= \int \sqrt{1 + f_y^2(x,y)}\, \mathrm{d}y \,/ \int \mathrm{d}y, \end{aligned}$$

where $f_x = \partial z/\partial x$, $f_y = \partial z/\partial y$. $\int \mathrm{d}x$ represents the length of the specimen in the horizontal direction, which in fact is parallel to the crack propagation for all the available datasets. $\int \mathrm{d}y$ represents the length of the specimen in the vertical direction (perpendicular to the crack propagation). One can easily see that if the profile is perfectly flat, then $R_x = 1$ (or $R_y = 1$).

Figure 5 compares the roughness parameters of the specimens of 8354 Run 8 (and one specimen from Run 5) with those of C39 specimens.



(a) 8354 Run8 Horizontal

(b) 8354 Run8 Vertical

(c) C39 Horizontal

(d) C39 Vertical

Figure 5: Comparison of roughness parameters of profiles of fractographs. (a) 8354 Run 8 R_x; (b) 8354 Run 8 R_y; (c) C39 R_x; (d) C39 R_y. For (a) and (b), the numbers in the vertical axis represent DK of the specimen. For (c) and (d), the numbers in the vertical axis represent the location of the specimen in μm. The box in each plot represents the 50 % of the population of the roughness parameters for the given fractograph. The dot in the center of the box represents the median of the roughness parameters for the given specimen. The circles represent extremely high values such as outliers.

6 Conclusion

We summarize our study on these fractography datasets as follows:

- For the two-class classification problems, in general, the larger the difference between the ΔK values of these classes, the lower the misclassification rates. This suggests that fracture surface elevation data definitely contains the information sensitive to ΔK.
- 2D classification on the Fourier domain is more stable than 1D classification, mainly because we can assume more statistical independence on the 2D image patches, and the Fourier magnitudes are translation invariant.
- It is still difficult, however, to single out the features only sensitive to the change of ΔK from the elevation data. The local variability of the data is so severe that even the two-class classification problems is tough (e.g., the misclassification rate ranges around 30% to 40%). We need to apply the majority rule on the classification results of many samples for correct classification as described in § 5.

7 Recommendations

Based on our experience on the above experiments, we recommend that the following actions should be taken.

1. Develop *robust* feature extraction and classification algorithms.
2. Develop an algorithm to remove undesired spikes.
3. Link local features to physical fracture models.
4. Build stochastic models.
5. Integrate multiple sensors and resolutions.

We now describe each project, which can be further divided into several subprojects.

7.1 Robust Feature Extraction and Classification

From our classification experiments, we realized that it is of critical importance to extract the features that are: 1) Distinguishing the data belonging to the different classes, and 2) Common and invariant within the data belonging to the same class. It is 2) that we need to address more for any tough classification problems such as the SRI fractography data. The reason for emphasizing the common features among the same class is both for robustness of the features and for a possibility of shedding a light to a better understanding of the underlying physics: we need to extract the features insensitive to the variations of the data in the same class.

Both the original and the probabilistic versions of the local discriminant basis (let us call them LDB and LDB+, respectively) methods tried to emphasize more about the aspect 1) above, although the criteria used in LDB+ implicitly and weakly used the aspect 2). In our projects, we will more explicitly explore the aspect 2).

7.1.1 Development of LDB++

We are currently implementing and refining the discriminant measure we described in § 2.6. Let $H(p_{0|v})$, $H(p_{1|v})$ be the entropy of class 0 and class 1 samples projected on a given basis vector v in the dictionary of bases such as the local Fourier dictionary or wavelet packet dictionary. Let $S(p_{0|v}, p_{1|v})$ be the symmetric relative entropy that is used for the LDB+ method. From the discussion in § 2.6, we propose to use the following new criteria:

$$M(p_{0|v}, p_{1|v}) = S(p_{0|v}, p_{1|v}) - A \max(H(p_{0|v}), H(p_{1|v})) + B \min(H(p_{0|v}), H(p_{1|v})).$$

Another possibility is:

$$M(p_{0|v}, p_{1|v}) = S(p_{0|v}, p_{1|v}) - A|H(p_{0|v}) - H(p_{1|v})| \cdot (H(p_{0|v}) + H(p_{1|v})).$$

A and B are positive constants balancing the relative importance one may want to impose on effects. We then search the basis that maximizes the sum of $M(p_{0|v}, p_{1|v})$ from the dictionaries of bases. In the experiment we conducted in § 5.4, we only tested the one of the above formula with $A = B = 0.5$. We need to conduct more tests and research how to find optimal A and B.

7.1.2 Subtraction of common structures using the joint best basis

We should also discard the truly common and shared coordinates between the two classes for any classification tasks since they are useless for that task. To do this more explicitly, we have reached a simple idea of common structure subtraction using the joint best basis algorithm. This strategy, although very simple, must be tried for any practical classification problems.

Step 0: Expand all the signals of both classes into a dictionary of bases.

Step 1: Find a joint best basis that is efficient to compress both classes.

Step 2: Starting from the most energetic coordinates, mark the coordinates that give us low discriminant values measured by relative entropy.

Step 3: Remove such energetic but non-discriminant coordinates from both classes of signals.

Step 4: Apply the LDB, LDB+, or LDB++ for the residuals.

This strategy should really enhance the genuine difference between these two classes.

7.1.3 The local Fourier dictionary, local shifts, and image segmentation

Another major development we would like to incorporate is the local Fourier dictionary as our basic dictionary for feature extraction. The local Fourier

dictionary has many attractive properties that the wavelet packets and local cosine dictionaries cannot have: the local shift invariance, the oriented feature extraction capability, to name a few (see e.g., [16]). These properties are particularly attractive for the fracture surface classification because: 1) the corresponding features within the same regions should be aligned properly before the LDB feature extraction; 2) we expect that the surface images with larger magnifications will reveal the striation patterns locally depending on the stress conditions. To treat the feature alignment problems, we recently proposed the idea of the Best Local Translation Basis (BLTB) using the local Fourier dictionary [16]. At this point, this algorithm is designed for 1D signals. We will implement and improve this algorithm for 2D images. For the oriented feature extractions and more generally, image segmentation based on the oriented patterns, we recently launched investigation of segmentation via sparsity constraints in the local Fourier dictionary. The idea here is to fully utilize the best basis algorithm using the sparsity constraint to capture the homogeneously and isotropically oriented regions as much as possible. We expect that the regions containing more than two different orientations should be split into finer boxes. Therefore, the idea is to detect the smallest boxes in the best basis partition patterns to identify the texture boundary, and the largest boxes as a part of the homogeneous texture regions, and then merge the smaller boxes into the larger boxes if they share the same texture. These algorithms, i.e., the BLTB and the texture segmentation via sparsity constraints can be significant by-product that may be generated from this research since it is easy to foresee such algorithms can be useful for the other context that is of DARPA's interest (e.g., ATR, other pattern recognition applications).

7.1.4 Robust entropy estimation

We have also realized that the robust entropy estimation from a given set of data is critically important for all of the projects above. Currently, we are

using the histogram-based entropy estimator of Hall and Morton [4] and its variants. These entropy estimators are good if: 1) the number of available samples are large; 2) the true pdf satisfies a certain regularity conditions (e.g., no heavy tails, etc.); and 3) the underlying distribution is 1D. Therefore, if one of these conditions are violated, we need to consider alternatives. The local Fourier dictionary immediately violates the condition 3) since their expansion coefficients are complex numbers (i.e., 2D in nature). Therefore, we need good and robust entropy estimator for the local Fourier dictionary. Even for the local cosine and wavelet packet dictionaries, the condition 1) and 2) may be violated from time to time. The entropy estimators we are currently examining are kernel-based density estimator and graph-theoretical density estimator (e.g., the one based on the minimum spanning trees). With our Ph.D. student, Jen-Jen Lin, we recently obtained a method to compute entropy without using the pdf [10]. This method uses the higher order cumulants instead of the pdf's. This also works for the multidimensional distributions. We have a plan to examine its stability and compare its performance with those of the pdf-based methods.

7.2 Spike Removal by Continuous Wavelets

As we already mentioned, the fractography elevation data contains the undesirable spikes that may influence almost all frequency components in the Fourier analysis. Since the roughness of the data contains useful information, which must be preserved, conventional denoising techniques such as wavelet thresholding of Donoho and Johnstone cannot be employed; those cannot distinguish rough features from spikes. We suggest the use of a continuous wavelet transform of these data, and the properties of wavelet coefficients related to local smoothness (the so-called Hölder regularity) of the data to detect and remove the spikes.

7.3 Linking Local Features to Physical Fracture Models

This is one of the key component of our future project, and will be conducted with the fractography and material science experts. If we understand the specific fracture features responsible for producing the "characteristic fingerprint" of a given load condition in a given microstructure, this may lead to a much more realistic fracture model. That is, if we understand how the microstructure is dissipating energy at specific load conditions, the fracture mechanists and material scientists may be able to generate a model that mimics those microstructural cracking phenomena. This would, in turn, generate a model that links the wavelet/Fourier analysis to physical reality. One thing to consider in this scenario would be the effect of grain orientation, with respect to the crack propagation direction, on the produced fracture features. This could be used if/when we consider crack propagation within specific, non-isotropic, grains.

It seems to us that the newer sets of dictionaries of bases, such as local Fourier, brushlets [13], and edgelets [1], which are capable of detecting and characterizing oriented and striped patterns, may be ideal tools to extract features and link them to the physical models. This project requires real close collaboration with fracture mechanists and material scientists.

7.4 Building Stochastic Models

While the fracture mechanists and material scientists develop the relevant fracture models, we also would like to build stochastic models that are able to generate as many fracture surfaces as one wishes, which are similar to the observed fracture surfaces of a given loading condition. If we can successfully build such models for different loading conditions, then the model parameters should reflect the difference of the loading conditions. This in turn may be incorporated to the physical fracture models of Section 2.3. We have a couple

of projects in this directions.

7.4.1 Simple stochastic models using the least statistically-dependent bases

Statistical independence is one of the most desirable properties as a coordinate system in terms of redundancy reduction. If the coordinates are statistically independent, modifying one coefficient does not affect the other coefficients at all. Therefore, if we can find such a coordinate system for a given class of images, we can use this system for a variety of purposes, e.g., compression, modeling, and simulation. In reality, however, it is difficult to obtain such a coordinate system because: 1) a given image may not consist of independent components in the first place; or 2) even so, numerical algorithms to obtain such a basis become extremely complicated and expensive. Based on these observations, we very recently developed an algorithm that quickly selects a basis (from the dictionaries) that is "closest" to the statistically-independent coordinate system in the sense of relative entropy [14], [15]. The simplest stochastic model of fracture surfaces is to assume that the LSDB coordinates are statistically independent, and each fracture surface obeys the stochastic process that can be specified by the product of empirical marginal pdf's of the LSDB coordinates. We can then easily simulate and sample new images from this model by: 1) independently sampling the LSDB coefficients using the standard sampling methods such as the inversion method or the rejection method; and 2) performing the inverse transform from the LSDB coordinates to the pixel coordinates.

7.4.2 Stochastic models incorporating dependence among local features

If the LSDB coordinates do not provide us with truly or nearly independent coordinates, we will incorporate the dependence among the LSDB coordinates. We will start with a simple dependence model such as the pairwise

dependence model, then develop more elaborate models such as the Markov chains or Markov Random Fields on the local features/LSDB coordinates.

The simplest dependence model is the pairwise dependence model that approximates the true pdf by a product of bivariate pdf's. The issue of this simple model then is how to make pairs of the LSDB coordinates. We will use the mutual information to quantify the dependence between the two coordinates. Once we make the pairs of the coordinates, we can sample and simulate the coefficients using the standard simulation techniques such as the rejection method. By looking at the simulation results, we know how important the dependence structure is for the texture models. For the coordinates of complex numbers, such as the local Fourier coordinates, we can also compute the mutual information on the magnitude and the phase separately, which is extremely interesting to examine.

If the pairwise dependence is not enough to model the images faithfully, we will proceed to build more elaborate probability models such as the Markov chain on the LSDB coordinates. Here, the simplest algorithm is to start from the most energetic coordinate, then find the most dependent coordinate on the first coordinate, then find the most dependent coordinate on the second coordinate, and repeat until we exhaust the indices. If, in the middle of the process, the dependence of all the remaining coordinates on the already-selected coordinates becomes too weak, then we can break the chain at this point, and then start the new chain from the most energetic coordinate that is still left. In both the simple pairwise model and the Markov chain model, if we could reduce the dimension from n to $m \ll n$, then it would significantly speed up the both pairing procedure ($O(m(m-1)/2)$ instead of $O(n(n-1)/2)$ for checking pairs) and the actual sampling from the models. In fact, this dimension reduction even may reduce the amount of unwanted noise in the images. We will compare the performance of the reduced models with that of the full models.

7.5 Integrating Multiple Resolutions and Sensors

We should incorporate all the pertinent data available to infer the loading condition that generated the observed fracture surfaces. So far, we used only the surface elevation data measured by the FRASTAscope, with the fixed magnification. We should be able to get more information about the pertinent features by looking at different resolutions. We will also examine SEM images, which are more commonly available to many fractographers. This availability implies that a large empirical rules and diagnostics are accumulated among the experienced fractographers. Our goal here is to interview these fractographers to find out what features they are looking for to infer the loading conditions and history. The FRASTAscope can also generate the contrast/reflectivity images of fracture surfaces. These images may be used to link the surface elevation data and the SEM images.

7.6 Summary of our recommendations

The above projects may lead to a better classification performance and better understanding of the underlying physics of the fatigue fractures, which is the key of the entire project. We also believe that this project will generate significant by-product that will be useful for other area of DARPA's interest.

References

[1] A. Averbuch, R. Coifman, D. Donoho, M. Israeli, and J. Waldén. The pseudopolar FFT and its applications. Technical report, Dept. Computer Science, Yale University, 1999.

[2] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Ann. Eugenics*, 7:179–188, 1936.

[3] K. Fukunaga. *Introduction to Statistical Pattern Recognition.* Academic Press, San Diego, CA, second edition, 1990.

[4] P. Hall and S. C. Morton. On the estimation of entropy. *Ann. Inst. Statist. Math.*, 45(1):69–88, 1993.

[5] T. Kobayashi and D. A. Shockey. Fracture analysis via FRASTA, part 2: Determining fracture mechanisms and parameters. *Adv. Materials Processes*, 140(6), 1991.

[6] T. Kobayashi and D. A. Shockey. FRASTA : A new way to analyse fracture surfaces, part 1: Reconstructing crack histories. *Adv. Materials Processes*, 140(5), 1991.

[7] T. Kobayashi and D. A. Shockey. Correlation of fracture surface topography with fatigue load spectrum. In *Proceedings of the 3rd National Turbine Engine High Cycle Fatigue Conference*, fév. 1998.

[8] T. Kobayashi, D. A. Shockey, G. Ogundele, D. D. McNabb, and D. Sidey. Deducing crack history in an aged boiler tube from fracture surface topography. *J. Testing Evaluation*, 22(4):309–317, juil. 1994.

[9] T. Kobayashi, D. A. Shockey, C. G. Schmidt, and R. W. Klopp. Assessment of fatigue load spectrum from fracture surface topography. *Int. J. Fatigue*, 19:S237–S244, 1997.

[10] J.-J. Lin, N. Saito, and R. A. Levine. Edgeworth expansions of the Kullback-Leibler information. Technical report, Division of Statistics, Univ. California, Davis, 1999. submitted to J. Amer. Statist. Assoc.

[11] S. Mallat. *A wavelet tour of signal processing.* Academic Press, 1998.

[12] G. J. McLachlan. *Discriminant Analysis and Statistical Pattern Recognition.* John Wiley & Sons, New York, 1992.

[13] F. G. Meyer and R. R. Coifman. Brushlets: a tool for directional image analysis and image compression. *Appl. Comput. Harmonic Anal.*, 4:147–187, 1997.

[14] N. Saito. Least statistically-dependent basis and its application to image modeling. In A. F. Laine, M. A. Unser, and A. Aldroubi, editors, *Wavelet Applications in Signal and Image Processing VI*, volume Proc. SPIE 3458, pages 24–37, 1998. Invited paper.

[15] N. Saito. The least statistically-dependent basis and its applications. In *Proc. 32nd Asilomar Conference on Signals, Systems, and Computers*, pages 732–736. IEEE, 1998.

[16] N. Saito. The local Fourier dictionary: a natural tool for data analysis. In M. A. Unser, A. Aldroubi, and A. F. Laine, editors, *Wavelet Applications in Signal and Image Processing VII*, volume Proc. SPIE 3814, pages 610–624, 1999. Invited paper.

[17] N. Saito and R. R. Coifman. Improved local discriminant bases using empirical probability density estimation. In *Amer. Statist. Assoc. 1996 Proceedings of the Statistical Computing Section*, pages 312–321.

[18] N. Saito and R. R. Coifman. Local discriminant bases and their applications. *J. Mathematical Imaging and Vision*, 5(4):337–358, 1995. Invited paper.

[19] C. G. Schmidt, T. Kobayashi, D. A. Shockey, and T. H. Flournoy. Correlation of fracture surface topography with fatigue load history. In *Proceedings of the ICAF '97*, juin 1997.

[20] W. N. Venables and B. D. Ripley. *Modern applied statistics with S-Plus*. Statistics and Computing. Springer-Verlag, 1994.

APPENDIX B

ANALYSIS OF A FAILED SPACE SHUTTLE COMPONENT

APPENDIX B

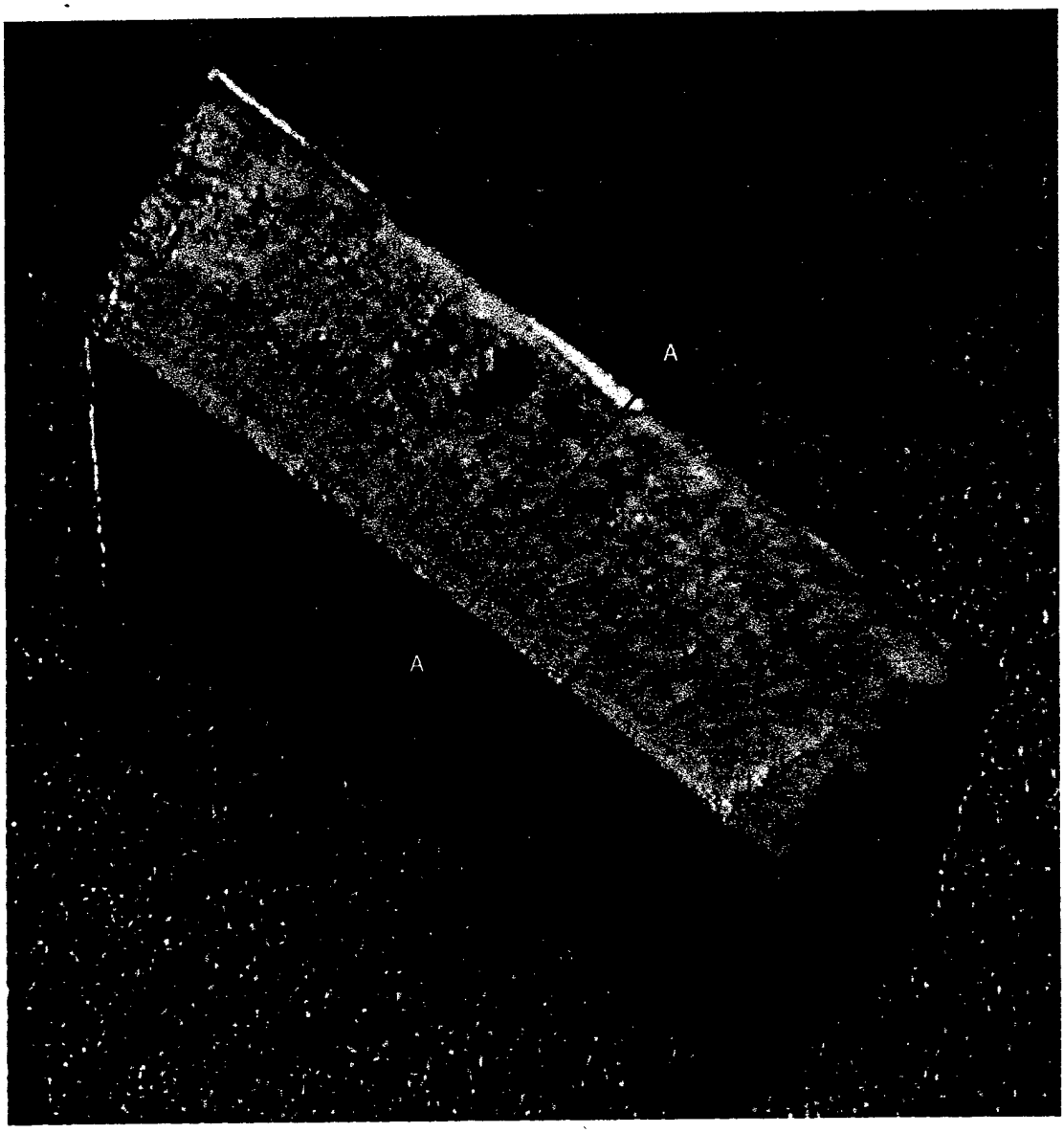
ANALYSIS OF A FAILED SPACE SHUTTLE COMPONENT

During testing of a high pressure pump in the space shuttle fuel tank, a crack about 6 inches long and 0.290 inch deep was found in a 0.070-inch blend radius of the turbine housing after 32 starts. A failure analysis team from Pratt & Whitney, Rocketdyne, and NASA Lewis looking into the cause and evolution of this crack wanted to know when in the test history the crack initiated, if the crack was unstable in the late stages of development, and whether propagation resulted from engine start/stop cycles or power level changes. To answer these questions, they hoped to identify and interpret features on the crack surfaces.

The housing, which functions as a pressure vessel for high temperature (up to 500°F) hydrogen and steam, is made of a wrought nickel-base alloy with a maximum average grain diameter of 9 μm. Chemical analysis showed a higher concentration of hydrogen near the surface than in the bulk of the material, and metallographic sections through the crack showed higher hydrogen concentration along the crack and at the crack tip. Engineering estimates of the strains indicated levels sufficiently beyond yield that linear elastic fracture mechanics principles would not apply. Several elastic-plastic models predicted the maximum stress intensity at the tip of the crack when it was 0.075 inch long to be 43 to 77 ksi$\sqrt{\text{in.}}$ A K_{max} of 77 ksi$\sqrt{\text{in.}}$ should produce crack growth in hydrogen.

Fractographic features were not distinctive enough to characterize the stress state, but macroscopically there were large changes in the surface texture. The team therefore sought other methods to analyze the fracture surface topography. A small section of one fracture surface (Section 6B from Unit 9-1, shown in Figure B-1) was sent to SRI for analysis.

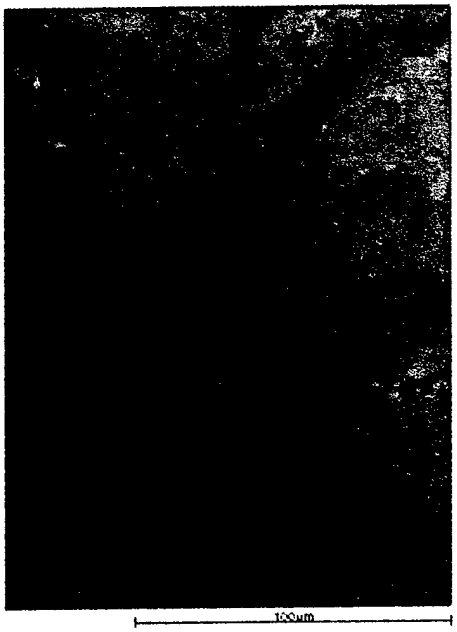
SRI used the FRASTAscope to produce topographic maps of 29 areas of the fracture surface along the line A-A. Each area was 139.8 μm in the direction of crack growth (i.e., in the through-thickness direction) and 177.0 μm in the perpendicular direction. This area was divided into á 600 by 800 pixel array and, hence, the data spacing was 0.233 μm in both directions. Figure B-2 shows contrast images and gray-scale topography images of selected areas.



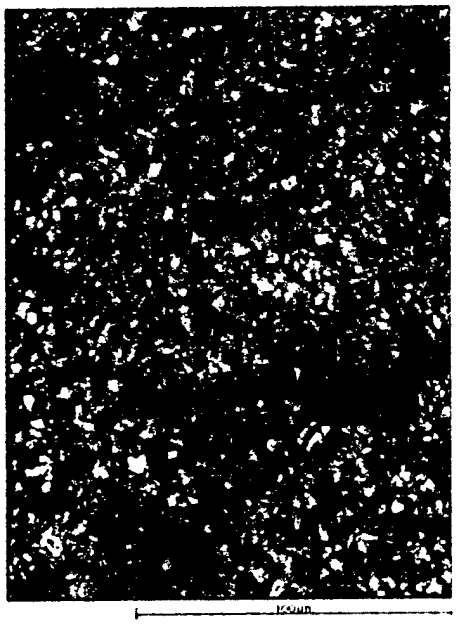
NM-312581-12

Figure B-1. Fracture surface from a Space Shuttle turbine housing. The topographies of areas along the line A-A were characterized and analyzed by FFT.

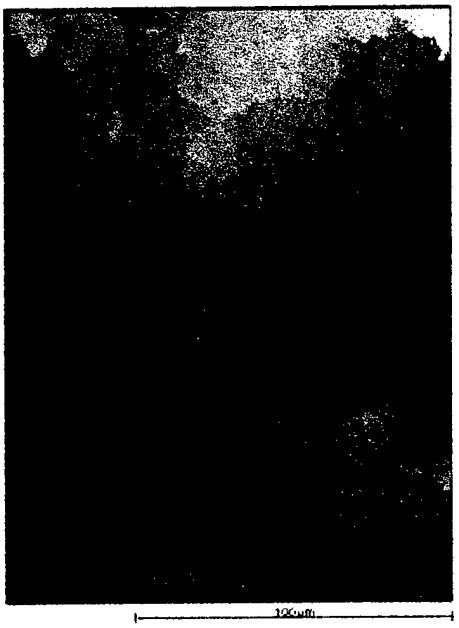
(a) Contrast image of the area at 69.3 μm from the inner edge



(b) Gray-scale topography image of the area at 69.3 μm from the inner edge



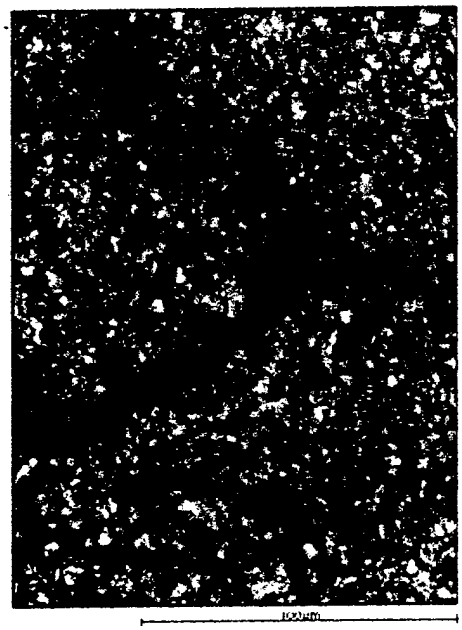
(c) Contrast image of the area at 2379 μm from the inner edge



(d) Gray-scale topography image of the area at 2379 μm from the inner edge

NM-312581-13

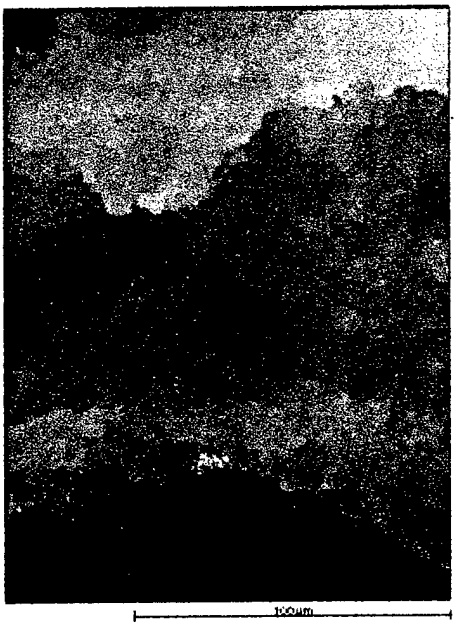
Figure B-2. A series of contrast and gray-scale topography images of selected areas along the line A-A in Figure B-1.



(e) Contrast image of the area at 2469 μm from the inner edge

(f) Gray-scale topography image of the area at 2469 μm from the inner edge

(g) Contrast image of the area at 2869 μm from the inner edge



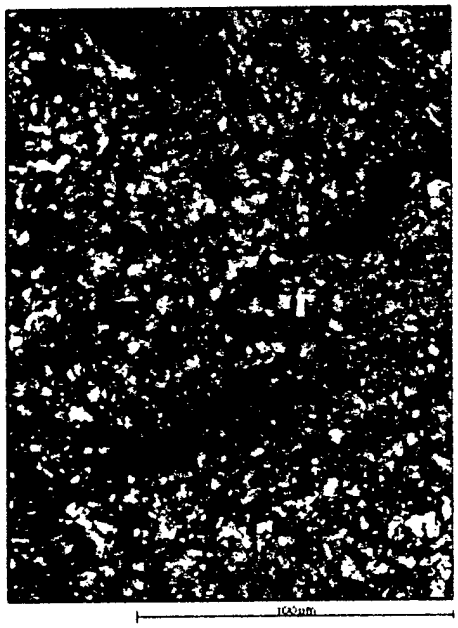
(h) Gray-scale topography image of the area at 2869 μm from the inner edge

NM-312581-14

Figure B-2. A series of contrast and gray-scale topography images of selected areas along the line A-A in Figure B-1. (Continued)

(i) Contrast image of the area at 4269 μm from the inner edge

(j) Gray-scale topography image of the area at 4269 μm from the inner edge



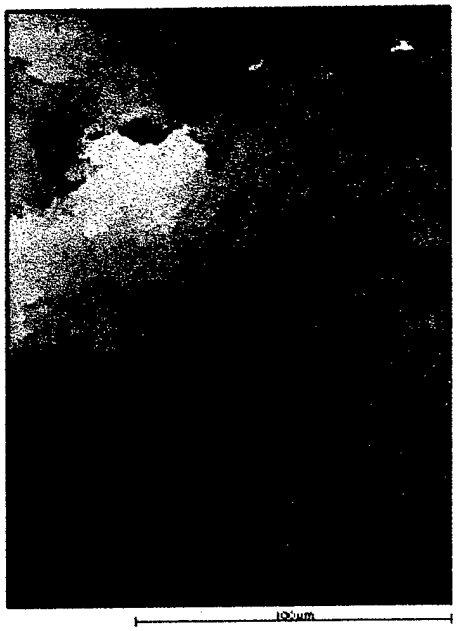
(k) Contrast image of the area at 4469 μm from the inner edge

(l) Gray-scale topography image of the area at 4469 μm from the inner edge

NM-312581-15

Figure B-2. A series of contrast and gray-scale topography images of selected areas along the line A-A in Figure B-1. (Continued)

(m) Contrast image of the area at 4659 μm from the inner edge



(n) Gray-scale topography image of the area at 4659 μm from the inner edge

NM-312581-16

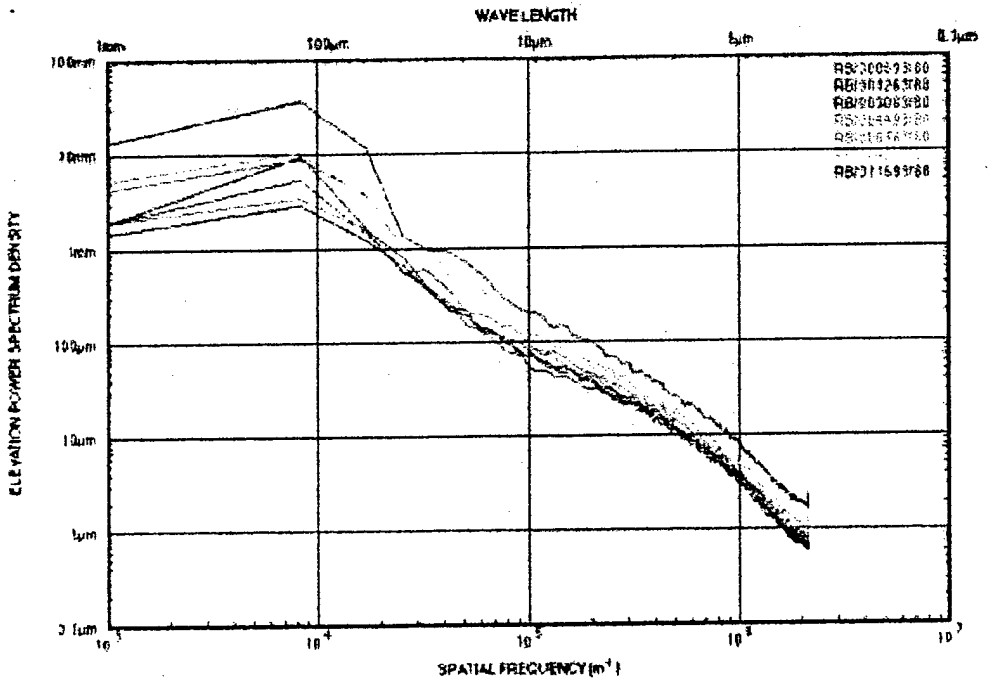
Figure B-2. A series of contrast and gray-scale topography images of selected areas along the line A-A in Figure B-1. (Concluded)

A fast Fourier transform analysis was applied to the elevation files, and the averaged elevation power spectrum density (EPSD) distribution was computed for each area. The results are plotted in Figures B-3(a), (b), and (c), where the parameter is the distance from the centroid of the area to the edge of the housing. Between the wavelengths of 0.5 and 10 μm, the EPSD curves are roughly parallel to one another and slightly concave. To examine the variation of surface roughness with crack depth, we plotted the value of the ESPD at the arbitrarily chosen wavelength of 5 μm versus distance from the housing edge, Figure B-4(a). A plot of the square root of the EPSD versus crack depth is shown in Figure B-4(b). These figures show two large peaks in fracture surface roughness.

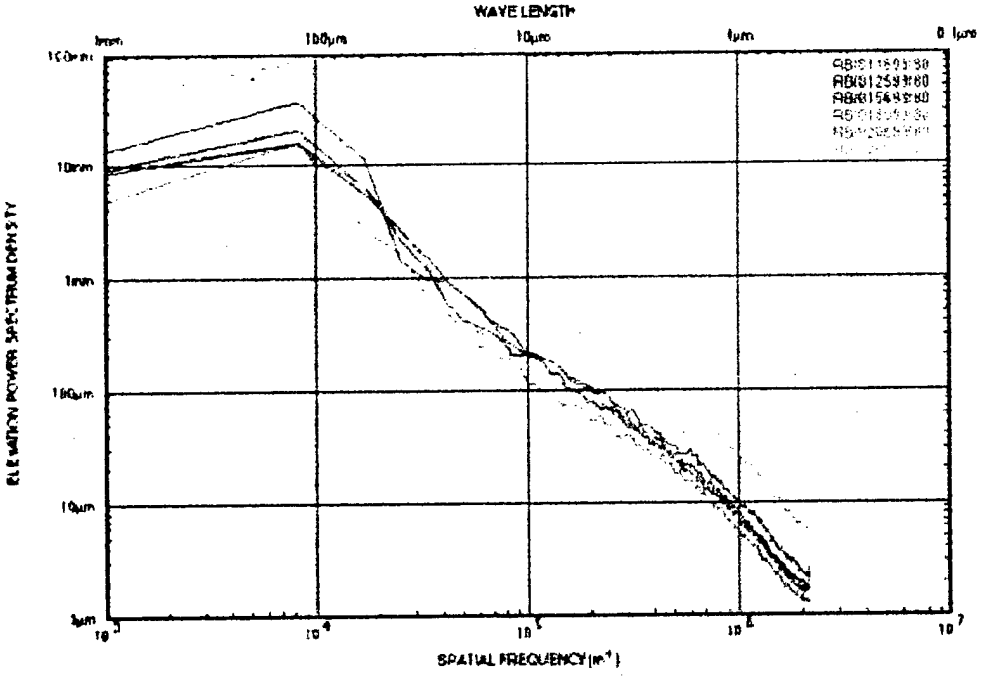
In light of the results presented in the body of this report, these peaks suggest loading excursions at crack depths of 2.5 mm and 4.5 mm. Because reference data for this nickel-base alloy have not been generated, we are unable to convert the EPSD values to quantitative estimates of K_{max} or ΔK. However, the slight concavity in the EPSD curves is reminiscent of the shape of the curves for the titanium alloy reported in the body of this report. Therefore, one might speculate that the R value ($\sigma_{min}/\sigma_{max}$) is closer to 0.1 than to 0.8.

Several smaller peaks are also evident in Figure B-4 and suggest smaller overloads. A similar analysis performed on center-cracked panels of aluminum alloy sheet showed similar EPSD peaks at crack lengths where overloads were applied.[9] Greater peak heights were obtained with larger overloads.

It would be of great interest to compare the results of Figure B-4 with the run time history (starts/stops and power changes versus time) of the engine.



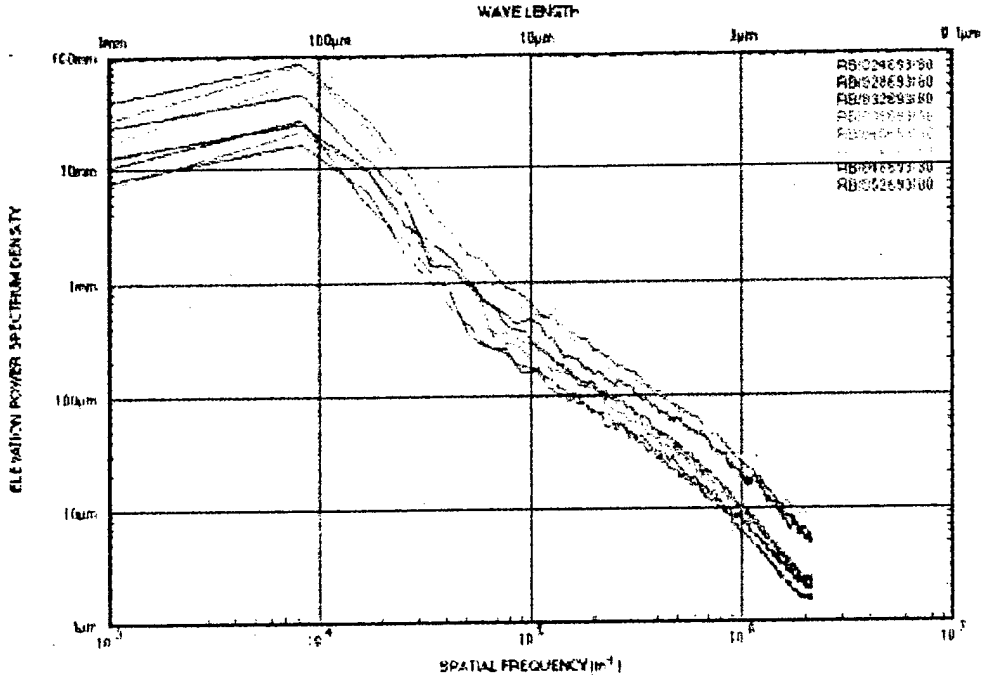
(a) Elevation Power Spectrum Density Plot of the First Seven Areas along the Line A-A



(b) Elevation Power Spectrum Density Plot of the Middle Six Areas along the Line A-A

NM-312581-17

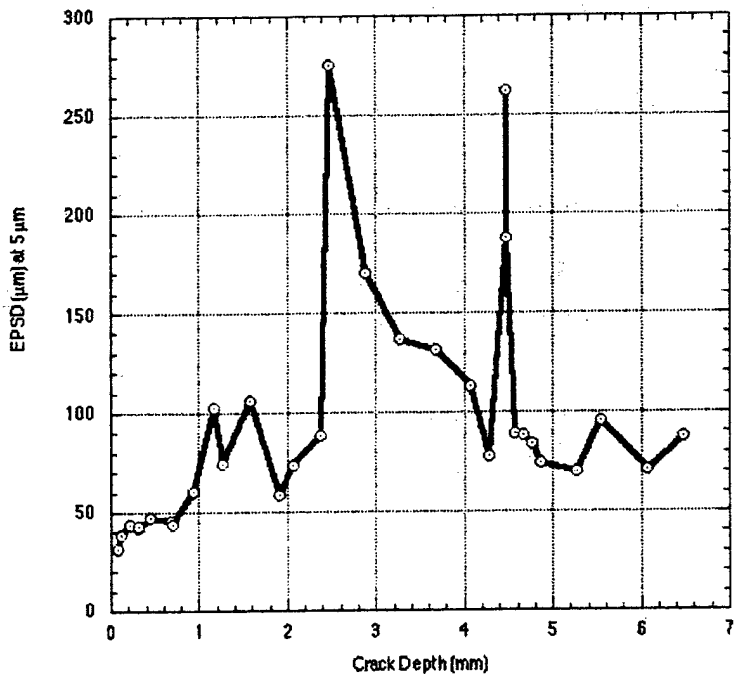
Figure B-3. Elevation power spectrum density plots of the areas along the line A-A.



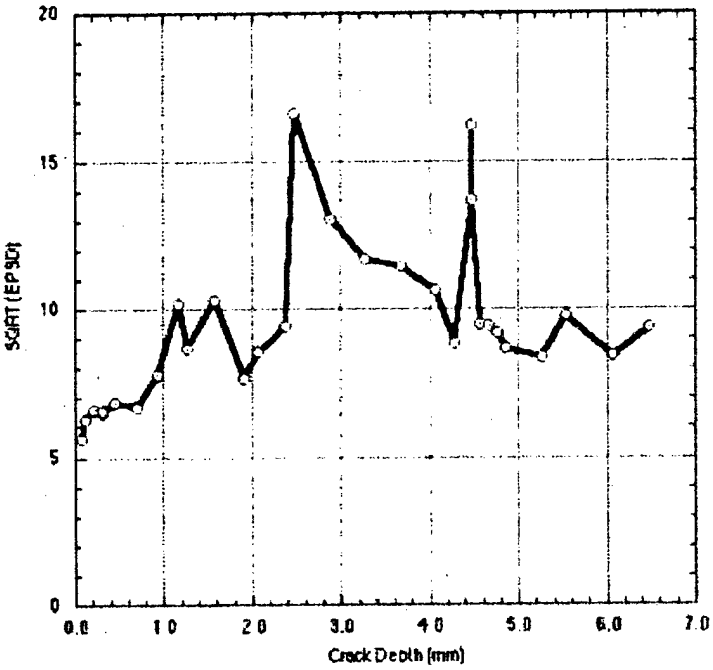
(c) Elevation Power Spectum Density Plot for the Last Eight Areas along the Line A-A

NM-312581-18

Figure B-3. Elevation power spectrum density plots of the areas along the line A-A. (concluded)



(a) EPSD values as a function of crack depth



(b) SQRT(EPSD) values as a function of crack depth

NM-312581-19

Figure B-4. Variation of EPSD and SQRT(EPSD) as a function of crack depth.